Heterogeneity and Persistence in Returns to Wealth*

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

First Version: December 2015 This version: August 2016

Abstract: Heterogeneity, persistence, and correlation with wealth of individuals' returns to savings are key to explain wealth concentration at the top. We provide the first systematic analysis of the properties of individual returns to wealth using twenty years of population data for Norway. The data consist of information on income from capital and asset values from 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 range of 500 basis points between the 10th and the 90th percentile. Second, heterogeneity in returns does not arise only from differences in the assets allocation between safe and risky assets. Indeed, returns are heterogeneous even within asset classes. Third, returns are positively correlated with wealth. Fourth, returns have an individual permanent component that explains almost 20% of the variation. Fifth, the individual permanent component accounts for the bulk of the correlation between returns an wealth for wealth below the 90th percentile; the correlation at the top reflects both compensation for risk and the correlation with the individual permanent component. Sixth, the permanent component of the return to wealth is also (mildly) correlated across generations. Finally, there is assortative mating in returns. We discuss the implications of these findings for the debate about the drivers of wealth concentration, its measurement, and the relation between income and wealth inequality.

Keywords: Wealth inequality, returns to wealth, heterogeneity, intergenerational mobility, assortative mating.

JEL codes: E13, E21, E24

*We thank Alberto Bisin, Fatih Guvenen, Gueorgui Kambourov, Daniele Massacci, Andrea Pozzi and Anton Tsoy for very fruitful discussions. We are grateful for useful comments from seminar participants at the 2016 ASSA meetings, EIEF, European University Institute, University of Chicago, UC Berkeley, Fudan University, HEC Montreal, UC Santa Cruz, Northwestern, Kellogg, NY FED, the 2016 QSPS meeting at Utah State, and the 2016 Cowles Macro/Labor conference at Yale.

[†]Statistics Norway, faa@ssb.no

[‡]Einaudi Institute for Economics and Finance (EIEF) and CEPR, guiso@tin.it

[§]Stanford University, davidem@stanford.edu

[¶]Stanford University and NBER, pista@stanford.edu

1 Introduction

Over time and across countries, the wealth distribution appears 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 U.S., for example, the top 0.1% hold about 20% of the economy's net worth. Moreover, tail inequality has tripled in little more than three decades (Saez and Zucman, 2016). These striking regularities led Vilfredo Pareto to introduce a statistical distribution, now bearing his name, to model long-tailed economic phenomena and to theorize about possible socio-economic factors that might generate them.

What produces the long tail of the wealth distribution and its extreme skewness is a topic of intense research. We review this vast literature in Section 2. One strand of literature, started by Aiyagari (1994) focused on the role played by idiosyncratic and uninsurable labor income (i.e., human capital) risk, leading households to accumulate assets for precautionary reasons (see Castaneda, Dıaz-Giménez, and Rıos-Rull (1998); Huggett (1996)). Several authors explored various additional mechanisms, such as non-homothetic preferences for bequests, heterogeneity in entrepreneurial talent, extreme skewness in the distribution of earnings for top earners, and heterogeneity in discount rates (e.g. Cagetti and De Nardi (2006); Castaneda et al. (2003); Kindermann and Krueger (2014); Krusell and Smith (1998); Quadrini (1999, 2000)). The success of these models in reproducing the amount of wealth concentration observed in the data is mixed. More recently, a wave of papers have shifted attention from heterogeneity in returns to labor to heterogeneity in returns to 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. *Type dependence* in the growth rate of wealth, i.e., a positive correlation between wealth and returns to wealth, can potentially explain the fast increase in tail inequality observed in the last three decades.

There is, however, scant evidence regarding 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) designed to explain the fast increase in tail inequality? More generally, what are the empirical properties of the returns to wealth? Data limitations have made answering these questions difficult: available survey data are plagued with measurement error, low response rates at the top of the wealth distribution, and limited to no longitudinal information. The goal of our paper is to fill this gap.

We study properties of returns to wealth using administrative tax records on capital income and wealth stocks for all taxpayers in Norway over two decades. Several properties of these data make them unique to address the questions above. First, measurement error and underreporting of wealth information are unlikely to be a problem because wealth data are generally collected through third parties (i.e., information provided by financial intermediaries) rather than being self-reported. Furthermore, the data have universal coverage, implying that there is exhaustive information on the assets owned by *all* individuals, including those at the very top of the wealth distribution. This is critical for a study of our sort, as leaving out the wealthy could seriously understate the extent of heterogeneity in returns to wealth, particularly if returns are correlated with wealth and if the extent of heterogeneity also varies with wealth. Most importantly, the data have an extraordinary long panel dimension, covering 20 years – from 1993 to 2013 – and various business cycles. This allows us to study within-person persistence in returns. In addition, because over a 20-year period (some) generations overlap and because we can identify parents and children, one can also study intergenerational persistence in returns to wealth. Finally, since we observe individuals before they marry, we can study whether returns to wealth persist across marital statuses and whether this reflects assortative mating on returns to wealth.

We find that returns to wealth exhibit substantial heterogeneity. For example, in the last year for which we have data (2013) the (value-weighted) average return on overall wealth is 3.7%, but it varies considerably across households (a standard deviation of 6.1%). There are also large differences between safe and risky asset returns. In 2013 the (value-weighted) average return to risky assets is 4.8%, almost double the return on safe assets (2.6%). However, the standard deviation of the former (11.7%) is one order of magnitude larger than the standard deviation of the latter (1.6%). Hence, we find returns heterogeneity even when we focus on safe assets, although the dominant source of returns heterogeneity originates from risky assets. Furthermore, heterogeneity in returns is not simply the reflection of differences in portfolio allocations between risky and safe assets and thus compensation for risk-taking mirroring heterogeneity in risk tolerance. Even conditioning on the share of risky assets in 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 point. The correlation between returns and wealth does not reflect merely risk-taking. We find that risk-adjusted measures of excess returns (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 idiosyncratic transitory variations as well as from a persistent component in returns to wealth. To separate the two components we estimate a panel data statistical model for the returns to wealth that includes an individual fixed effect. To account for the heterogeneity that is explained by observable factors, we control for time, demographics, the level of wealth (capturing investment-size effects on returns), and the share of wealth in risky assets (capturing compensation for risk). The individual fixed effect measures the component of unobserved heterogeneity that persists over time. Finally, there is a component of heterogeneity that is unobserved but unsystematic (good/bad luck, etc.). We find that observable characteristics, alone, explain roughly 12% of the variability in returns to wealth. Adding individual fixed effects triples 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-10th percentile difference of 6.4 percentage points. Interestingly, dispersion in the distribution of returns fixed effects characterizes not only the population but also its subgroups: the distribution is more spread out and shifted to the right for people at the top of the wealth distribution compared to people at the bottom; for business-owners compared to non-owners; and for people with higher levels of economics/business education.

Having established the presence of large and systematic heterogeneity in asset returns across individuals, we turn to analyzing intergenerational and intramarital 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 the parent's asset return, while positive for a good 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.

We also find evidence of assortative mating in returns *conditional* on assortative mating in wealth. High-return singles tend to marry individuals who also earn above-average returns. Moreover, we find that post-marriage household returns mostly resemble the pre-marriage return of the highest-return spouse. However, the lowest-return spouse also plays a role, providing a rationale for assortative mating on returns to wealth. Strikingly, we document that the weight played by the highest-return spouse is higher if that spouse is the male.

As far as we know, this is the first paper that provides systematic evidence on individual returns to wealth over the entire wealth distribution, characterizes their properties, documenting the extent of cross sectional heterogeneity (both observed and unobserved), its correlation with the level of wealth, and its long-run persistence (within person, across generations, and across marital status). Bach et al. (2015) perform an exercise similar 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 several 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 returns persistence. Third, we observe all components of financial wealth, including shares of private businesses which is the dominant source of wealth for the very top quantiles of the wealth distribution. Fourth, we can study heterogeneity and persistence in returns to wealth over and above the intra-generational dimension they focus on. Indeed, our paper is the first to provide systematic evidence on persistence in returns within and across generations as well as across marital status. The first two features are critical for explaining the long thick tail in the wealth distribution. Bisin et al. (2015) is the only paper we know of that estimates the extent of persistent heterogeneity and intergenerational persistence in returns to wealth. But estimates are obtained from structural estimation of a life cycle model of wealth accumulation, calibrated on U.S. data. We instead provide direct evidence from observed returns. 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 reasonable calibrated models of wealth inequality should account for. We also find that heterogeneity in returns varies over time. While heterogeneity in returns matters for explaining the level of wealth inequality at the top, variation over time in heterogeneity may matter for explaining 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 that 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)). None of the theoretical papers on wealth inequality has studied the implications of assortative mating in returns to wealth for wealth inequality and mobility, probably because while assortative mating on income is widely documented, our is the first study to show that people sort not only on wealth but also on returns to wealth.

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 on 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 additional dimensions of persistence: intergenerationally and across spouses. Section 8 concludes discussing some implications of our findings.

2 Heterogeneity in returns and the distribution of wealth

Absent 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 that feature. Yet, wealth seems to be uniformly 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) gets close to matching the distribution of wealth in the US, it requires that the top 0.25% of income earners to earn between 400 and 600 times more than the median earner. As Benhabib and Bisin (2016) notice, 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 thus their heterogeneity difficult to assess. Hence, one has 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 - e.g. Heaton and Lucas (2000); Vissing-Jorgensen and Moskowitz (2002)), rather than with higher than average discount rates. An alternative route followed in the attempt to match the thick tail in the distribution of wealth has been to allow explicitly for entrepreneurship and idiosyncratic returns to investment, as in Quadrini (2000) and Cagetti and De Nardi (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 allows to match 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.

While idiosyncratic returns from entrepreneurship are one source of heterogeneity in returns to wealth that can help explain wealth concentration, heterogeneity in returns to wealth can arise from other sources. For example, Guvenen (2009) introduces return differentials by allowing all households to trade in a risk-free bond, but restricting 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]

The heterogeneous stochastic returns approach to explain wealth concentration at the top has been more recently systematically developed and sharpened by Benhabib, Bisin and various coauthors in a sequence 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 returns to wealth. Each household is endowed at birth with a rate of return on 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.

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

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 (as shown in their theorem 1). The latter is instead increasing in the degree of heterogeneity in asset returns. And because the wealthy are on average those endowed with a high rate of return, their model generates endogenously 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 of the drivers of wealth inequality highlighting the specific role of returns heterogeneity. To assess quantitatively 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), compared to other factors, they calibrate their overlapping generations model to U.S. 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 U.S. 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 but modest. Yet, even this relatively small amount of persistent heterogeneity is able to play 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 a much smaller top 1% wealth share of 5.7%, a top 5% wealth share of 9.5% (vis-à-vis an observed level of 60.3%), and wealth shares at the bottom of the distribution that are abnormally inflated. Hence, returns heterogeneity appears a key factor for matching the empirical wealth distribution. Interestingly, as we will see the extent of heterogeneity and of its intergenerational persistence that we estimate is remarkably similar to the one calibrated by Benhabib et al. (2015).

Gabaix, Lasry, Lions, and Moll (2015) are interested not only in the amount of wealth concentration in the steady state, but also on the speed of the transition across steady states.

They show that while the Benhabib et al.'s 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 *type dependence* in the growth rate of wealth, i.e., high-wealth individuals have faster random growth rates of wealth than low-wealth individuals. Since the growth rate of wealth coincides with the return to wealth (absent saving or borrowing), the Gabaix, Lasry, Lions, and Moll (2015)'s model requires that returns to wealth are positively correlated with the level of wealth.

Despite the theoretical appeal, explanations for 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 problems of the models that rely on heterogeneity in discount rates. How reasonable is the heterogeneity and persistence estimated 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? Differently from individual discount rates, however, individual returns on wealth have the great advantage that they can be observed (though not easily). 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. We start by using a rich longitudinal database that covers every Norwegian resident from 1967 to 2013. For each year, the database provides relevant demographic information (gender, age, marital status, educational attainment) and geographical identifiers. For the period 1993-2013 - the one we focus on here - we can link this database with tax records containing individual information on asset holdings and liabilities (such as real estate, financial assets, private businesses, 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 are measured as of December 31 of each year. While tax records typically include information on income, they rarely (if ever) contain information on wealth. In Norway's case, this happens because of a wealth tax mandating taxpayers to report in their tax filing not only their incomes but also their asset holdings.

