The Gender Investment Gap over the Life-Cycle

Annika Bacher*

July 26, 2024

Click here for the most recent version.

Abstract

Single women invest less in risky assets than single men. This paper analyzes the determinants of the "gender investment gap" based on a structural life-cycle framework. The model can rationalize the gender investment gap without introducing gender heterogeneity in preferences. Rather, lower deterministic income and larger household sizes shift the composition of single women toward poorer households who invest less in risky assets (composition effect). Additionally, future outcomes of both variables (that cannot easily be controlled for in regressions) make single women more vulnerable to financial shocks and decrease their optimal equity share even conditional on state variables (policy effect).

Keywords: Household Finance, Life-Cycle, Gender, Portfolio Choice
JEL: E21, G11, G50, J16

*BI Norwegian Business School: annika.bacher@bi.no; address: Nydalsveien 37, 0484 Oslo; phone: +4792026933. I am indebted to my advisors, Russell Cooper and Thomas Crossley, for their continuous guidance and support. Moreover, I thank Tarun Ramadorai, two anonymous referees, my discussants Francisco Gomes, Oliwia Komada, and Francesco Maura as well as Jesús Bueren, Dean Corbae, Mariacristina De Nardi, Eirik Eylands Brandsås, Andreas Fagereng, Philipp Grübener, Michael Haliassos, Lukas Nord, Jirka Slačálek, and workshop participants at the ECB, the European University Institute, University of Wisconsin-Madison, University of Minnesota, Norges Bank, BI Norwegian Business School, the 2020 virtual ASSET conference, the 2020 EWMES, the 2021 RES Annual Conference, the 2021 GRAPE Gender Gaps Conference, the 2021 ESMES, the 2021 EFA Doctoral Tutorial, the 2021 VfS Annual Meeting, and the 2023 Bocconi Workshop on Gender Equality for comments and suggestions that helped to greatly improve this work. All remaining errors are mine. This paper is part of the research activities at the Centre for Household Finance and Macroeconomic Research (HOFIMAR) at the BI Norwegian Business School. Part of this project was conducted while I was visiting the European Central Bank, whose hospitality is gratefully acknowledged.

1 Introduction

Single women are less likely to participate in the stock market than single men and if they do, they allocate a smaller share of their portfolio toward risky assets. In the presence of an equity premium, a less risky portfolio translates (ceteris paribus) into lower wealth levels. This paper studies the sources of the so-called "gender investment gap" based on a structural life-cycle framework. I show that such a framework is able to match the empirical gender investment gap *without* introducing gender heterogeneity in preferences. Rather, lower deterministic income and larger household sizes of single women are the main determinants of the gap.

Both factors contribute to the gender investment gap through a composition and policy effect: first, they shift the composition of single women toward lower wealth households who are less likely to invest in risky assets (composition effect). Second, even conditional on state variables (such as wealth), they decrease single women's optimal equity share because larger future household sizes increase future optimal consumption and, together with lower future deterministic income, make single women more vulnerable to financial shocks, resulting in less financial risk-taking (policy effect). Consequently, reduced form regressions that control for household observable characteristics but do not take into account future outcomes fail to fully explain the empirical gender investment gap.

I first document life-cycle profiles of asset holdings and portfolio choices for single men, single women, and couples using survey data on U.S. households. My empirical findings confirm the gender investment gap: women are less likely to participate in the stock market and allocate – conditional on participating – a lower share of their portfolio toward risky assets. The gender gap is strongest among young households and declines as individuals age. All differences are statistically different from zero, even after controlling for a wide range of observable characteristics that have been shown to affect investment behavior.

To uncover which factors account for the unexplained part of the gap and to quantify the relative importance of each channel, I develop a life-cycle model of portfolio choice that allows for differences in household structure (single or couple) and gender. Individuals can get married and divorced. Single men and single women differ in their income profiles (both the deterministic and stochastic part), the number of individuals who live in their household, their marital transition probabilities, the (expected) characteristics of their partner in the event of marriage as well as survival probabilities and out-of-pocket medical expenditures during retirement. I restrict preference parameters to be identical across all types of households.

I calibrate the model using the Survey of Consumer Finances (SCF) for financial choices and the Panel Study of Income Dynamics (PSID) for labor income and demographic characteristics. The model matches well the life-cycle profiles of asset holdings and equity shares for both single men and single women. Moreover, in line with the empirical evidence, reduced form regressions on model simulated data that control for household observable characteristics fail to fully explain the gender investment, in particular among young households, even though the underlying data generating process assumes homogeneous preferences across men and women.

By means of counterfactual exercises, I show that gender heterogeneity in deterministic income and the average number of household members ('household sizes') are the most important determinants of the gender investment gap. Deterministic income matters both through its level and slope. Single women's deterministic income is lower than that of single men, making them less likely to participate in risky asset markets. Additionally, the gap in deterministic income is largest early in life, preventing single women from participating in the stock market when young. As a result, they earn lower returns in expectation, accumulate less wealth, and become less likely to participate later on. In contrast, gender heterogeneity in the stochastic component of income contributes to widening the gap.

I then decompose the gender investment gap into a *composition* and *policy* effect. The composition effect explains how much of the gap arises from differences in observable characteristics, that is in the distribution of individuals across the state space. The policy effect describes how much of the gap can be accounted for by differences in policy functions for the equity share conditional on state variables. Since all agents in the model are forward looking, heterogeneity in policy functions (conditional on the state vector) arises from heterogeneity in future outcomes.

On aggregate, the composition effect (that is, heterogeneity in the sample composition) explains most of the gender gap in participation rates. In contrast, lower female conditional equity shares among young households mainly arise due to gender differences in policy functions.

When further analyzing the importance of individual model elements for the policy effect, I find that lower future female deterministic income contributes to single women taking less financial risk, even conditional on current state variables. As long as labor income has some bond-like characteristics, the optimal equity share is decreasing in the ratio of the present value of human capital (i.e., the present value of future expected income) over current assets (Merton 1969, 1971). Hence, if a man and woman have the same current assets, it is optimal for the woman to choose a smaller equity share if her (deterministic) income is lower than that of the man in future periods, that is, if she is endowed with less human capital.

Similarly, larger female household sizes – which mainly arise through a higher likelihood of having children living in the same household – affect single women's policy functions both with regard to participation rates and conditional risky shares. Larger household sizes enter the model through equivalence scales in the utility function, lowering agents' effective discount factor. Hence, all else equal, larger future household sizes of single women decrease the present value of their future expected income, which lowers the optimal equity share.

Intuitively, labor income works as a hedge against financial risk to insure consumption. Single women who anticipate to have lower income and to consume more relative to their wealth (through larger household sizes) in future periods are therefore more exposed to financial risk, reducing their willingness to invest in the risky asset. Moreover, gender differences in deterministic income and household sizes are largest early in the life-cycle, explaining why reduced form regressions that do not take into account future outcomes fail to fully explain the gender investment gap especially among young households.

Lastly, I provide direct empirical support for the main model mechanisms. First, I show that controlling for proxies of deterministic income further reduces the gender investment gap in the data. Second, by complementing the analysis with data from the New York Fed Survey of Consumer Expectations, I show that single women expect lower future earnings and more non-spousal household members than single men.

While the focus of the paper is on stock market investment, its implications go beyond that specific application. A large literature has documented that women earn less in real estate markets and choose less risky portfolio compositions in retirement accounts.[1] Beyond financial markets, there exists evidence that women, and in particular single mothers, sort into less risky occupations, changing their trajectory of lifetime earnings (e.g. Bertrand 2011; DeLeire and Levy 2004). Dohmen et al. (2011) combine survey and experimental evidence to show that women behave more risk averse with regard to career choices and financial outcomes. In turn, lower asset returns have been linked to slower wealth accumulation and financial vulnerability of women, especially during old age (e.g. Neelakantan and Chang 2010; Goldsmith-Pinkham and Shue 2023).

However, to correctly design and evaluate policies that aim at promoting female financial security, it is important to understand whether gender differences in risk-taking arise from underlying variation in preferences or from societal constraints that result in women making less risky choices.[2] If differences are purely preference-driven, both men and women behave

[1] See e.g. Sunden and Surette (1998); Agnew et al. (2003); Arano et al. (2010); Säve-Söderbergh (2012) on retirement accounts and Andersen et al. (2021); Girshina et al. (2022); Goldsmith-Pinkham and Shue (2023) on gender gaps in real estate markets.

[2] Such policies may include more generous child support payments for single parents, subsidized childcare, but also programs directed at promoting women's career and income progression (e.g. female quotas).

optimally without any room for welfare improvements. In contrast, if women face different constraints than men, removing these constraints can change women's perception about their lifetime income trajectory or future optimal consumption and subsequently result in more risky investments and faster wealth accumulation in expectation.

A similar argument applies to the correct cost-benefit evaluation of such policies. For example, the impact of a policy that aims at closing the gender wage gap on female wealth accumulation gets amplified by encouraging women to invest in more risky assets that pay on average higher returns. Hence, the implementation of such a policy may be less costly than previously assumed as it generates higher (capital) tax revenues and further weakens women's dependence on government transfers, in particular during old age.

Related Literature. This paper contributes to several strands of the literature. First, it adds to a literature documenting gender differences in investment behavior and financial choices. There is widespread consensus that women invest less risky than men. Jianakoplos and Bernasek (1998) document lower equity shares among single women than single men in U.S. data. Sunden and Surette (1998) and Agnew et al. (2003) show that women in the U.S. choose lower equity allocations in retirement saving plans. Arano et al. (2010) cannot confirm significant gender differences in retirement accounts for U.S. single households but do so for married individuals. Barber and Odean (2001) find that single men trade more often in risky assets and attribute this result to male overconfidence. Säve-Söderbergh (2012) documents that even though women do not exclude stocks more frequently from their pension contribution plan, they allocate a smaller share into risky assets. Almenberg and Dreber (2015) show that the gender investment gap in Sweden prevails until today. Ke (2018) attributes cross-country differences in stock market participation rates to gender norms, showing that countries with strong gender norms exhibit lower female participation rates. Moreover, several papers document that women earn lower returns in real estate markets (Andersen et al. 2021; Girshina et al. 2022; Goldsmith-Pinkham and Shue 2023). My paper adds to this literature by being the first work to analyze the gender investment gap through the lens of a structural framework.

Second, I relate to an experimental literature which finds that women choose less risky portfolio allocations in investment games (Eckel and Grossman 2008; Croson and Gneezy 2009; Charness and Gneezy 2012) as well as to survey evidence documenting that women rate their willingness to take risk lower than men, even after controlling for a wide range of observable characteristics (e.g. Dohmen et al. 2011). Both findings can lead to the conclusion that women are more risk averse than men. At first, my results seem to contradict this literature because my model replicates the gender investment gap without introducing gender heterogeneity in risk aversion. However, also in the current framework, reduced form regressions on model simulated data that control for observable characteristics fail to fully explain the gender investment gap. The structural analysis then reveals that both current and future deterministic income and household sizes can explain the observed gap in investment choices, rather than innate differences in risk aversion. Hence, my paper confirms prior results on gender heterogeneity in risk-taking, it simply differs in the interpretation of the underlying sources that drive these patterns.

Third, I relate to a literature that explores how family-related shocks affect portfolio allocation and savings. Cubeddu and Ríos-Rull (2003) study the role of marriage and divorce on wealth accumulation in a dynamic setting. Love (2010) is the first paper to present a joint life-cycle framework of marital status and portfolio choice. He finds that married investors hold more risky portfolios than singles. In the event of divorce, stock holdings increase for men whereas they decline for women. Hubener et al. (2015) extend the analysis by incorporating endogenous labor supply and realistically calibrated social security benefits. Christiansen et al. (2015) empirically address the heterogeneous impact of family shocks on portfolio choices across gender using administrative panel data from Denmark. Similar to Love (2010) for the U.S., their findings suggest that the fraction of risky assets in women's portfolios increases after marriage whereas it declines after divorce. For men, this relationship points in the opposite direction. Along the same lines, Bertocchi et al. (2011) find in an empirical framework that the marital gap in stock holdings is larger for women than for men in Italy. While all these papers show that family-related shocks affect portfolio choices heterogeneously across gender, neither of them quantifies the importance of such shocks for gender differences in investment behavior over the life-cycle.

More broadly, my paper extends a literature that studies life-cycle patterns of household finances. For a review of that literature, see Poterba and Samwick (2001) and Gomes (2020). Typically, life-cycle models of portfolio choice predict very high participation rates and equity shares to be strongly declining in age, which is at odds with the data. The literature has proposed several mechanisms to explain this discrepancy. The most prominent ones are costs associated with stock market investment (Vissing-Jorgensen 2002; Gomes and Michaelides 2005; Alan 2006), the illiquid nature of housing (Cocco 2005), lack of financial literacy (Lusardi and Mitchell 2014), or accounting for cyclicality of higher order moments in the income process (Catherine 2022; Shen 2024). However, so far little focus has been on gender and marital status as an additional source of heterogeneity that helps to bring life-cycle models of portfolio choice closer to the data.

Roadmap. The remainder of the paper is structured as follows. Section 2 presents empirical observations on gender-specific portfolio choices. Section 3 introduces the structural model. Section 4 presents the calibration strategy and Section 5 the quantitative results. In Section 6, I analyze the mechanisms that drive the model results. Section 7 performs several robustness checks and Section 8 concludes.

2 The Gender Investment Gap in the Data

In this section, I first describe the data and sample selection criteria. Next, I document empirical patterns of portfolio choices for single men, single women, and couples over their life-cycle.