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 the 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.[2] In particular, steady-state wealth inequality is likely to be sensitive to even 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. Because of this, the data do not suffer from the standard measurement errors that characterizes household surveys, where individuals self-report income and assets components (as for instance in the U.S. Survey of Consumer Finances) and confidentiality issues censor extreme asset holdings above certain thresholds. Third, the Norwegian data have a very long panel dimension: this is indispensable to identify persistent heterogeneity in returns which is the central goal of this paper. And because the data cover the whole relevant population, they are free from attrition, except the unavoidable ones arising from mortality and out-of-country migration. Fourth, unique identifiers allow us to match parents with their children. This feature, together with the long panel dimension of the data, is key to study intergenerational persistence in returns to wealth, which may be an important determinant of wealth inequality in the tail (Benhabib et al. (2011); Bisin et al. (2015)). Fifth, since we observe individuals before they marry or form joint tax units, we can study whether returns to wealth persist across marital statuses and whether there is assortative mating on returns to wealth. Coupled with the documented tendency of the wealthy to marry similarly wealthy partners, this can be an important and so far unstudied mechanism of persistence in wealth inequality. 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 have also some shortcomings: one, not surprising, is the measurement of the value of private

[2] Wealth concentration in Norway is high. In 2012, the top 0.1% owned about 10% of all net worth in the economy. For comparison, in the U.S. the top 0.1% owned about 22% of all the net worth in the economy.

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.[3] Below we briefly describe the administrative tax records on 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 in the Internet Appendix.

3.1 Administrative wealth and capital income records

Norwegian households are subject to both an income tax and a wealth tax. Each year they are required to report their incomes and complete information on wealth holdings to the tax authorities. Tax record data are available on an annual basis since 1993.[4] The collection of tax information is mostly through third-parties. In particular, employers must send both to the individual and to the tax authorities information on earned labor income; 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. For individuals who own no stocks, the tax authority pre-fills a tax form and sends it to them for approval. If the individual does not reply, the tax authority considers the information it has gathered as approved. In 2009, nearly 2 million individuals (60 percent of the Norwegian tax payers) belonged to this category. Individuals who own stocks have to fill

[3] The main components of wealth that are left out of our analysis are housing and private pension wealth (and their related returns). Contributions to private pension funds are capped to USD 1,500 annually and not subject to the wealth tax (and hence do not appear in the tax records). However, 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 reason. 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 most likely overstating 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 2010-13 (when housing data is complete and accurate), the correlation between financial wealth and total wealth (financial wealth + housing wealth - debt) ranges between 0.98 and 0.99.

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

in the tax statement - including calculations of capital gains/losses and deduction claims. The statement is sent back to the tax authority which, as in the previous case, receives all the basic information from employers and intermediaries and can thus check its truthfulness and accuracy. Stockholders are treated differently because the government wants to save on the time necessary to fill in more complex tax statements. The fact that financial institutions supply information on their customer's financial assets directly to the tax authority greatly reduces the scope for tax evasion, and thus non-reporting or under-reporting of assets holdings is likely to be negligible.[5]

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 with 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 (held directly or indirectly through mutual funds, $w^{m,l}$) 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 with the wealth tax - what we label the "assessed" value. This value does not necessarily correspond to the "market" value of those shares - i.e., the realization price if they were to be sold in the market. The value of unlisted stocks is obtained as the product of the equity share held in the firm and the value of the company as reported in the company's tax returns. The latter, in turn, excludes net present value calculation of the firm or goodwill. 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 underreport their value. Consistent with this, the (log) assessed value is strongly correlated with the firm (log) book value (correlation 0.88, Figure IA.6) and in more than 50% of the cases the assessed value exceeds the books value (which would be inconsistent with the goal of minimizing the tax bill). Medium- to

[5] 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.

large-size firms (with turnover above NOK 5 millions, or USD 500k) are mandated 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}). Because dividends and capital gains/losses on listed and private firms are, for tax purposes, reported jointly, 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 median and below the top 1%. The share in private business is strongly increasing with wealth above the 95th percentile and carries a very large weight, close to 90%, for the top 0.01%. Because returns to private businesses are largely idiosyncratic, given the strong correlation between exposure to private business and wealth, lack of information on private business holdings is likely to grossly understate returns heterogeneity.

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.[6]

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, 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). The 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 as the ratio of income from capital and the average stock of wealth at the beginning *and* end of year, i.e.:[7]

[6] One worry is that this measure (and the ones presented below) may induce spurious positive correlation between returns and wealth if the wealthier exhibit higher propensity to save out of wealth. In particular, suppose that y_{it} is the sum of capital income out of initial wealth ($\sum_j r_t w_{it}^j$) 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 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 is likely to be *downward*.

[7] To see the importance of this adjustment, consider an individual who has beginning-of-period wealth $w_{it} = 0$ and after six months invests USD 100 in a money market account at 10% interest rate, earning $y_{it} = 5$. End-of-period wealth is thus $w_{it+1} = 105$. The naive measure of return (1) is undefined. The adjusted return measure is instead r_{it}^A=9.52%, much closer to the actual 10% return. The adjusted measure works well also when people withdraw for consumption purposes. Consider an individual who has beginning-of-period wealth $w_{it} = 100$ invested in a 10% money market account. After 9 months, the individual withdraws and spends USD 50, so that capital income is $y_{it} = 8.75$. End-of-period wealth is $w_{it+1} = 58.75$. The naive measure of return (1) would be lower than the actual one, 8.75%. The adjusted return measure is, instead, r_{it}^A=11%, much closer to the actual 10% return.

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

We use this adjustment both when we compute the returns on safe assets, $r_{it}^{s,A} = \frac{i_{it}}{(w_{it}^{s}+w_{it+1}^{s})/2}$, as well as when we measure returns on risky assets, $r_{it}^{m,A} = \frac{d_{it}+g_{it}}{(w_{it}^{m}+w_{it+1}^{m})/2}$. The 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.[8] 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 qualitatively different 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 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 relatively to what they would get if they were to sell it. There is no simple way to correct for this problem. To show that none of our results depend on private equity, we consider an alternative measure that excludes private equity owners, and is hence defined as:

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

[8] 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 Christmas 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=50%. This overstatement is less likely to happen for large accounts.

The third limitation of our data is that we observe capital gains/losses when they are realized, rather than when 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 we observe enough realizations of the capital gains.[9]

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 individual accrued capital gain/loss from stockholding 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 the return to total wealth (2) – our baseline measure, which has the advantage of being based on information directly available from the tax records. In Section 5.3 we show that our main findings are not sensitive to adopting the alternative measures of returns (3) and (4).

All returns statistics we report are at the individual, not household level. This way we account for the fact that while households form and dissolve, individuals can be observed as

[9] The average return over a holding period of T years of an asset that is sold at T is the same whether the average return is computed using the annual return $R(t) = \frac{y_t}{P_t} + \frac{P_{t+1}}{P_t}$, with capital gains computed on an accrual basis, or when the annual 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. On the other hand, if the measured return on a risky asset for individual i is $R_i(t) = \frac{y_t}{P_t} + I_i \frac{P_{t+1}}{P_1}$ where $I_i = 1$ if i sells the asset at t (and we observe the capital gain) and zero otherwise, then clearly at each point in time this induces some cross sectional heterogeneity in measured returns, because in each year only a fraction of the individuals realizes the capital gain. In the Internet Appendix we show that this type of heterogeneity is nevertheless contained and can only explain a small fraction of the heterogeneity in returns to risky assets that we measure in each year.

they cycle through different marital arrangements. When individuals are single, the formulae above apply without modifications. When individuals are married, we assume that spouses share household wealth and capital income equally. This is consistent with Norwegian laws requiring that family assets are split equally between spouses in case of divorce. In this case, we first compute the return to household wealth, and then assign to each spouse this return and the per-capita household wealth.

3.4 Descriptive statistics

Table **??** shows summary statistics for our data. For 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 marital status (50% are married). About 80% of 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 digs into statistics describing wealth levels and composition. In 2013, 45% of the Norwegian households had some risky assets in their portfolio. One in nine owned shares of 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 **??** 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%) exceeded substantially 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. Similarly, the average returns from risky assets are 3.5% and 6.9% in the two cases. The larger difference in the value-weighted case is explained, as we shall see, 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 to model returns to wealth that will help shed light on these stylized facts.

4.1 Returns to wealth are heterogenous

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. The average return on wealth using the baseline measure is 3% with a standard deviation of 5% (Table **??**, panel C).[10] 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 on poorly-performing assets. Among risky assets holders, returns are more heterogeneous.

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 π_{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:

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

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

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

[10] The coefficient of variation in the Norwegian case is thus larger than that calibrated with U.S. data by Benhabib et al. (2015), who find an average return of 3.4% with a cross sectional standard deviation of 2.7%. However, the calibration by Benhabib et al. (2015) refers to average individual returns over the lifecycle. We will discuss measures that are comparable to the ones they report in Section 5.

Heterogeneity in returns is induced by differences in risk aversion and thus in (compensated) risk taking measured by the risky share.[11] Equation (5) suggests that conditioning on having the *same* share of risky assets in portfolio, total returns on wealth should be similar across investors. That is, the cross sectional standard deviation of returns, given π_{it}, should be close to zero. In Figure 3 we use again 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 $\pi_{it} = 0$ (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 formal controlled regression, discussed in Section 5). Note that some of the heterogeneity in Figure 3 may come from holdings of a private business with very idiosyncratic returns. 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 of 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 the holding of private business wealth (see e.g,. Carroll (2000), Vissing-Jorgensen and Moskowitz (2002); Kartashova (2014) and others).

Heterogeneity in returns is present in all years and its extent varies overtime. 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. Cross sectional heterogeneity (standard deviation) of returns on total wealth does not move with average returns. Figure 5 shows the patterns for returns on safe and risky assets. Heterogeneity on the latter covaries closely with average returns; heterogeneity in returns on risky assets is much higher, much more volatile and less correlated with average returns.

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

[11] 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.

in different percentiles of the wealth distribution using data for 2013. The differences in returns across wealth levels are large. Median returns for households at the 10th and 90th percentile of the wealth distribution are 0.7% and 2.6%, respectively. Hence, moving from the 10th to the 90th percentile of the wealth distribution the median return almost quadruples, suggesting that the correlation between returns and wealth holdings can potentially have large effects on wealth inequality.[12] How large requires new investigation. Indeed, recent calibrated models of wealth inequality by Benhabib et al. (2015) and Krusell et al. (2015) allow for heterogeneity in returns to wealth but assume absence of correlation between returns and wealth. Such correlation is invoked by Gabaix, Lasry, Lions, and Moll (2015) to explain the fast increase in tail inequality observed in many countries.

Note that returns declines at the top 1% of the distribution. As the red (crossed) line shows, this is entirely accounted for by private equity holders (who are over-represented in the top percentiles of the distribution, see Figure 1). It is plausible that private businesses use 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 distribute dividends for signaling purposes.[13]

Panel B of Figure 6 shows that the positive correlation between returns and wealth holds both for risky as well as for safe assets (and again, the slight decline at the very top is entirely accounted for by private equity holders). This rules out that the returns wealth correlation arises only because of participation costs in risky assets markets. Differences in returns on safe assets depending on wealth levels is instead consistent with differences in remuneration on deposits depending on investment size; for instance, in 2008 this ranges between 3% per year for deposits less than 7,000 dollars to 5.5% for deposits larger than USD 35,000 (see Internet Appendix, Figure IA.1).

The correlation between returns and wealth is not specific to a given year. It appears as 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

[12]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".

[13]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.

wealth levels varies considerably over the sample.[14]

In general, correlation between returns and wealth may arise because of fixed entry costs in risky assets that preclude participation by low wealth households. This is indeed consistent with a large 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, it may simply reflect the fact that wealthy investors are more risk tolerant, have a riskier portfolio, and hence receive a premium for greater risk-taking. Finally, there are important economies of scale in wealth management. 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 can access better information about where the market is heading, and hence reap higher returns on average. But perhaps the simplest explanation for the positive correlation between returns and wealth is that wealthier household have a higher exposure to risk, i.e., invest a larger share of their wealth in risky assets. Moreover, conditioning on the latter, wealthier households may invest in riskier listed companies rather than in index mutual funds. One way of checking whether risk-taking is behind the correlation documented in Figure 6 is to 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 on 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 initial year for which we have data). Clearly, wealthier individuals reap higher returns for *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 feature that will be confirmed in controlled regressions (Section 6).

Finally, we notice 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

[14]Fagereng et al. (2016) show that even a small positive correlation between returns and wealth can significantly overstate inequality measures when wealth is estimated by capitalizing income from tax returns as in Saez and Zucman (2016). We discuss this evidence in Section 7.

wealth distribution. In 2013 heterogeneity is relatively high at low levels of wealth and is 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 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 the 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 fixed-effect 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.