2.1 The Sample

I use the waves from 1989 until 2016 of the Survey of Consumer Finances (SCF) to measure portfolio choices. The SCF is a triennial repeated cross-sectional survey sponsored by the Federal Reserve Board. For income and demographic characteristics, I work with data from the Panel Study of Income Dynamics (PSID) spanning from 1989 until 2017 (Panel Study of Income Dynamics 2021). The PSID is a longitudinal panel survey of private households in the U.S. running from 1968 until today.[3] It oversamples low-income families (the 'SEO' sample) and immigrant families (the 'immigrant' sample).

I combine two datasets for my analysis because I need detailed portfolio choice information and panel data on household income (to estimate the income processes). While the SCF collects the former, it does not follow the same household over time. In contrast, asset information in the PSID is only reported in some waves and lacks precise information on the portfolio composition of the household. To nevertheless increase confidence in their comparability, I show in Appendix A.1 that life-cycle profiles of variables that are available in both datasets look very similar.[4]

To ensure the representativeness of the U.S. population, I drop all families belonging to the two sub-samples in the PSID and weigh each observation by the provided survey weights in both datasets. I restrict the sample to singles between 30 and 65 years and couples whose average age lies within the same range. All financial variables are converted into 2007 dollars using the CPI-U and winsorized at the 1^{st} and 99^{th} percentile.

I define a single woman to be a family unit with a female head and no spouse present. Single men are defined accordingly. Couples include legally married and cohabiting households. In

[3] Because the Survey of Consumer Finances starts in 1989, I restrict my data sample taken from the PSID to the waves from 1989 until 2017. Data was collected annually until 1997 and afterward every two years.

[4] Combining multiple datasets to estimate structural models is not uncommon in the literature. For instance, Cooper and Zhu (2016) combine the PSID and SCF to estimate the effect of education on stock market investment. Borella et al. (2023) use the PSID together with the Health and Retirement Study (HRS) to study lifetime outcomes during working life and retirement.

total, the PSID sample consists of 100,851 individual-year observations (82,558 for couples, 7,069 for single men, and 11,224 for single women) that correspond to 2,099 unique single women, 1,622 unique single men, and 11,342 individuals who live in couples. The SCF includes information on 23,496 individuals in couples, 4,089 single men, and 6,155 single women.

2.2 Life-Cycle Profiles of Portfolio Allocation

I define net worth as total assets minus total debt. Total assets include financial assets, real estate, and vehicles. Total debt consists of mortgages, credit card balances, installment loans (e.g. education or vehicle loans), and other forms of debt such as loans against pensions or life insurances. Risky assets include direct stock holdings, corporate and foreign bonds, the fraction of mutual funds that include the former, as well as the fraction of retirement accounts which is invested in stocks.[5]

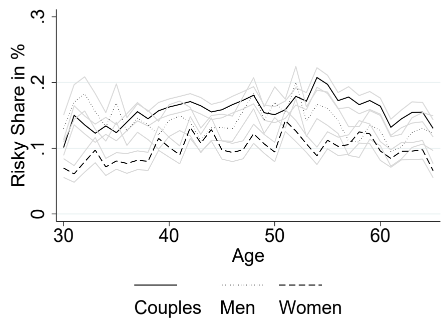
Figure 1a displays the life-cycle profiles of equity shares for single men, single women, and couples.[6] The equity share is defined as the fraction of net worth that is invested in risky assets. It combines the extensive margin (whether or not the household owns any risky assets) with the intensive margin (conditional on holding risky assets, what share of net worth is allocated to them). Figure 1b and Figure 1c separately plot the stock market participation rate and conditional risky share. The gender difference in equity shares is statistically different from zero, in particular during young age, as displayed by the confidence bands and corresponding regression coefficients in Table 1. Moreover, the gender investment gap is primarily driven by lower participation rates of single women, as opposed to a lower risky share conditional on participating (Figures 1b and 1c).

On average, the equity share of single women is around 4%-points lower than that of single

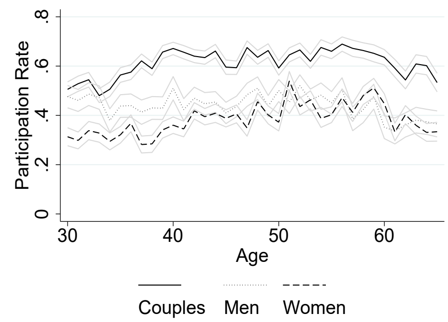
[5] In Appendix A.2, I show that my results are robust to adopting a tighter definition of risky assets that excludes retirement accounts.

[6] To account for cohort effects, Appendix A.2 replicates Figure 1 for individuals born within a relatively short time-frame.

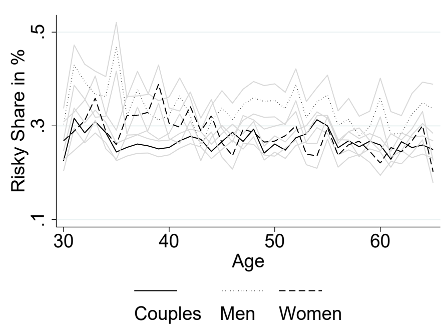
Figure 1: Life-Cycle Patterns of Household Finances (Data)



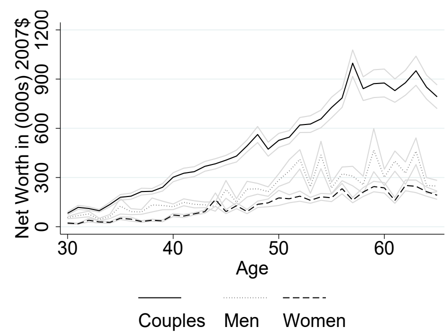
(a) Equity Share



(b) Participation Rate



(c) Conditional Risky Share



(d) Net Worth

Notes: Figure 1 plots the life-cycle profiles of equity shares, stock market participation rates, conditional risky shares, and net worth for singles and couples, including 95% confidence intervals. All figures display averages of the pooled sample by age and household type. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF). Risky assets are defined as direct stock holdings, corporate and foreign bonds, the fraction of mutual funds that include the former, and the fraction of retirement accounts which is invested in stocks.

men which – given an average male equity share of 14% – corresponds to being ca. 30% lower. Both the gender gap in equity shares and stock market participation rates converge toward the entry of retirement.

As indicated by the black solid line in Figure 1a, couples have on average a higher equity share than singles which is driven by the extensive margin (see the black solid lines in Figures 1b and 1c, respectively). Finally, Figure 1d shows that single women accumulate less net worth than single men. This gap is often referred to as the "gender wealth gap". Throughout their working life, the gender gap in net worth is on average around $90,000 and diverges as households grow older.