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, such as age, common shocks (time effects), etc. 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: $r_{igt} = r_t^s + \pi_{igt}(r_t^m - r_t^s)$, where π_{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 π_{it}, and their interaction would absorb all the existing variation. If some individuals can invest also in a private businesses, as in Quadrini (2000) and Aoki and Nirei (2015), the return on wealth can be written as $r_{igt} = r_t^s + \pi_{igt}^{m,l}(r_t^m - r_t^s) + \pi_{igt}^{m,u}(r_{igt}^{m,u} - r_t^s)$, where $\pi_{igt}^{m,l}$ and $\pi_{igt}^{m,u}$ 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 for the share of wealth invested in listed stocks and in private businesses separately. To capture any direct correlation between returns and wealth due for instance to investment size effects, we add to the specification a full set of dummies for the individual wealth percentiles computed using lagged wealth values.

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 being 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. These may arise from systematic differences in the composition of the risky portfolio and thus in returns on wealth, from differences in ability in portfolio management or in opportunities to access investment alternatives - including persistent differences in private businesses productivity. The error term e_{igt} captures unsystematic idiosyncratic variation in returns reflecting "good/bad luck". This representation allows as to decompose idiosyncratic variation in returns to wealth as $var(u_{igt}) = var(f_{ig}) + var(e_{igt})$.

Our data allow us to study two additional dimensions of persistence in returns to wealth: across generations and across spouses. Because we observe several generations in our data, we can study intergenerational persistence in returns fixed heterogeneity 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 - whose properties (cross sectional variance and intergenerational persistence) can in theory explain the thickness in the distribution of wealth as shows by Benhabib et al. (2011). The aforementioned variance decomposition into $var(f_{ig})$ and $var(e_{igt})$, together with intergenerational persistence in f_{ig} is key for the design of optimal capital income taxation (Shourideh (2014)).

Our data also allow us to study persistence in returns as people move across marital status (i.e., from singlehood to marriage). In the same way that assortative mating on individual incomes exacerbate household income inequality, assortative mating on individual wealth and returns to wealth may exacerbate household wealth inequality. We study these additional dimensions of persistence in Section 7.

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 the results 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. Observable heterogeneity in wealth returns is captured by demographics (gender, municipality fixed effects, age dummies, number of years of education, a dummy for economics or business education, employment, and marital status dummies), year fixed effects (to capture aggregate variation in returns), and the lagged shares in listed risky assets and in private business out of total wealth. To capture the relation between returns and wealth in a flexible way we add a full set of wealth percentile dummies. To avoid spurious correlation arising from the wealth accumulation equation, which implies that next year's wealth is positively correlated with current returns, we compute wealth percentiles using lagged wealth. 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 13.6 basis points higher return), while having received economics or business education is associated with 11 basis points higher returns. Because education is a permanent socio-economic feature, its effect cumulates. A systematic difference in returns of 25 basis points enjoyed by Econ 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 is over and above 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 references 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 U.S. data.

Not surprisingly, portfolio shares in listed stocks 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 significantly larger than the effect of the share in listed shocks. 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 firms by the same amount is associated with a much larger increase in returns on wealth of 186 basis points. This finding is consistent with the idea that because private business wealth is highly concentrated, it yields a premium. It runs contrary to that of Vissing-Jorgensen and Moskowitz (2002), who find no evidence that private businesses earn a premium with respect to public equity using data from the US SCF; but 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. Time fixed effects, though not shown, are always significant, as are age dummies.

Yet, overall observable characteristics explain only about 8% of the variance of individual returns to wealth. This limited fit (or the larger role of unobservable heterogeneity) is remarkable because, as noticed, 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.[15]The fit of the model improves (the R^2 increases from 0.079 to 0.117) but the size and significance of the other effects are otherwise unchanged.

[15]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 of retained profits. A reform passed in 2006, but anticipated since at least 2001, introduced taxation of distributed dividends at 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.

The third column adds the individual fixed effects to the specification in column 1.[16] 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 off individuals who change portfolio composition over time. The effect of the share in listed stocks is now larger and that on 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 and 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.[17] The last 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 of individual returns to wealth.

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 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.2 in the Internet Appendix).

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 include 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

[16]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.

[17]Some of the R^2 increase comes 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 0.11); 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.

unchanged: the sign and size of the covariates - 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 increases substantially 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 IA.4 and Figure IA.5), reassuring that the properties of persistent heterogeneity that we identify and discuss below are robust to the way returns to wealth are measured.

5.4 Persistent heterogeneity

Figure 9 plots the empirical distribution of the individual fixed effects (from the estimates in column 4, Table 2, measured in deviations from the overall mean of returns, 3.16% from Table **??**, panel C). The distribution has a long right tail (a skewness coefficient of 1.87)[18] and is quite dispersed, with a standard deviation of 2.8 percentage points and a 90th-10th percentile difference of 6.4 percentage points. It also shows considerable excess kurtosis (13.5).

One interesting question is whether the persistent component of wealth returns is associated with observable characteristics that, a priori, one may consider economically relevant. Figure 10 plots the distribution of fixed effect for business owners and non-owners (first panel); top vs. bottom wealth groups (second panel); more vs. less educated individuals (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 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 as well as among people at the bottom of the wealth distribution. But the distribution of the permanent component of returns is

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

more spread out and returns are on average higher among the wealthy, with differences in 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 general 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 more shifted to the right, consistent with a positive correlation with education. Table 4 shows summary statistics of the distribution of the return fixed effects for the total sample and for various population subgroups (using our baseline measure). The last two columns shows summary statistics of the distribution of the fixed effects obtained when we drop the private equity holders (return measure B, equation 3), and when we use a measure of returns that include unrealized capital gains (return measure C, equation 4). In both cased, the estimated standard deviation is similar to the baseline but both distributions are somewhat more skewed and exhibit more kurtosis.

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. In Internet Appendix Table IA.5 we report the estimated variances of the two components for different samples and specifications. In our baseline specifications (with controls), 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 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. 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. Average returns are, as documented above, increasing in wealth, and at an increasing 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 for explaining the correlation between returns and wealth at the very top. In particular, the bottom panel of Figure 11 shows that a key role is played by the share of wealth held in risky assets, which 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). The role of return fixed effects, in contrast, 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 top percentiles of the wealth distribution with an increase in average return of 28 basis points for every 10 percentile increase in wealth. Some simple calculations may illustrate the separate importance of observables vs. fixed effects in driving the correlations between wealth and returns. The average return to wealth increases by 83 basis points as wealth increases from the 75th to the 90th percentile; 80% of this increase (66 basis points) is explained by the increase in the fixed effects and very little by compensation for (more) risk taking. On the other hand, 2/3 of the 300 basis points increase in average return as wealth climbs from the 90th to the 100th percentile reflects compensation for greater exposure to risk and only 1/3 the increased values of individual fixed effect.

To dig deeper into the relation between returns, wealth, and risk-taking, Table 5 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. The first column only controls for initial wealth; it documents that risk adjusted returns are strongly increasing with the individual wealth percentile, again implying that the correlation between wealth and returns is not merely relecting compensation for risk. 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 business changes very little the relation with wealth. Interestingly, more educated individuals display higher Sharpe ratios and so do individuals with a business or economics degree; compared to those with investments in private equity (the excluded category) holders of private businesses attain lower Sharpe ratios, particularly those with longer exposure to it. This feature may reflect poorer diversification of idiosyncratic risk among private business holders that is only partially compensated with higher monetary returns or that is compensated with non-monetary, and thus unmeasured, returns. Findings however are unchanged when we exclude private equity holders (third column).

Macroeconomic papers in the spirit of Benhabib et al. (2011) face the problem that the key object of interest, the empirical distribution of persistent differences in returns, is typically unknown. Our estimates can be used to shed some light on this matter. Table 4 shows correlation coefficients between average fixed effects and wealth percentiles, and the slope parameter of regressions of average fixed effects and wealth percentiles. Using these data, a summary characterization of the distribution of the return fixed effects (ignoring moments higher than the second) is $f_{ig} \sim (mean = 3\% + 0.028(P_{iw} - 50), Sd = 2.8\%)$ where P_{iw} is the wealth percentile of individual i and 3% is the average return on wealth over the sample period (Table **??**, panel C). Because our regressions control for individual wealth, the correlation between the return fixed effects and wealth cannot be due to wealthier people having access to more remunerative investments. It is instead consistent with the idea that those with 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).[19]

In sum, the decomposition results in Figure 11 and the estimates in Table 5 suggest that the positive correlation between returns and wealth is driven primarily by compensation for risk and by positive correlation between wealth and unobserved heterogeneity, possibly capturing compensation for ability to generate returns. Their relative importance, however, varies over the spectrum of the distribution. Below the 90th percentile, the correlation is almost all due to unobserved heterogeneity; above it, 2/3 of the correlation reflects compensation for greater exposure to risk, while the rest reflects fixed unobserved heterogeneity.

7 Other dimensions of persistence

7.1 Intergenerational persistence in returns to wealth

The Norwegian data contain both an individual identifier and a family identifier. Hence, it is possible to link individuals across generations. To focus on a sharper case, we look at fathers

[19] In the appendix 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.

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, one would like 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, in the second, a coefficient of 0.10 with a standard error of 0.004. This suggests that most of the intergenerational correlation in returns to wealth is the 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 revert quickly to the mean.

For comparison, a regression of the father's wealth percentile rank on the wealth average percentile rank of the child has a coefficient of 0.29 (s.e. 0.006) (the dashed line in the graph).[20] Thus, intergenerational correlation in returns is three times weaker than that in wealth. Furthermore, while intergenerational correlation of returns weakens or even turn

[20]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).

negative at the top of parents returns, the opposite is true for the correlation in wealth across generation 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 kids imitate the investment strategies of their parents, or that they inherit from their parents traits 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 just reflect intergenerational correlation in wealth or aggregate shocks to returns. The positive correlation in the second panel of Figure 12 (between the child's and the father's return fixed effects), 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 6 using children's and fathers' return percentiles; 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. Results are unaffected when individual controls are added (third column). Including individual fixed effects in the last column flattens the relation even further, but raises considerably the fit (the R^2 increases from 0.06 to 0.36) suggesting that the intergenerational correlation in returns is driven by the permanent component of returns. Results are confirmed dropping private business owners from the sample and using alternative definition of returns (Internet Appendix Table IA.6). Intergenerational persistence is also detected if we use Sharpe ratios of fathers and children (Table 7) confirming that it is risk-adjusted returns that correlate across generations.

Table IA.2 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 to the bottom quintile. Overall, our data suggest substantial persistence and heterogeneity in returns within a generation but smaller persistence across generations. This result is similar to that found by Benhabib and Bisin (2016). In their exercise there is little evidence of intergenerational persistence in returns. In our case, with a substantial amount of statistical power, we find an economically small but statistically significant degree of persistence.

7.2 Assortative mating in returns to wealth

The labor literature has documented significant assortative mating by education, income, and parents' wealth (Eika, Mogstad, and Zafar (2014); Greenwood, Guner, Kocharkov, and Santos (2016); Lam (1988); Charles, Hurst, and Killewald (2013)). As far as we know, there is no evidence on assortative mating on personal wealth or returns to wealth. Both can be studied with our data. Differently from assortative mating on stable characteristics such as education or even parents' wealth (which do not vary because of marriage and can thus be measured *after* people have married), detecting assortative mating on personal wealth and returns to wealth requires that these variables are observed *before* marriage, which explains the practical absence of any empirical evidence on it. After marriage, individual wealth and returns to wealth are hard if not impossible to separate. The large sample size and the long time span of our data allow us to identify and follow individuals for several years before they get married and thus test whether individual wealth and pre-marriage returns of future spouses or partners are correlated as implied by assortative mating.

Figure 13 documents assortative mating in returns to wealth. To construct this picture, we focus on a sample of individuals who make a transition from singlehood to marriage or cohabitation at some point during our sample period (some of the singles may be separated or divorced). We start by computing the average return during pre-marriage years. In Norway people share a tax ID if they are married or if they are cohabiting and sharing the care of a child. To avoid contamination induced by the fact that what appears as singlehood may be childless cohabitation, we experiment dropping two or four years before firstly observing individuals sharing a tax ID.[21] Figure 13 shows that pre-marriage average returns to wealth are positively correlated. This remains true regardless of whether we drop the two or four years preceding marriage (or the birth of a child to cohabiting couples). However, the extent of the correlation is lower than what perfect sorting would imply (a scenario represented by the 45 degree line in the figures). As in the case of intergenerational correlation, assortative mating in returns may reflect assortative mating in attributes such as education or wages (which is well documented in the literature) or assortative mating on wealth (on which there is instead no evidence in the literature, mostly due to lack of data). Indeed, Figure 14 shows

[21] Some childless cohabiting couples may have tax incentives to marry, as the wealth tax has an allowance deduction that applies to each individual separately.

that spouses sort on their pre-marriage wealth.