2.3 Regression Coefficients over the Life-Cycle

The empirical gender differences in portfolio choices reported in Figure 1 can arise due to differences in observable characteristics or due to unobserved factors such as preferences. As a first exercise to quantify the importance of the former, Table 1 reports the results of reduced form regressions that control for household observable characteristics. I run Tobit regressions (to account for non-participating households) of the equity share on a gender dummy, age polynomials, and gender interacted with age (Column (1)). In Column (2), I additionally control for observable characteristics that have shown to be important predictors of portfolio choices. Following Christelis et al. (2013), I control for the education of the individual, the overall number of household members, the inverse hyperbolic sine transformation of non-asset income, and year fixed-effects. Non-asset income includes labor earnings, social security benefits, welfare payments, income from unemployment or worker's compensation, as well as child support and alimony payments. Column (3) furthermore includes the inverse hyperbolic sine transformation of safe assets, which are defined as all net worth that is not invested in risky assets. Column (4) additionally controls for occupation and industry fixed effects. The last three rows of Table 1 report the marginal effects of being a single woman on the equity share along with their standard errors at different ages.[7]

The coefficient for being a single woman is negative and statistically significant across all specifications and becomes smaller as I include more controls. Similarly, the interaction term of gender and age is largest in the first column (least controls) and declines across columns. The marginal effect of being a single woman ("ME") at age 30 is negative and significant across all four columns. As individuals age, this "negative" effect of being a woman on the equity share becomes smaller and, for some specifications, turns insignificant.

Thus, the unexplained part of the gender investment gap (i.e., the part that is not accounted for by household observable characteristics) is strongest among young households and declines

[7] Appendix A.4 lists the corresponding specifications for the participation rate and conditional risky share.

along the life-cycle. To further explore which factors are driving this unexplained part and to quantify their relative importance, Section 3 builds a structural model of gender and portfolio choice. Having a structural model helps to accommodate non-linearities and account for factors that cannot be easily controlled for in reduced form specifications, such as future outcomes of variables.

3 A Life-Cycle Model of Portfolio Choice

This section develops a stochastic life-cycle model with women and men (denote gender by $i = \{f, m\}$) who live either as singles ($\mathcal{S}$) or married couples ($\mathcal{M}$). Life is split into two stages: working age and retirement. Time is discrete and the model period is one year. Agents start their life at age 30, retire at 65, and live at most until age 85, i.e. $j \in \{30, 31, ..., 65, ..., 85\}$. At age 30, agents are ex-ante heterogeneous in terms of education θ which can take two values ($\theta = \{l, h\}$) and refers to having at least 12 years of schooling or not in the data.

During working age, households are subject to uninsurable labor income shocks that depend on their gender, marital status, and the aggregate state of the economy. When being single, individuals decide how much to consume (c_i), how much to save in a safe asset (a_i^s), and how much to save in a risky asset (a_i^r).[8] Borrowing and short-selling are not allowed. While the risky asset pays an equity premium, its return is uncertain and varies with the aggregate state. Couples decide jointly on their level of consumption ($c_{\mathcal{M}}$) and how much to save in both assets ($a_{\mathcal{M}}^s, a_{\mathcal{M}}^r$). Singles face an exogenous marriage probability that depends on their gender, age, and education. Likewise, couples face an exogenous divorce probability that varies by age and both spouses' education.

During retirement, agents face age- and gender-dependent medical expenditures and are subject to longevity risk. Upon dying, they value leaving bequests. They receive a flat pension that depends on their last income during working age. They can live as singles or

[8] I abstract from modeling housing explicitly. See Appendix B for details.

Table 1: Regression Coefficients & Marginal Effects – Equity Shares of Singles

	(1) Equity Share	(2) Equity Share	(3) Equity Share	(4) Equity Share
single woman	-0.3359$^\diamond$	-0.2151$^\diamond$	-0.1671$^\diamond$	-0.1239$^\diamond$
	(0.0131)	(0.0113)	(0.0123)	(0.0140)
single woman*age	0.0052$^\diamond$	0.0032$^\diamond$	0.0026$^\diamond$	0.0015$^\diamond$
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
age	-0.0995$^\diamond$	-0.0508$^\diamond$	-0.0483$^\diamond$	-0.0283
	(0.0135)	(0.0146)	(0.0165)	(0.0185)
$age^2 * 100$	0.2457$^\diamond$	0.1357$^\diamond$	0.1188$^\diamond$	0.0739
	(0.0292)	(0.0316)	(0.0356)	(0.0408)
$age^3 * 10000$	-0.1947$^\diamond$	-0.1123$^\diamond$	-0.0968$^\diamond$	-0.0592$^\diamond$
	(0.0206)	(0.0222)	(0.0249)	(0.0291)
high education		0.2698$^\diamond$	0.2469$^\diamond$	0.1309$^\diamond$
		(0.0020)	(0.0019)	(0.0042)
number of HH members		-0.0556$^\diamond$	-0.0558$^\diamond$	-0.0509$^\diamond$
		(0.0018)	(0.0020)	(0.0030)
non-asset income		0.0324$^\diamond$	0.0288$^\diamond$	0.0257$^\diamond$
		(0.0007)	(0.0006)	(0.0007)
safe assets			0.0218$^\diamond$	0.0146$^\diamond$
			(0.0012)	(0.0009)
constant	1.2163$^\diamond$	-0.0393	-0.1261	-0.4092
	(0.2027)	(0.2168)	(0.2398)	(0.2749)
Observations	10,244	10,240	10,240	7,607
Year FE	No	Yes	Yes	Yes
Industry & Occupation FE	No	No	No	Yes
ME for women at age 30	-0.1814$^\diamond$	-0.1179$^\diamond$	-0.0892$^\diamond$	-0.0790$^\diamond$
	(0.0065)	(0.0062)	(0.0064)	(0.0075)
ME for women at mean age (47)	-0.0919$^\diamond$	-0.0616$^\diamond$	-0.0441$^\diamond$	-0.0551$^\diamond$
	(0.0032)	(0.0036)	(0.0035)	(0.0042)
ME for women at age 65	-0.0011	-0.0045	0.0016	-0.0267$^\diamond$
	(0.0031)	(0.0028)	(0.0032)	(0.0018)

Notes: Estimations are based on Tobit regressions on the sample of individuals who live in households with no spouse present. Source: SCF waves 1989 until 2016. Equity Share = Unconditional risky share, with the risky share being defined as the fraction of net worth that is invested in risky assets. *single woman* is a dummy indicating that the household head is a woman. *high education* is a dummy equal to one if the household head has more than 12 years of education. *safe assets* refers to all net worth that is not invested in risky assets. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

couples and their marital status is fixed. If one spouse living in a couple dies, the surviving spouse continues his or her life as a single with a fraction of the couple's assets. As before, households have a portfolio choice between a safe and risky asset.

3.1 Preferences

All households have time-separable CRRA preferences over a consumption good c. The period flow of utility for singles and couples is given by:

$$\text{Singles:} \quad u(c) = \frac{\eta_{ij}\left(\frac{c}{\eta_{ij}}\right)^{1-\gamma}}{1-\gamma} \qquad \text{Couples:} \quad u(c) = \frac{\eta_{\mathcal{M}j}\left(\frac{c}{\eta_{\mathcal{M}j}}\right)^{1-\gamma}}{1-\gamma} \tag{1}$$

where γ is the coefficient of relative risk aversion and η is an equivalence scale that adjusts for household size. The term η is allowed to vary by age j and family type (couple, single man, single woman).