To detect assortative mating on returns over and above assortative mating in wealth or other traits, we estimate the following model:

$$r^h_{it-pre} = \lambda r^w_{it-pre} + Z'_{it-pre}\mu + \varphi_{it-pre} \tag{8}$$

where r^j_{it-pre} is the average return to wealth for spouse $j = \{h, w\}$ measured in the years before marriage or shared custody of a child (excluding the two years before we firstly observe the individuals sharing a tax ID), Z_{it-pre} is a vector of controls, and φ_{it-pre} an i.i.d. error term.

Table 8 shows the results of the estimates. Before delving on assortative mating on returns to wealth, we document extensive assortative mating on personal wealth (see column 1). Pre-marriage wealth of the two future spouses appears to be strongly correlated with a slope coefficient of 0.24. This results is distinct from Charles et al. (2013), who document that spouses sort on the basis of *parental* (not personal) wealth. The other columns show results for the returns to wealth. The second column shows a positive correlation between the pre-marriage returns to wealth of the two spouses, with a slope of 0.12. This is consistent with people choosing to marry an individual with a similar ability to generate returns out of wealth. In column (3) we control for the possibility that the association between returns is spuriously due to assortative mating on other traits, such as education (length and type), aggregate effects (which we control for using year of marriage dummies), or life cycle stage (which we control for using age at marriage). The effect declines in magnitude but not in statistical significance. In light of the positive correlation between returns and wealth, the positive correlation between the returns of the perspective spouses could reflect the assortative mating based on wealth documented in the first column. The other columns of Table 8 add increasingly finer controls to make sure that assortative matching in returns is not spuriously induced by assortative mating on wealth. In column (4) we classify each spouse by whether they have above or below median pre-marriage wealth and then add dummies for whether both are above median ("both rich"), both below median ("both poor"), and so forth (with "both poor" being the excluded category). Controlling for wealth pairings lowers slightly the slope of the mating relation in returns (from 0.08 to 0.07), but significance remains high. The remaining two columns add even more granular controls for assortative mating on wealth: 25 dummies corresponding to the pairings of five wealth quintiles (column (5)), or 10,000 dummies corresponding to the pairings when spouses are classified according to their wealth percentiles (column (6)). Results remain similar, implying that assortative mating on returns is not just a reflection of assortative mating in wealth. Even within a narrow wealth group (say, husband and wife both in the top percentile), the husband's pre-marriage return to wealth is positively associated with the wife's pre-marriage return.

Assortative mating on returns to wealth can be predicated on two possibilities. First, it may reflect yet another trait people match on – such as preferences for risk, entrepreneurial spirit, etc.. Second, it may be justified by a demand for wealth preservation: the wealth of a high-return individual may be threatened by the poor return of his/her spouse. For this to be the case the low-return spouse must contribute to the wealth management of the households once assets are jointly owned after marriage. If instead, following marriage, the management of household finances is taken care of by the high-return spouse there is no scope for assortative mating. To test whether assortative mating on returns is justified by this mechanism, we regress the post-marriage (household) return against the pre-marriage returns of the two spouses. In particular, we consider the regression:

$$r_{it-post} = \beta_0 + \beta_1 min\{r^h_{it-pre}, r^w_{it-pre}\} + \beta_2 max\{r^h_{it-pre}, r^w_{it-pre}\} + Z'_{it}\theta + \nu_{it}$$

where $r_{it-post}$ denotes the post-marriage *household* return, r^j_{it-pre} is the pre-marriage returns of spouse $j = h, w$, Z is a vector of controls and ν an error term. Results of the estimates of this model are shown in Table 9. Interestingly, both the lower and higher pre-marriage return contribute to predicting the post-marriage household return. Their effect barely changes when adding controls. The return of the spouse with the higher pre-marriage return has the strongest effect with a coefficient that is more than five times that of the return on the spouse with the lower pre-marriage return. But the latter matters, leaving room for some demand for assortative mating. Letting $r^L = min\{r^h_{it-pre}, r^w_{it-pre}\}$ and $r^H = max\{r^h_{it-pre}, r^w_{it-pre}\}$, we can summarize the effect of pre-marriage returns on post-marriage returns as $r_{it-post} = (\delta r^L + (1-\delta) r^H)$, where $\delta = \frac{\beta_1}{\beta_1+\beta_2}$ is the weight assigned to the spouse with the lower return. Using the estimates in column (4), the weights are 0.92 for the highest return and 0.08 for the lowest.

The last column adds interaction terms between the maximum and minimum return with a dummy for whether the highest return is that of the male, to account for potential differences across genders in the weight of the lowest and highest pre-marriage return. The results show that when the spouse with the higher pre-marriage return is the male, the effect of the higher pre-marriage return becomes higher (0.25) and that of the lower pre-marriage return close to zero, implying $\delta = 0$. When instead the spouse with the highest pre-marriage return is the female, the two effects are respectively 0.19 and 0.04. The weights on the lowest and highest return become $\delta = 0.18$ and $(1-\delta) = 0.82$. In other words, males carry a higher weight than average both when they are better at generating returns as well as when they are not.

8 Discussion and Conclusions

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

Measurement of wealth trends Saez and Zucman (2016) have revived the debate around the medium term dynamics in wealth inequality, particularly the dynamics in the shares of wealth at the very top of the distribution. Lacking time-series of comprehensive data on wealth holdings for the U.S. like those we have available for Norway, they use tax records on income from capital to obtain underlying wealth figures and recover trends in wealth shares at the top. 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 correlation between wealth and returns change over time (which is another feature of the data). In another paper (Fagereng et al. (2016)) 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 generate significant deviations between measures of inequality based on imputed and actual wealth.

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, using World Bank data, in Denmark and Switzerland the income share of the top 10% is around 25%, much lower than the 40% corresponding share in the U.S..[22] Yet, the top 10% wealth share is 76% in Denmark and

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

71% in Switzerland, which is even higher than the figure for the U.S. (70%, Davies et al., 2011). The comparison between the U.S. and Norway is even more striking: in 2012 and using comparable definitions, the top 0.1% income share in Norway is just above 1 percent and the top 0.1% wealth share 10%; on the other hand, the top 0.1% income share in the U.S. is 8% - eight times larger than that in Norway - while the top 0.1% wealth share is just twice as large (22%).[23] Heterogeneity in returns to wealth may rationalize the puzzle of why a country with much lower concentration of income at the top than another 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 shares 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. One interesting question is what happens when the income distribution has thick tails and returns are heterogeneous at the same time. Can one type of heterogeneity dominate the other? This is relevant to solve the above puzzle. Simulation results by Benhabib et al. (2015) imply that the tail is dominated by persistent returns heterogeneity, suggesting that return heterogeneity can give rise to similar wealth inequality in two countries with very different levels of income inequality. Interestingly, Benhabib et al. (2015)'s simulations show that a degree of returns heterogeneity close to the one we estimate for Norway can fit well the top wealth shares in the U.S.. It would be interesting to see whether this would still be the case when combining the same amount of returns heterogeneity with the features of the Norwegian – rather than American – income distribution.

Taxation of capital income and taxation of wealth Our findings speak also to the emerging debate on capital income and wealth taxation. In most standard models, without heterogeneity in returns, taxing capital income and taxing the stock of wealth are equivalent. However, when returns are heterogeneous, taxing income from capital and taxing capital can have

[23] Top income shares for the U.S. are from the Wealth and Income Database: http://www.wid.world/#Database; the U.S. top wealth shares are from Saez and Zucman (2016). For Noway we compute top income share and wealth shares from the registry data using definitions that are as close as possible to those of Saez and Zucman (2016). The top 0.1% wealth share we find is also comparable to what Epland and Kirkeberg (2012) find.

important efficiency implications as shown by Guvenen et al. (2015). In fact, holding constant tax revenue, replacing a capital income tax with a wealth tax tends to widen the after-tax heterogeneity in returns - i.e. after-tax returns on wealth become more spread out when wealth is taxed than when capital income is taxed. Intuitively, taxing capital income lowers disproportionately the after-tax return of individuals with high rates of returns; 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 - if it is persistent - and its extent. Our results inform on both dimensions; the extent of measured persistent heterogeneity suggest that the efficiency concern of capital income taxation raised by Guvenen et al. (2015) are of practical relevance. Furthermore, when returns have also a transitory idiosyncratic component, in addition to the permanent one, the relative importance of the two sources of cross sectional heterogeneity are relevant for the progressivity of capital income taxation (Shourideh, 2014). Our variance decomposition provides information to assess empirically how far actual taxation of capital income is from optimal.

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 understood, our data reveal features that have so far been neglected in models that emphasizes 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 modeling of returns from public equity and from private equity in order to be better able to understand the drivers of wealth inequality. We also document persistence in returns across marital statuses, because people sort also on the basis of pre-marital returns to wealth and because pre-marriage returns of both spouses affect the return to wealth of the family. We are unaware of any model that accounts for mating on returns to wealth and allocation of wealth management responsibility within the family. Yet, they are potentially relevant for heterogeneity in returns to wealth and thus for wealth concentration. We plan to assess the role of these properties for wealth inequality in future work.

More generally, the effects on wealth inequality and 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 allows explicitly for returns heterogeneity in human and non-human capital as well as in key preference parameters can allow 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

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

Aït-Sahalia, Y. and L. P. Hansen (2009). *Handbook of financial econometrics, Vol 1: tools and techniques.* Amsterdam: Elsevier Science.

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.

Aoki, S. and M. Nirei (2015). Zipf's law, pareto's law, and the evolution of top incomes in the US. Working paper.

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.

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

Black, S. E., P. J. Devereux, P. Lundborg, and K. Majlesi (2015). Poor little rich kids? 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.

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. Dıaz-Giménez, and J.-V. Rıos-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.

Charles, K. K., E. Hurst, and A. Killewald (2013). Marital sorting and parental wealth. *Demography 50*(1), 51–70.

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.

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 *Studies in the Economics of Aging*, pp. 331–362. University of Chicago Press.

Eika, L., M. Mogstad, and B. Zafar (2014). Educational assortative mating and household income inequality. NBER Working Paper.

Epland, J. and M. I. Kirkeberg (2012). Wealth distribution in Norway. Evidence from a new register-based data source. Reports, Statistics Norway.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2016). 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. Working paper.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2016). Technology and the changing family: A unified model of marriage, divorce, educational attainment and married female labor-force participation. *American Economic Journal: Macroeconomics, 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, 06). Portfolio choice and asset prices: The importance of entrepreneurial risk. *The Journal of Finance 55*(3), 1163–1198.

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., A. Smith, J. Hubmer, et al. (2015). The historical evolution of the wealth distribution: A quantitative-theoretic investigation. Working paper.

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

Lam, D. (1988). Marriage markets and assortative mating with household public goods: Theoretical results and empirical implications. *Journal of Human Resources*, 462–487.

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. (1999). The importance of entrepreneurship for wealth concentration and mobility. *Review of income and Wealth 45*(1), 1–19.

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.

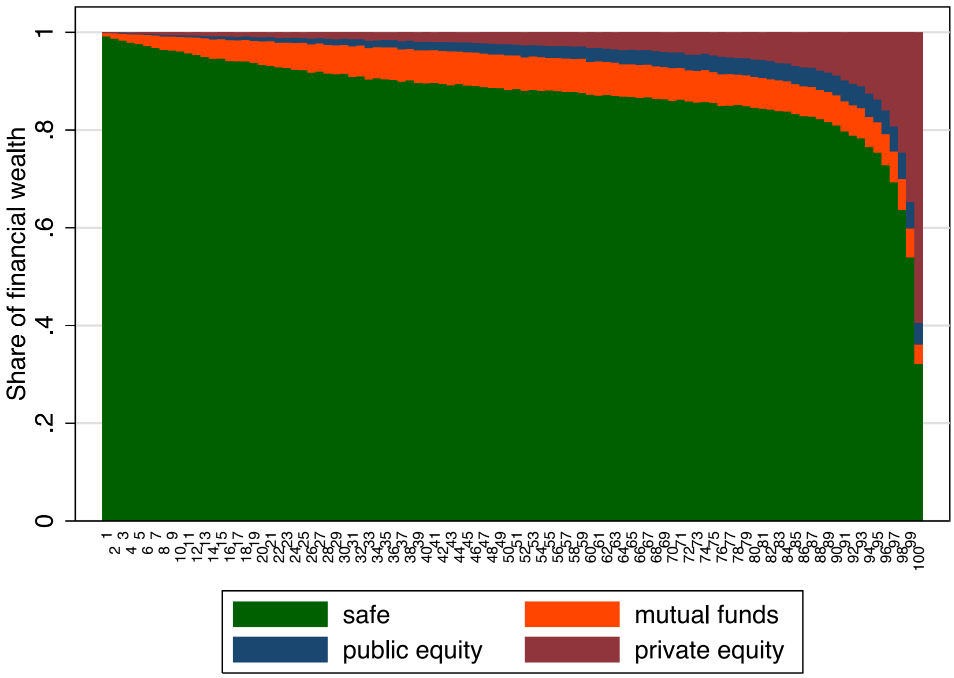
Vissing-Jorgensen, A. and T. J. Moskowitz (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review 92*(4), 745–778.