Bequest Motive. Upon dying, households derive utility from leaving bequests as in De Nardi (2004):

$$\phi(a') = L\frac{(\omega + a')^{1-\gamma}}{1-\gamma} \tag{2}$$

where a' denotes the bequeathed assets, ω captures the luxuriousness of the bequest motive, and L governs the bequest intensity. Couples value leaving bequests if they both die within the same period. Whenever only one spouse dies, the surviving spouse continues life as a single and values leaving bequests in the case of his or her own death.

3.2 Dynamics

Aggregate State. The economy is characterized by an aggregate state Ω that governs labor market conditions and stock returns, as explained in detail below. The aggregate state can take two values, referring to booms and recessions ($\Omega = \{b, r\}$). Recessions realize with probability p_{rec} and are not persistent.

Asset Returns. The safe asset pays a time-invariant return r_s. The return of the risky asset depends on the aggregate state Ω and follows a mixture normal distribution:

$$r_r = \begin{cases} \mathcal{N} \sim (\mu_r^{\text{boom}}, \sigma_r^2) & \text{if } \Omega = b \text{ with probability } (1 - p_{\text{rec}}) \\ \mathcal{N} \sim (\mu_r^{\text{rec}}, \sigma_r^2) & \text{if } \Omega = r \text{ with probability } p_{\text{rec}} \end{cases} \tag{3}$$

In expectations, the risky asset pays an equity premium and it holds that $\mu_r^{\text{boom}} > \mu_r^{\text{rec}}$.

Income Profiles. The income profiles vary between singles and couples. Conditional on marital status, income y_{ij} at age j for gender i can be split into a deterministic and stochastic component and is expressed as:

$$y(i, j, \theta, \tilde{y}(\Omega)) = \bar{y}_i \theta_i \xi_{ij} \tilde{y}_{ij}(\Omega) \tag{4}$$

The term $\bar{y}_i$ denotes a constant, θ_i is the education premium, and ξ_{ij} stands for an age-specific component. The term $\tilde{y}_{ij}(\Omega)$ represents the stochastic component of income and consists of a transitory and persistent part, with the latter depending on the aggregate state Ω:

$$\tilde{y}_{ij} = z_{ij} + \epsilon_{ij} \quad \text{with} \quad z_{ij} = \begin{cases} \mu_z^{\text{boom}} + \rho_i z_{i,j-1} + \nu_{ij} & \text{if } \Omega = b \text{ with probability } (1 - p_{\text{rec}}) \\ \mu_z^{\text{rec}} + \rho_i z_{i,j-1} + \nu_{ij} & \text{if } \Omega = r \text{ with probability } p_{\text{rec}} \end{cases} \tag{5}$$

The terms ϵ_{ij} and ν_{ij} are independent zero mean random shocks with variances $\sigma_{\epsilon i}^2$ and $\sigma_{\nu i}^2$, respectively. The parameter $\rho_i \in (0, 1]$ captures the persistence of shock ν_{ij}. To keep the process stationary, I impose that $\mu_z^{\text{boom}} = \left(\frac{-p_{\text{rec}}}{1 - p_{\text{rec}}}\right) \mu_z^{\text{rec}}$. Neither μ_z^{boom} nor μ_z^{rec} depends on gender or marital status.

Within couples, the transitory shocks ϵ_{fj} and ϵ_{mj} are assumed to be correlated with $\rho_{\sigma_{\epsilon f}, \sigma_{\epsilon m}} = 0.3$.[9] Spouses live in the same area, are likely to work in similar industries, and may thus be

[9] By setting the correlation to 0.3, I follow Borella et al. (2023) who estimate an empirical correlation between initial wage draws for newly formed couples in U.S. data of 0.22 for the age group 25-34, 0.36 for ages

subject to correlated labor market shocks. Lastly, I follow Huggett and Kaplan (2016) and impose a flat labor income tax rate of $\tau = 0.27$.

Out-of-Pocket Medical Expenditures. When being retired, agents are subject to medical expenditures m_{ij} that are a deterministic function of age and gender. Because individuals face survival risk and medical expenditures are strictly increasing in age, deterministic medical expenditures impose a risk in that agents are uncertain whether they survive until the age when the medical bills are due.

Marriage and Divorce. Singles get married with an exogenous probability $\mu(i, j, \theta)$ that depends on their gender i, age j, and education θ. Conditional on meeting a partner p, the probability of meeting a partner with education θ_p and stochastic income realization $\tilde{y}_p$ is defined as $\Pi(.) = \Pi(\theta_p, \tilde{y}_p | \theta_i, \tilde{y}_i)$. Both partners have the same age. Individuals are always matched to a partner with the mean empirical amount of assets conditional on age, gender, and education. This specification generates assortative mating along asset holdings close to the data. Couples face an exogenous divorce probability $\lambda(j, \theta_f, \theta_m)$ that depends on age and both spouses' education. Upon divorce, assets are split equally with 10% being exogenously destroyed.[10] There are no alimony payments.

3.3 Stock Market Participation Cost

Agents have to pay a fixed cost S^F_j each period if they choose to invest part of their savings in the risky asset. I allow for the possibility that this cost varies by age j. For example, as Catherine (2022) notes, young households may face lower costs because they are automatically enrolled in retirement plans or tend to be more financially literate. Moreover, following Vissing-Jorgensen (2002), participation costs have to be paid each period irrespective of the history of stock holdings. The main advantage of modeling participation costs as a flow variable rather than an entry cost (see e.g. Alan (2006) or Cooper and Zhu (2016)) is that

35-44, and 0.42 for couples above 45 years.

[10] This splitting rule is motivated by the data. In the PSID, the median financial wealth of singles one period after a divorce is 45% of the former couple's wealth, regardless of gender.

flow costs do not require introducing stock holdings as a state variable.

3.4 Timing

At the beginning of period t, households learn their stochastic income realization (which depends on the aggregate state), marital status, and asset level (which depends on their stock market return and the aggregate state). After observing all shock realizations, households decide on how much to consume and save in the risky and safe asset, forming expectations over future shocks. When investing part of their savings in the risky asset, they have to pay the stock market participation cost S_j^F in the current period t.

3.5 Recursive Formulation

I express the problem recursively by defining six value functions: the value function for singles, the value function for couples, and the value function for an individual living in a couple, each during working age and retirement. The latter is the relevant object when computing the present value of marriage for a single whereas the value function for couples determines the optimal allocation of resources within couples across time (Borella et al. 2020). Because the stock market participation cost has to be paid per period and given the i.i.d. nature of the risky asset and aggregate state, I can combine labor income with safe and risky assets into one "cash-on-hand" variable: $a = (1-\tau)y(i,j,\theta,\tilde{y}(\Omega)) + (1+r_r(\Omega))a_r + (1+r_s)a_s$.

Singles – Working Age. The state variables of a single agent are their gender i, age j, education θ, cash-on-hand a, and current stochastic income realization $\tilde{y}$. Their value function can be expressed as:

$$\begin{aligned} V^S(i,j,\theta,a,\tilde{y}) = \max_{a'_s \geq 0, a'_r \geq 0, c \geq 0} \frac{\eta_{ij}\left(\frac{c}{\eta_{ij}}\right)^{1-\gamma}}{1-\gamma} &+ (1-\mu(i,j,\theta))\beta \mathbb{E} V^S(i,j+1,\theta,a',\tilde{y}') \\ &+ \mu(i,j,\theta)\beta \mathbb{E} \hat{V}^C(i,j+1,\theta,\theta_p,a'+a'_p,\tilde{y}',\tilde{y}'_p) \quad (6) \end{aligned}$$

subject to:

$$a'_r + a'_s + c = a - \mathbb{1}_{a'_r > 0} S_j^F \quad \text{with} \quad a = (1-\tau) y_s(i, j, \theta, \tilde{y}(\Omega)) + (1 + r_r(\Omega)) a_r + (1 + r_s) a_s \tag{7}$$

The labor income and risky return processes are defined in Section 3.2. The term η_{ij} denotes an equivalence scale that accounts for changing household sizes over the life-cycle. $\hat{V}^C$ expresses the value for individual i of getting married to partner p. Single individuals take the expected value over their stochastic income realization, asset return, and the aggregate state when staying single whereas they form expectations over their stochastic income realization, asset return, the aggregate state, and their specific partner in case of getting married.

Singles – Retirement. The state variables of a retired single are their gender i, age j, education level θ, cash-on-hand a, and the last income realization before retirement ($\hat{y}$).

$$V_R^S(i, j, \theta, a, \hat{y}) = \max_{a'_s \geq 0, a'_r \geq 0, c \geq 0} \frac{\eta_{ij} \left(\frac{c}{\eta_{ij}} \right)^{1-\gamma}}{1-\gamma} + \beta \psi_{ij} \mathbb{E} V_R^S(i, j+1, \theta, a', \hat{y}) + \beta (1 - \psi_{ij}) L \frac{(\omega + a')^{1-\gamma}}{1-\gamma} \tag{8}$$

subject to:

$$a'_r + a'_s + c = a - m_{ij} - \mathbb{1}_{a'_r > 0} S_j^F \quad \text{with} \quad a = (1-\tau) pen_s(\hat{y}) + (1 + r_r(\Omega)) a_r + (1 + r_s) a_s \tag{9}$$

The return process for the risky asset is defined in Section 3.2. The terms ψ_{ij} and m_{ij} denote age- and gender-dependent survival probabilities and medical expenditures. Retired singles take the expected value over asset returns, the realization of the aggregate state, and their likelihood of survival.

Couples – Working Age. The state variables of a couple that consists of a woman f and man m are their age j, education of both spouses θ_f, θ_m, joint cash-on-hand a, and both stochastic income realizations $\tilde{y}_f, \tilde{y}_m$. Their value function reads as:

$$
\begin{aligned}
V^C(j,\theta_f,\theta_m,a,\tilde{y}_f,\tilde{y}_m) = \max_{a'_s\geq 0, a'_r\geq 0, c\geq 0} & \frac{\eta_{\mathcal{M}j}\left(\frac{c}{\eta_{\mathcal{M}j}}\right)^{1-\gamma}}{1-\gamma} \\
& + (1-\lambda(j,\theta_f,\theta_m))\beta\mathbb{E}V^C(j+1,\theta_f,\theta_m,a',\tilde{y}'_f,\tilde{y}'_m) \\
& + \lambda(j,\theta_f,\theta_m)\beta\sum_{i=f,m}\mathbb{E}V^S(i,j+1,\theta_i,0.9\frac{a'}{2},\tilde{y}'_i)
\end{aligned} \tag{10}
$$

subject to:

$$
a'_r + a'_s + c = a - \mathbb{1}_{a'_r>0}S^F_j \quad \text{with} \quad a = \sum_{i=f,m}(1-\tau)y_c(i,j,\theta,\tilde{y}(\Omega)) + (1+r_r(\Omega))a_r + (1+r_s)a_s \tag{11}
$$

Couples take the expected value over both partners' stochastic income realizations, the aggregate state, and asset returns when staying married, and the respective individual's stochastic income realizations, asset returns, and the aggregate state when getting divorced. The labor income and risky asset return processes are defined in Section 3.2.

Couples – Retirement. The value function of a retired couple is:

$$
\begin{aligned}
V^C_R(j,\theta_m,a,\hat{y}_m) = & \max_{a'_s\geq 0, a'_r\geq 0, c\geq 0} \frac{\eta_{\mathcal{M}j}\left(\frac{c}{\eta_{\mathcal{M}j}}\right)^{1-\gamma}}{1-\gamma} + \beta\psi_{jf}\psi_{jm}\mathbb{E}V^C_R(j+1,\theta_m,a',\hat{y}_m) \\
& + \beta\sum_{i=f,m}\psi_{ij}(1-\psi_{-ij})\mathbb{E}V^S_R(i,j+1,\theta_m,\delta_i a',\hat{y}_m) \\
& + \beta(1-\psi_{jf})(1-\psi_{jm})L\frac{(\omega+a')^{1-\gamma}}{1-\gamma}
\end{aligned} \tag{12}
$$

subject to:

$$
a'_r + a'_s + c = a - \sum_{i=f,m} m_{ij} - \mathbb{1}_{a'_r>0}S^F_j \quad \text{with} \quad a = (1-\tau)pen_c(\hat{y}_m) + (1+r_r(\Omega))a_r + (1+r_s)a_s \tag{13}
$$

Retired couples take the expected value over the aggregate state, asset returns, and individual survival probabilities, with the return process for the risky asset being defined in Section 3.2.

Value to an individual of becoming a couple. The value of an individual in a couple is the present discounted value of the individual's utility in the event of marriage (Borella et al. 2020). It is needed to compute the value of single i for getting married to partner p. Variables denoted with a $\hat{hat}$ express optimal allocations computed with the value function for couples, given the respective state variables. The value of an individual in a retired couple $\hat{V}_R^C$ is defined accordingly.

$$\hat{V}^C(i,j,\theta_i,\theta_p,a,\tilde{y}_i,\tilde{y}_p) = \frac{\eta_{\mathcal{M}j}\left(\frac{\hat{c}}{\eta_{\mathcal{M}j}}\right)^{1-\gamma}}{1-\gamma} + (1-\lambda(j,\theta_i,\theta_p))\beta\mathbb{E}\hat{V}^C(i,j+1,\theta_i,\theta_p,a',\tilde{y}_i',\tilde{y}_p') + \lambda(j,\theta_i,\theta_p)\beta\mathbb{E}V^S(i,j+1,\theta_i,0.9\frac{a'}{2},\tilde{y}_i') \tag{14}$$

4 Estimation & Calibration

As standard in the literature, I estimate and calibrate the model in two steps (Gourinchas and Parker 2002; Cagetti 2003). First, I estimate all parameters that can be cleanly identified from the data and pre-set some parameters to values from the literature. In the second step, I calibrate the remaining structural parameters using the Simulated Method of Moments (SMM), taking the parameters from the first stage as given.