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.

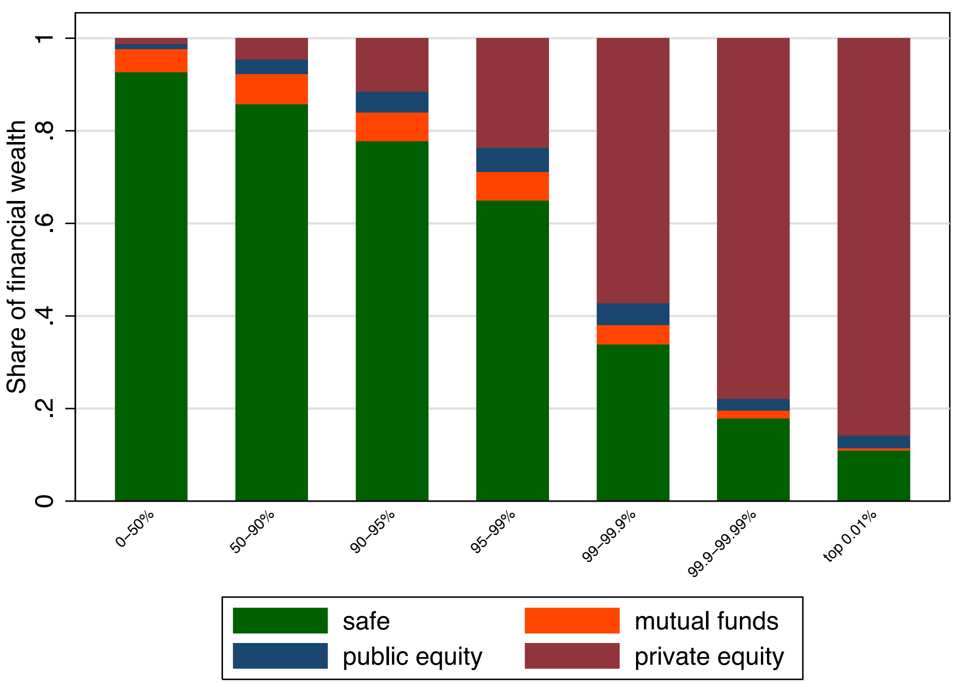
Figures

Figure 1. Portfolio Composition: by percentile

(a) All wealth percentiles



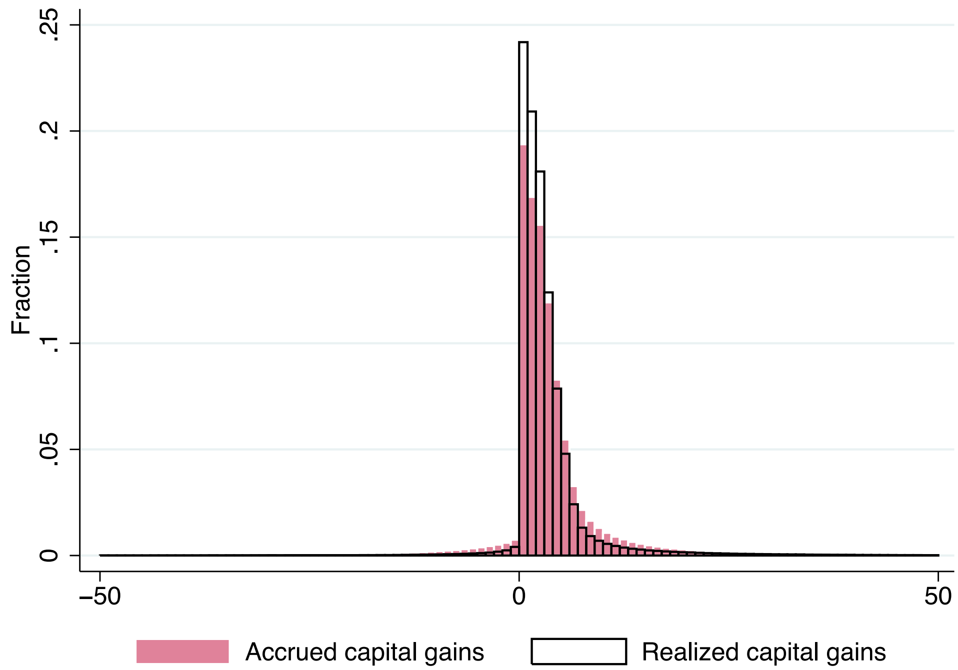
(b) Selected fractiles



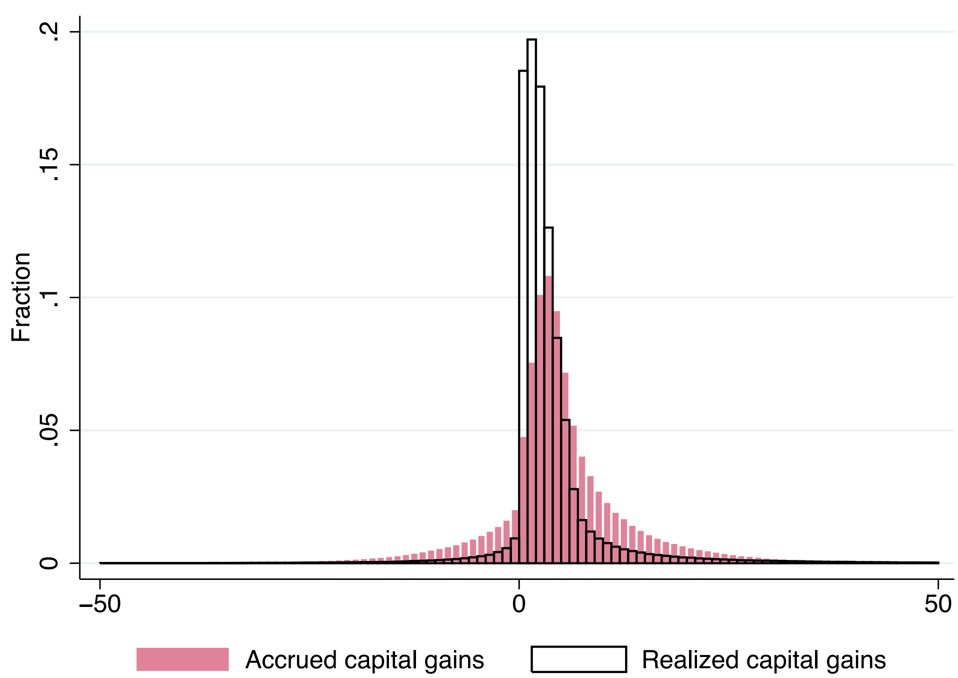
Notes: The figure shows the asset composition of Norwegian households wealth (the sum of listed stocks, mutual funds, private business wealth and liquid assets). Panel A shows the assets allocation by wealth percentiles; the second panel zooms on the composition of selected fractiles at the top. Data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

Figure 2. Distribution of returns on wealth

(a) Full sample

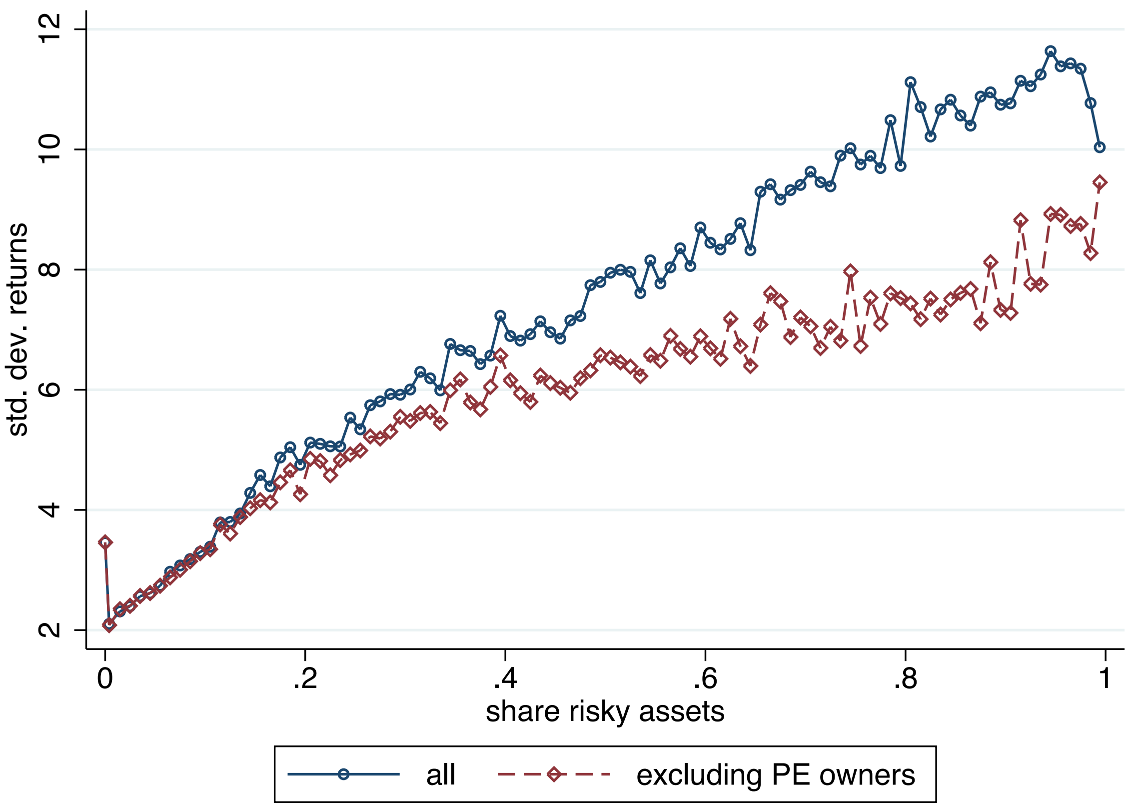


(b) Sample of risky assets holders



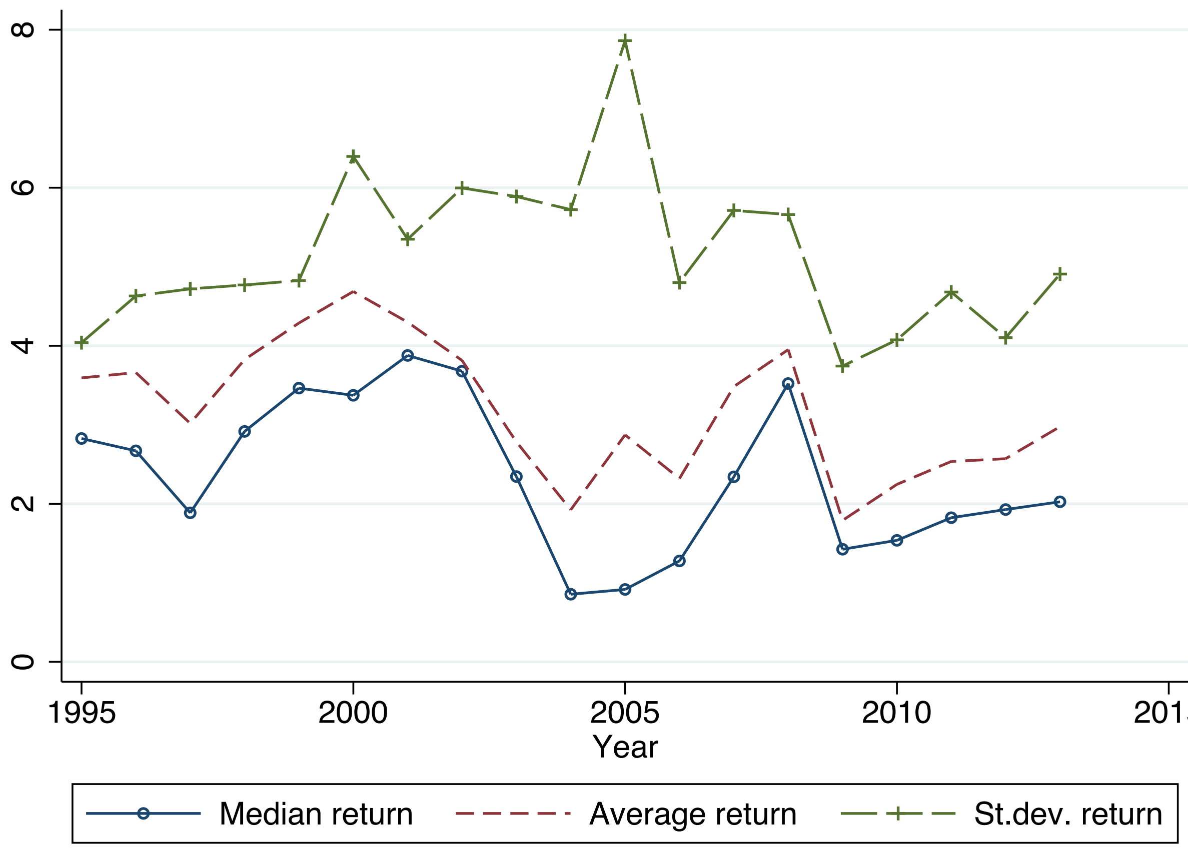
Notes: The figure shows the histograms of the individual returns on wealth across all years of the sample, 1995-2013, using the baseline definition of returns (with realized capital gains) and that with imputed capital gains. Returns are in percent, bin size equals 1.