4.1 First Stage Estimation

Income Profiles. Figure 2 shows the life-cycle profiles of the deterministic income component by gender and marital status from the PSID. Income is expressed as annual income out of labor earnings (including income from farms and businesses), social security benefits, and transfers (including child support and alimony payments). For singles, I include labor earnings, social security benefits, and transfers from all members of the household. For couples, I assign each spouse their own labor income, social security benefits, transfers, and add half of that from other household members.[11] Moreover, I winsorize the top and bottom percentile of earnings and drop observations who, according to the described measure, report zero annual income (in the case of couples, if they report zero overall income).

I follow Borella et al. (2020) and first split the sample by marital status and then separately regress the inverse hyperbolic sine of income for an individual of gender i at age j,

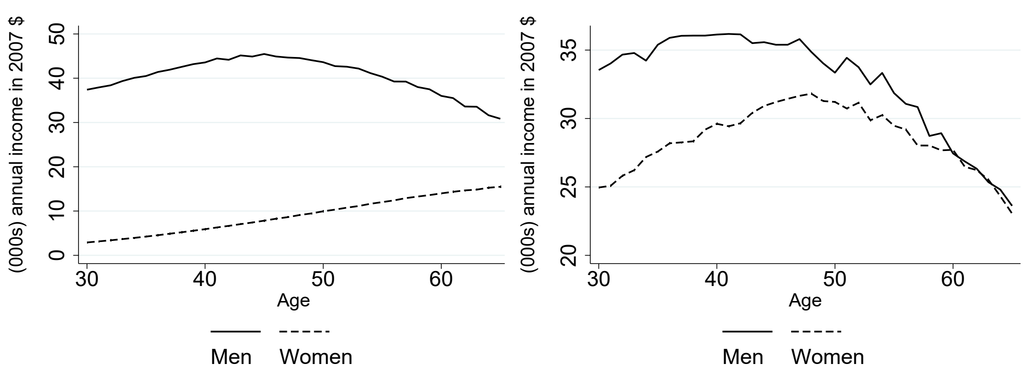
$$income_{ij} = \alpha + \beta_1 age_{ij} + \beta_2 age_{ij}^2 + \beta_3 woman_i * age_{ij} + \delta_i + u_{ij} \tag{15}$$

on a fixed effect δ_i, age, age^2, and an interaction term of gender and age. To obtain shifters for gender and education, I regress the sum of the fixed effect and residual on fully interacted dummies of gender and education:

$$\delta_i + u_{ij} = \gamma_0 + \gamma_1 woman_i + \gamma_2 educ_i + \gamma_2 woman_i * educ_i + w_{ij} \tag{16}$$

where $educ_i$ is a dummy taking the value one if the respective individual has more than 12 years of schooling.

Figure 2: Life-Cycle Profiles of the Deterministic Income Component



(a) Couples (b) Singles

Notes: Figure 2 plots the life-cycle profiles for the deterministic income component by gender and marital status. Data is from the waves 1989 until 2016 of the Panel Study of Income Dynamics (PSID).

[11] In some waves, the PSID does not report transfer income and social security benefits separately for spouse and household head. In these cases, I allocate half of the overall reported measure to the wife and the other half to the husband.

The coefficients from these income equations (reported in Table 11 in Appendix C.2) inform me about the deterministic income component in the model. Note that the estimated age gradients are partly driven by variation in hours worked and transitions in and out of the labor force, as opposed to differences in wages. For example, Borella et al. (2023) document that average hours worked of single women between age 30 and 45 grow faster than those of single men, whereas single men are more likely to drop out of the labor force beyond age 45, contributing to the documented life-cycle patterns for single households in Figure 2b.

I estimate the parameters governing the stochastic income component using the minimum distance estimator from Guvenen (2009).[12] Table 2 summarizes the results. My point estimates imply a slightly less persistent income process for single women than for single men, whereas the variance of both the persistent shock σ_ν^2 and transitory shock σ_ϵ^2 is lower for single women. That is, the overall variance of single women's income process is lower than that of single men's, which may for example arise due to single women sorting into more stable occupations (Bertrand 2011). When solving the model, I discretize the income processes using the Rouwenhorst method (Rouwenhorst 1995).

I set the mean of the persistent part of the stochastic income component during recessions (μ_z^{rec}) to -0.09. Thereby, I ensure that the correlation between stock market and human capital returns is positive but does not exceed estimates from previous literature, in particular Huggett and Kaplan (2016).[13] Moreover, by setting $\mu_z^{\text{rec}} = -0.09$ the ex-post correlation between labor income shock realizations and risky asset returns is around 0.16, which aligns with Shen (2024) who introduces a correlation between innovations to permanent income shocks and stock returns of 0.15.

Marital Transitions. Marital transitions are defined as the likelihood of getting married

[12] Details on the estimation strategy can be found in Appendix C.1. When estimating the stochastic part of the income process, I drop individuals who report zero income to avoid unrealistically high estimates for the income volatility, in particular among married women.

[13] See Appendix C.1 for details on the definition and estimation of human capital returns and their correlation to the risky asset.

Table 2: Estimation Results – Stochastic Income Component

Parameter	Men	Women	Men	Women
	Singles		Couples	
ρ	0.9522	0.9348	0.9353	0.9273
	(0.0079)	(0.0095)	(0.0046)	(0.0045)
σ_ν^2	0.0899	0.0856	0.0809	0.1583
	(0.0115)	(0.0118)	(0.0051)	(0.0092)
σ_ϵ^2	0.1689	0.1547	0.1412	0.2851
	(0.0341)	(0.0217)	(0.0109)	(0.0177)

Notes: Standard Errors in parentheses obtained with bootstrapping (2000 replications).

(respectively divorced) within the next period conditional on not being married (respectively being married) in the current period. I estimate the following logit function, separately for couples and singles:

$$\zeta_{h,j+1} = \frac{exp(\boldsymbol{x}'_{hj}\boldsymbol{\beta^s})}{1 + exp(\boldsymbol{x}'_{hj}\boldsymbol{\beta^s})} \tag{17}$$

where $\zeta_{h,j+1}$ denotes the probability for household h in period j of being married (respectively divorced) next period. As explanatory variables ($\boldsymbol{x}$), I include gender (for the marriage probability), age, education, and a dummy for waves after 1997 to account for the switch from annual to biannual frequency in the PSID. Table 12 in Appendix C.3 reports the corresponding regression coefficients ($\boldsymbol{\beta^s}$). Finally, I estimate the matching of spouses in terms of income and education (summarized by the term Π) non-parametricly from the PSID.