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



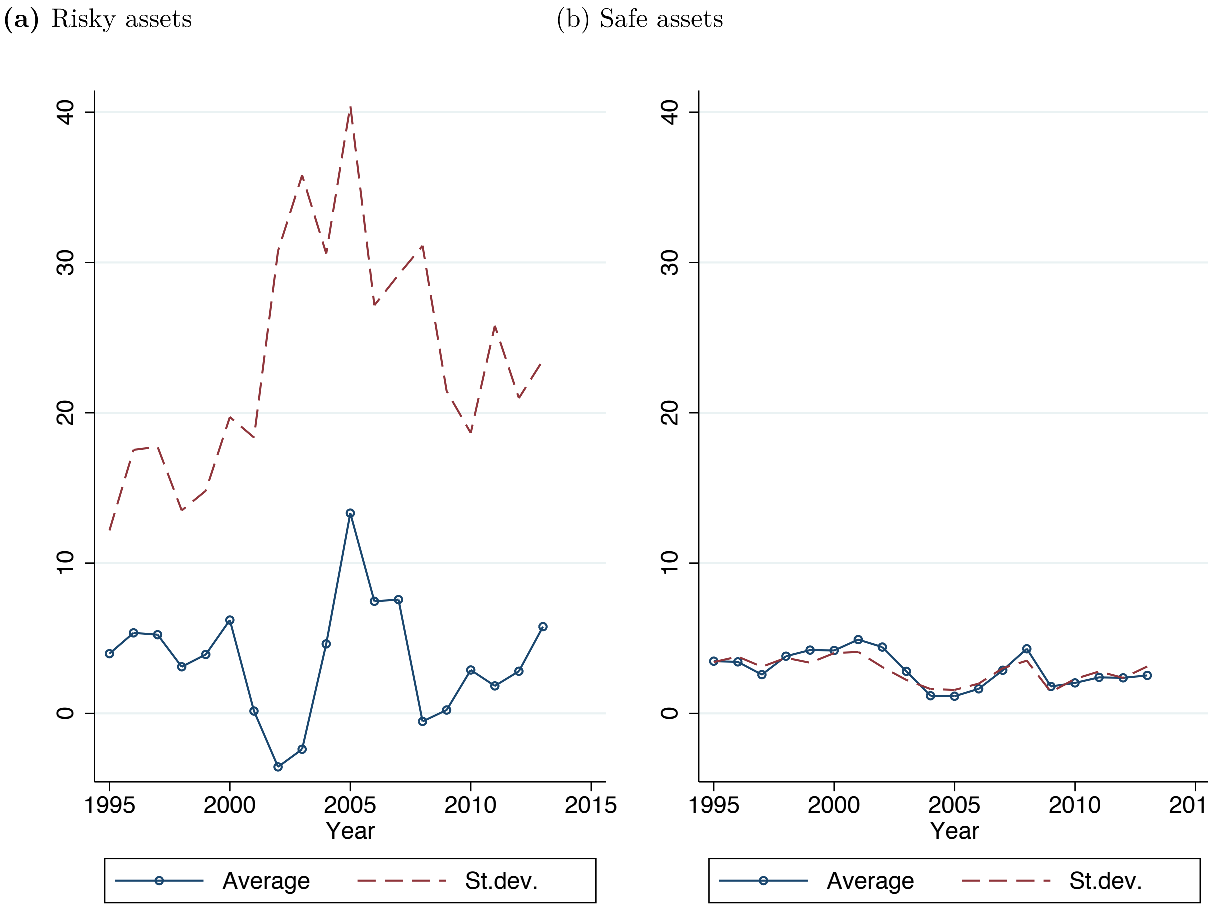
Notes: The figure plots the cross-sectional standard deviation of individual returns to wealth in 2013 by value of the share of wealth in risky assets (directly and indirectly held stocks plus private equity wealth) for the full sample (blue line) and excluding private equity holders (red line). Standard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

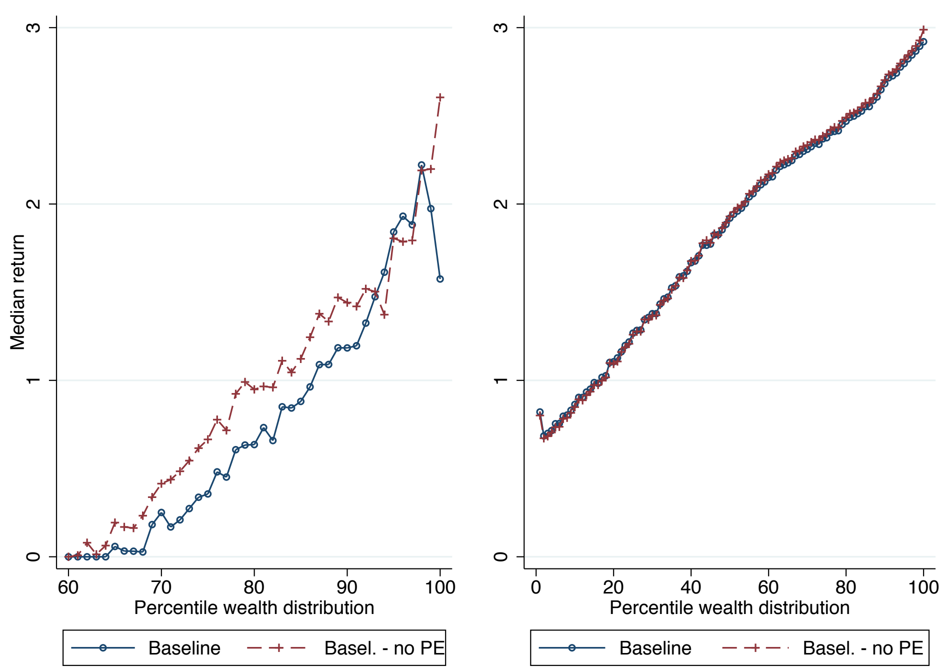


Notes: The figure plots the time patterns over the sample years of the cross sectional median and standard deviation of individual returns on wealth separately for risky (the sum of public and private equity) and safe assets. Returns are computed using our baseline definition based on realized capital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

Figure 6. The correlation between returns and wealth

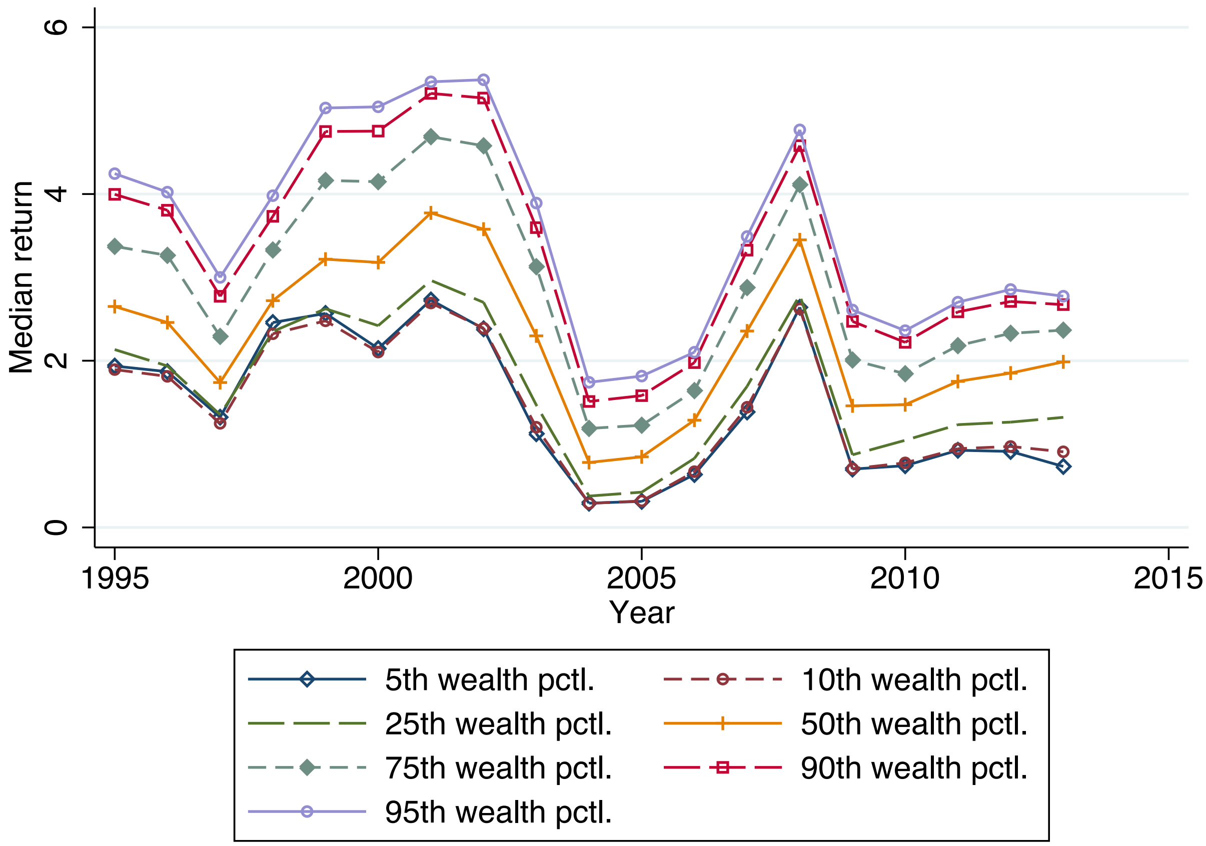
(a) All assets

(b) Risky and safe assets



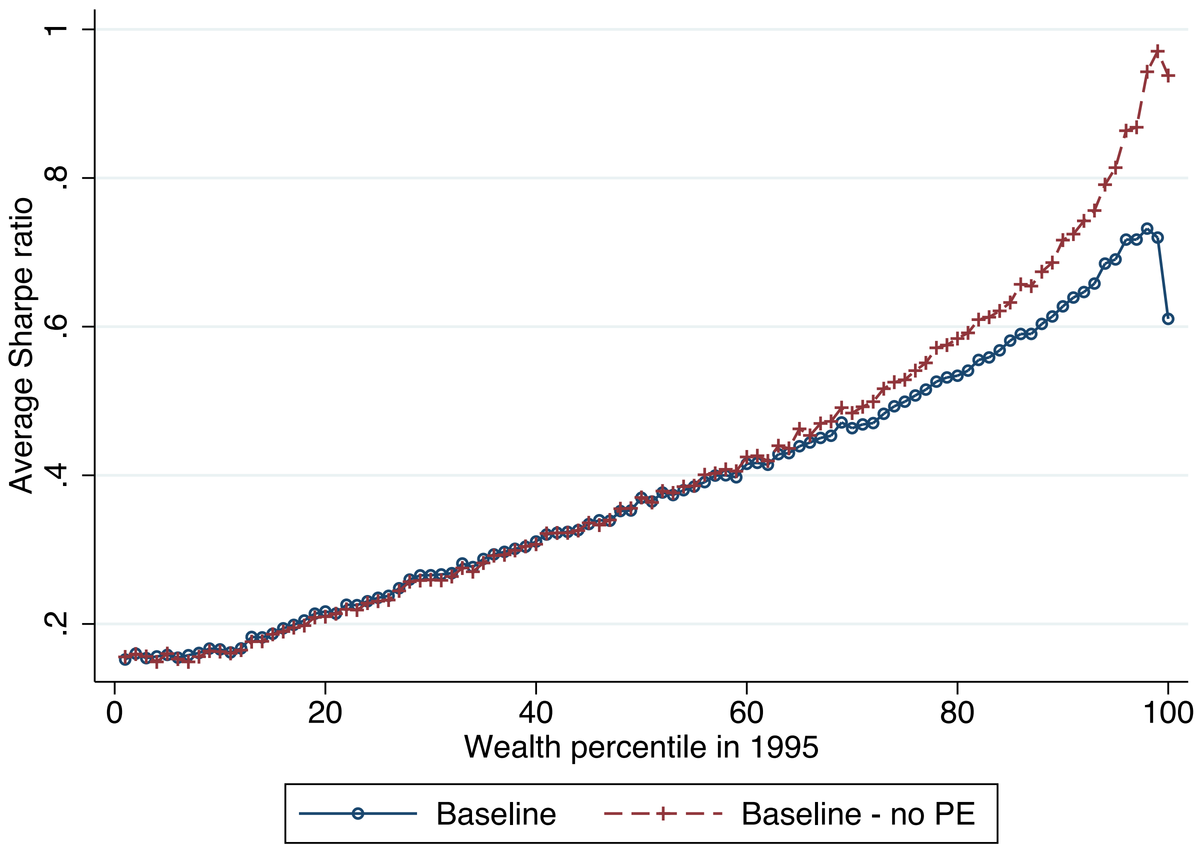
Notes: The figure shows the relation between returns on wealth and wealth percentiles in 2013. Panel (a) shows the relation for the returns on all assets, for the full sample (blue line) and excluding the private equity holders (red line). Panel (b) shows the relation distinctly for risky (left figure)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



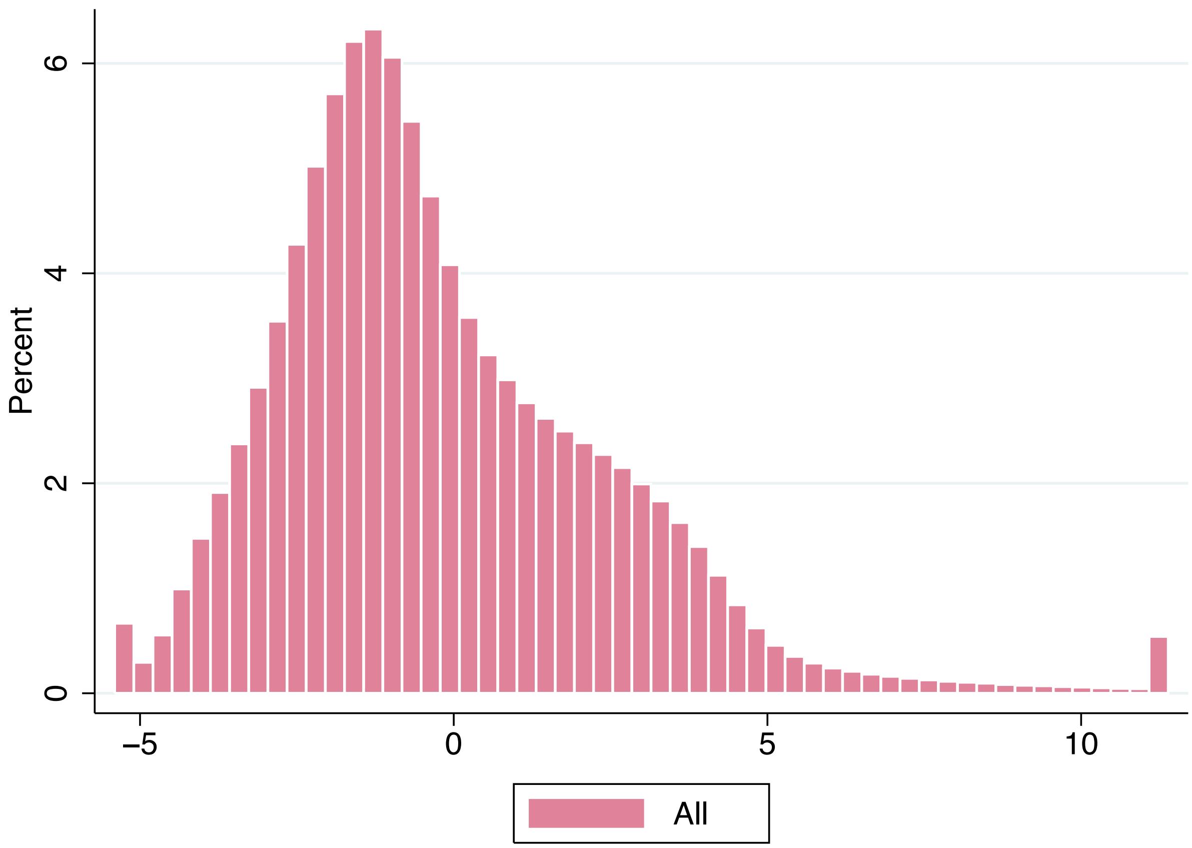
Notes: The figure plots the time pattern of median returns of individual returns on wealth over our sample period for different percentiles of total wealth. It shows both evolution of dispersion and correlation with wealth over time. Figures are in percent.

Figure 8. The Sharpe ratio and the level of wealth



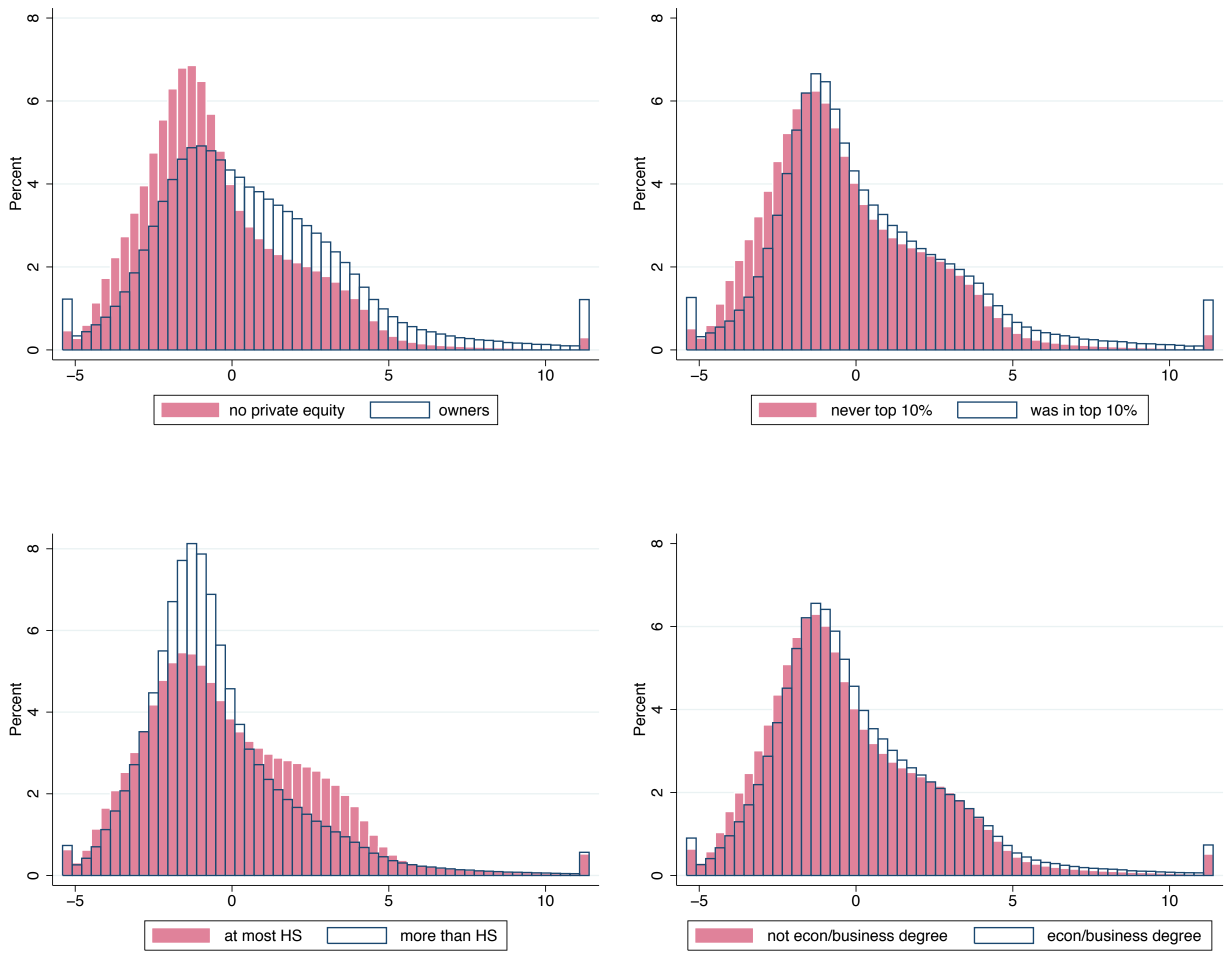
Notes: The figure shows the average cross sectional Sharpe ratio of individual wealth portfolios by wealth percentile. The Sharpe ratio is obtained by first computing deviations of individual returns on wealth from the return on the safe asset (the annualized real 3-month rate on Norwegian T-bills); 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.

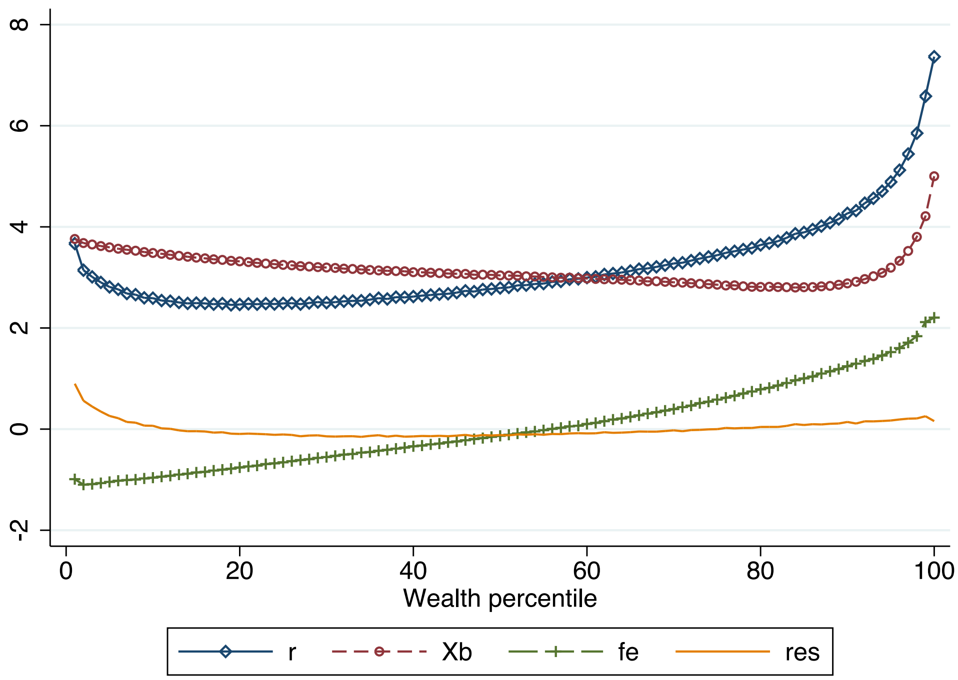
Figure 10. The distribution of estimated return fixed effects, stratifying by selected characteristics.



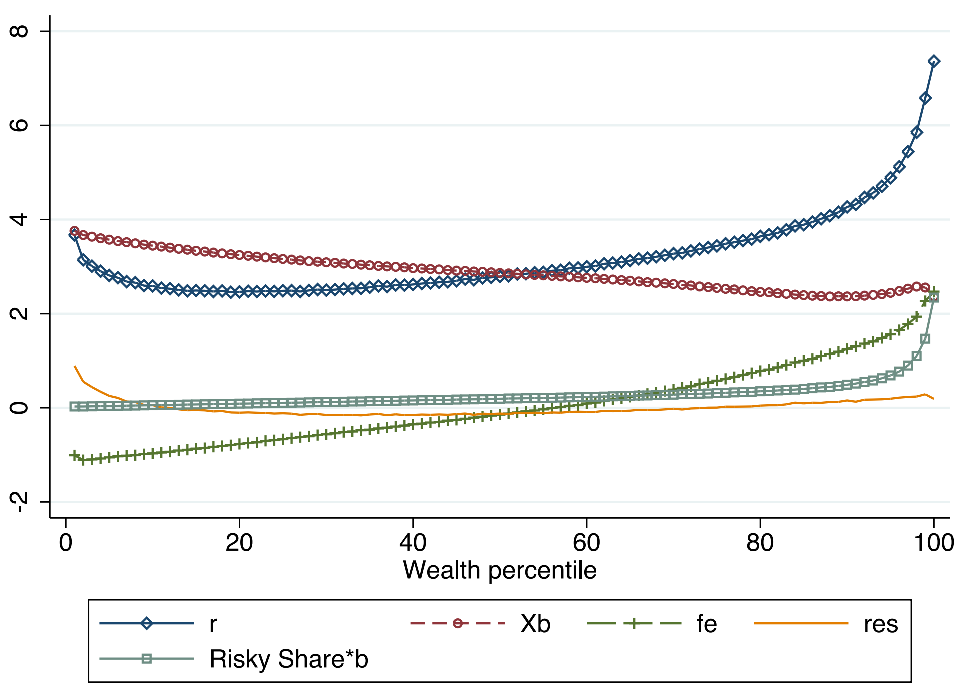
Notes: The figure shows the histogram of the estimated fixed effects in the wealth return regression using estimates in Table 2, column 3 for various subgroups of the population. For comparison it also shows the histogram for the whole population. Values above the 99.5 percentile have been grouped in 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)



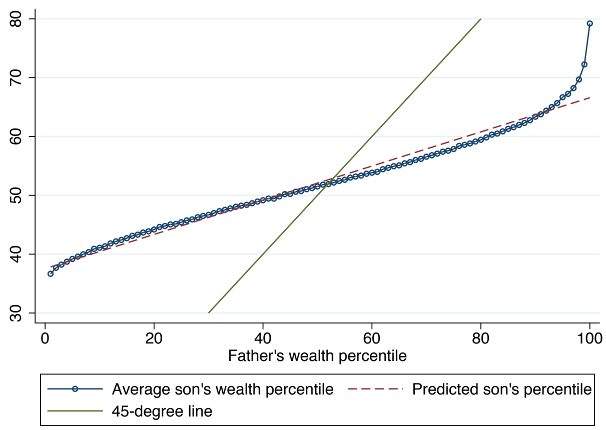
(b)



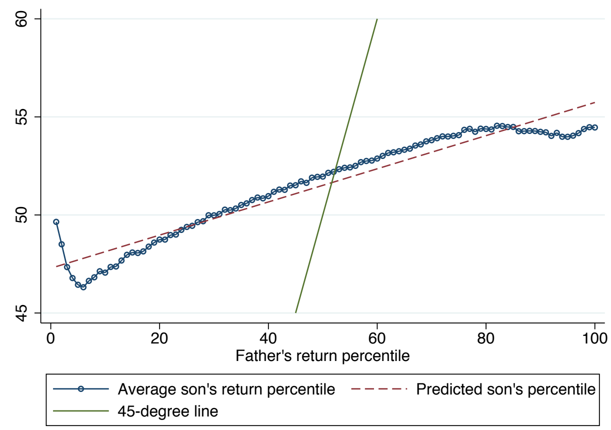
Notes: The figure shows the contribution to the predicted average return for each wealth percentile of the observed (regression controls) and unobserved (fixed effects) components and the residual using the estimates in Table 2, column 3. The first panel lumps together all the observable components. 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

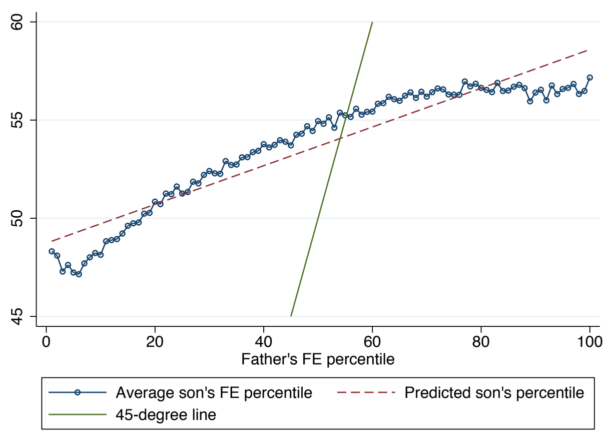
(a) Wealth



(b) Returns



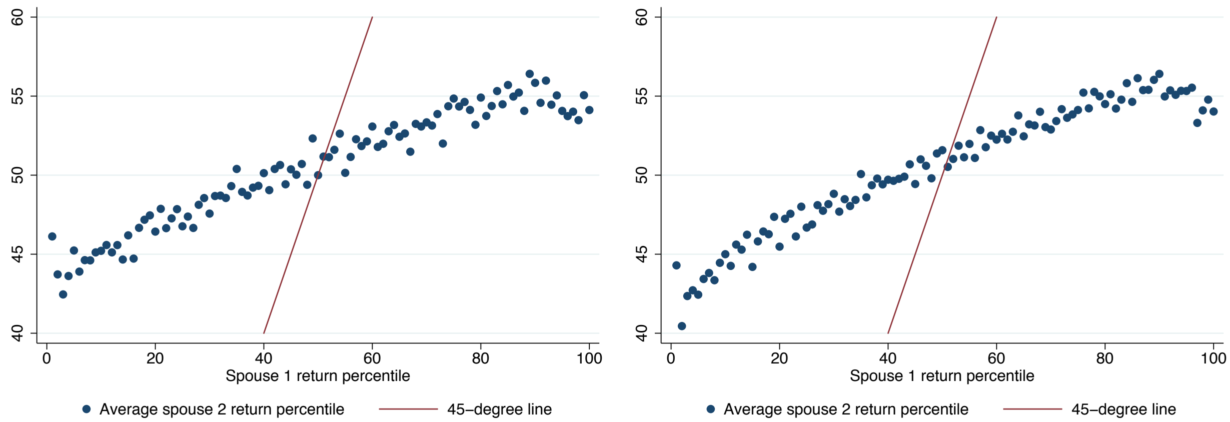
(c) Fixed Effects



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.

Figure 13. Assortative mating on returns to wealth

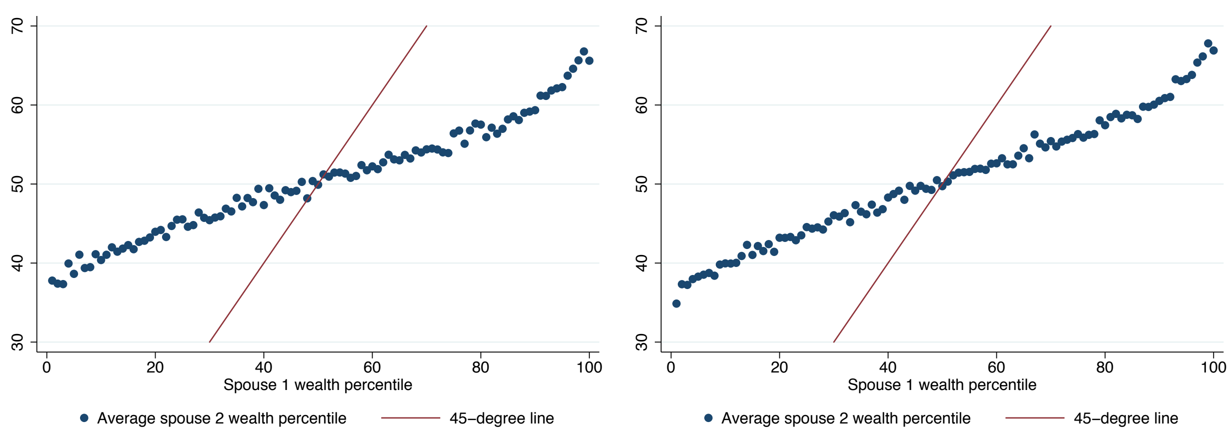
(a) Mean returns 4 years pre marriage   **(b)** Mean returns 2 years pre marriage



Notes: The figure shows the rank correlation between spouses before-marriage returns to wealth percentile. The left figure computes returns using data before 4-years from marriage; the right figure using data before 2-years from marriage. The red line is the 45 degree line.

Figure 14. Assortative mating on wealth

(a) Mean wealth 4 years pre marriage   **(b)** Mean wealth 2 years pre marriage



Notes: The figure shows the rank correlation between spouses before-marriage wealth percentile. The left figure computes wealth using data before 4-years from marriage; the right figure using data before 2-years from marriage. The red line is the 45 degree line.

Tables

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	
2.98	(4.91)	5.78	(23.50)	2.52	(3.12)

Value weighted averages (st. dev.) of returns					
Total assets		Risky Assets		Safe Assets	
3.65	(6.14)	4.82	(11.70)	2.63	(1.61)

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 at the Oslo stock exchange and mutual funds.

Table 2. Estimates of returns to wealth

	(1) Returns b/se	(2) Returns b/se	(3) Returns b/se	(4) Returns b/se	(5) Returns b/se
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). 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 b/se	(2) Returns C b/se	(3) Returns B b/se	(4) Returns C b/se
Lagged risky share	0.628***	5.096***	0.779***	3.223***
	(0.008)	(0.011)	(0.011)	(0.017)
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
Year FE	yes	yes	yes[1]	yes[1]
Age FE	yes	yes	yes	yes
County FE	yes	yes	yes	yes
Demographic controls	yes	yes	yes	yes
Lag. wealth percentile	yes	yes	yes	yes
R-squared	0.058	0.122	0.225	0.197
N	44,399,241	44,399,241	44,399,241	44,399,241

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. All regressions include interactions between time effects and the portfolio shares in risky assets and private businesses. 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

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

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. The difference in the number of observations between column 2 and column 6 comes from the following selection choices: in column 2 we start from the baseline sample and exclude anyone who own private equity ever in the sample. In the last column the fixed effects are obtained by excluding individual-year observations with private equity. For this reason, some people who are excluded from the sample in column 2 are included in the last column; moreover the estimated fixed effects are different in the two columns.

Table 5. Sharpe ratio estimates

	(1)	(2)	(3)
Wealth perc. 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		0.000***	0.000***
		(0.000)	(0.000)
Educ. Years		0.018***	0.015***
		(0.001)	(0.001)
Educ. Years Squared		-0.000***	-0.000***
		(0.000)	(0.000)
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
Sample	Full	Full	Never business Owners
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 a 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 ration between the first and the second. Robust standard errors in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.

Table 6. Intergenerational persistence in returns to wealth.

	(1) Child ret. percentile b/se	(2) Child ret. percentile b/se	(3) Child ret. percentile b/se	(4) Child ret. percentile b/se
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 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 adds also education and age; the last column also 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 7. 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 Sharpe ratio on the fathers' Sharpe ratio. The first column report 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.

Table 8. Assortative mating on returns

	(1) Wealth pctile b/se	(2) Return pctile b/se	(3) Return pctile b/se	(4) Return pctile b/se	(5) Return pctile b/se	(6) Return pctile b/se
Wealth pctile spouse	0.244***					
	(0.002)					
Return pctile spouse		0.119***	0.088***	0.074***	0.073***	0.073***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Wealth Controls:						
Poor&Rich				-0.002		
				(0.002)		
Rich&Poor				0.142***		
				(0.002)		
Rich&Rich				0.149***		
				(0.002)		
Constant	0.378***	0.440***	0.377***	0.342***	0.321***	0.409***
	(0.001)	(0.001)	(0.078)	(0.080)	(0.079)	(0.081)
Wealth controls		no	no	Rich x Poor	5 x 5	100 x 100
Age ind.		no	yes	yes	yes	yes
Marital year ind.		no	yes	yes	yes	yes
Education length/type ind.		no	yes	yes	yes	yes
R-squared	0.059	0.014	0.058	0.118	0.139	0.200
N	160,860	160,860	160,860	160,860	160,860	160,860

Notes: The table shows regressions documenting assortative mating in wealth (first column) and in returns (remaining columns). In the first column the left hand side variable is the wealth percentile of the husband 4-years before marriage and the explanatory variable is the wealth percentile of the spouse before marriage. In the other columns the left hand side variable is the return percentile of the husband 4-years before marriage. Robust standard errors in parenthesis; ***p-value<0.01, **p-value<0.05, * p-value<.10.

Table 9. Determinants of post-marriage returns

	(1) ret__post__mar b/se	(2) ret__post__mar b/se	(3) ret__post__mar b/se	(4) ret__post__mar b/se	(5) ret__post__mar b/se
Pre-marital return of spouse with lower returns	0.014	0.026***	0.023**	0.021**	0.037***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.012)
Pre-marital return of spouse with higher returns	0.229***	0.231***	0.232***	0.232***	0.191***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)
Age at marriage, female			-0.019***		
			(0.002)		
Age at marriage, male			0.011***		
			(0.002)		
Male*Return of higest					0.061***
					(0.012)
Male*Return of lowest					-0.035**
					(0.018)
Constant	2.097***	2.194***	2.465***	2.266***	2.303***
	(0.024)	(0.038)	(0.053)	(0.063)	(0.063)
Indicator controls:					
Marital year	no	yes	yes	yes	yes
Ages at marriage	no	no	no	yes	yes
R-squared	0.031	0.036	0.036	0.037	0.038
N	164,154	164,154	164,154	164,154	164,154

Notes: The table shows regressions of the return to wealth of the family on pre-marriage returns to wealth of the two spouses. The specification allows for different marginal effects of the return of the spouse with the lowest and highest return (columns 1-4) and for differences in the marginal effect of the highest return when the spouse is the male or the female. Pre-marriage returns are computed using data prior to 4-years before marriage. Robust standard errors in parentheses; ***p-value<0.01, **p-value<0.05, * p-value<.10.