Out-of-Pocket Medical Expenditures. I take the parameters for medical expenditures by age and gender from Borella et al. (2020) who estimate deterministic out-of-pocket medical expenditures from HRS data. They find higher medical expenditures for men at the start of retirement but a steeper gradient for women, especially after age 76.

Survival Probabilities. I take gender-specific death probabilities from the Life Tables of the U.S. Social Security Administration.[14] The death probability at age j is defined as the

[14] All tables are available under this link.

probability to die within the next year conditional on having survived up to age j. I compute the inverse of those probabilities and take average values between the years 1990, 2000, and 2010, corresponding to my sample period. For couples, if the husband dies, the surviving wife keeps 60% of the household's assets, whereas a surviving husband keeps 70% of the household's assets to account for sharply increasing medical expenses in the year prior to death, as well as for bequests to non-spousal heirs.[15]

Asset Returns. The annual return rate of the risk-free asset is 2%, taken from Catherine (2022). During booms, the return of the risky asset is drawn from a normal distribution with mean $\mu_r^{\text{boom}} = 11.5\%$ and variance $\sigma_r^2 = (0.1758)^2$. During recessions, the mean return to the risky asset is $\mu_r^{\text{rec}} = -24.5\%$. A recession occurs each period with probability $p_{\text{rec}} = 14.6\%$. Hence, in expectations, the equity premium is 5%. The values for p_{rec}, μ_r^{boom}, and μ_r^{rec} are again taken from Catherine (2022) whereas σ_r^2 reflects the variance of the annual total return index of the S&P 500 from 1989 until 2016.

Pension Payments. Pension payments are flat and 60% of the medium persistent income realization at age 65. To ensure that on average pensions are correlated with lifetime income, I assume that individuals in the lowest productivity state at 65 receive 60% of the income that corresponds to the lowest transitory realization of the medium persistent state, singles with a medium productivity receive a pension corresponding to the medium transitory realization of the medium persistent state and so forth. Couples receive a common pension that is 1.7 times higher than that of single men.

Equivalence Scales. To calibrate the equivalence scales η, I compute the average number of household members by age and family type from the PSID and then apply the OECD equivalence scale: I assign a weight of 1 to the first adult household member, a weight of 0.7 to all other adult members, and a weight of 0.5 to each child.[16]

[15] I choose these values because Jones et al. (2020) document that households who experienced the death of one spouse have around 30% lower wealth than couples who did not experience a death, and that surviving single men are on average wealthier than surviving single women.

[16] Figure 14 in Appendix C.4 plots the resulting predicted household sizes by family type over the life-cycle.

Initial Conditions. The initial distribution over asset holdings in the model mimics the distribution of net worth across individuals at age 30 in the SCF. Similarly, the initial distribution of couples and singles, as well as the matching of spouses across education levels is taken from individuals at age 30 in the PSID. Finally, I set the fraction of high- and low-educated individuals by gender to be the average share of individuals with more or less than 12 years of schooling in the PSID.

4.2 Second Stage Calibration

I borrow the bequest parameters from Cooper and Zhu (2016) which results in $L = 0.128$ and $\omega = 0.73$. To impose structure on the age-profile in the stock market participation costs (S_j^F), I assume that the highest participation cost (i.e., at age 65) is \$300 above the cost for young households. Moreover, the costs start to increase linearly from age 45 onward. These restrictions leave me with one participation cost parameter to calibrate internally (S_{young}^F). Taking the parameters from the first stage as given, I then calibrate the remaining parameters $\Theta = \{\beta, \gamma, S_{\text{young}}^F\}$ using the Simulated Method of Moments. The exercise is to find $\hat{\Theta}$ that solves the following optimization problem:[17]

$$\mathcal{L} = \min_{\Theta} \left(\frac{M^s(\Theta) - M^d}{M^d} \right) W \left(\frac{M^s(\Theta) - M^d}{M^d} \right)' \tag{18}$$

where W represents a weighing matrix, M^d are the moments derived from the data and $M^s(\Theta)$ their theoretical counterparts derived from model simulations. I take the relative deviation of simulated moments from their data targets as input in the objective function to account for different units (%-points vs. \$ values) across empirical moments.

Parameter Identification & Choice of Moments. I exploit heterogeneity in net worth to identify the discount factor β. Once households cross the threshold of stock market participation, the participation cost becomes irrelevant for their decision on how much to

[17] Given the computational complexity of the framework, I can only solve the model for a limited number of iterations. See Appendix D.1 for further details.

invest in the risky asset. Taking this discrepancy into account, I identify the coefficient of risk aversion γ by exploiting heterogeneity in the portfolio share conditional on participating. The stock market participation cost S^F_{young} serves as the parameter of interest to match participation rates. I target the life-cycle profiles of single men and single women, resulting in 216 moments (36 years $\times$ 3 variables $\times$ 2 household types), with only three parameters.

The Weighting Matrix W. I first calibrate the second-stage parameters by using a slightly modified identity matrix ($W = \mathcal{I}$). Given the paper's focus on equity share, I place less weight (30%) on asset profiles than on participation rates and conditional risky shares. In a second run, I follow Cooper and Zhu (2016) and use the inverse of the variances of my moment conditions as a (diagonal) weighting matrix to assign a lower weight to less precisely estimated data moments ($W = \frac{1}{\mathcal{V}}$).

5 Quantitative Results

5.1 2nd Stage Parameters

Table 3 reports the calibrated second-stage parameters. The parameter values across the two specifications are very similar. The calibration with the modified identity matrix ($W = \mathcal{I}$) finds an annual stock market participation cost of $288 for young households which increases to $588 at age 65. Moreover, my estimates suggest a coefficient of relative risk aversion of $\gamma = 5.07$. This value is in line with previous papers of portfolio choice that allow for a small dependence between stock returns and labor income shocks. For example, Shen (2024) estimates a degree of risk aversion of $\gamma = 5.6$. Catherine (2022) finds a CRRA coefficient of $\gamma = 6$ and an annual stock market participation cost of $250 which is close to my results. In addition, my value for the discount factor β (0.917) is well within the range of previous studies. Cooper and Zhu (2016) estimate a discount factor of 0.869, Fagereng et al. (2017) of 0.77, Catherine (2022) of 0.96, and Shen (2024) of 0.98 for stockholders and 0.92 for non-stockholders.

Table 3: 2nd Stage Parameters

W	β	γ	S^F_{young}	S^F_{old}
$\mathcal{I}$	0.917	5.07	\$288	\$588
$\frac{1}{\mathcal{V}}$	0.918	5.31	\$228	\$528

Notes: Table 3 lists the values for internally calibrated model parameters. W denotes the weighting matrix, as explained in Section 4.2.

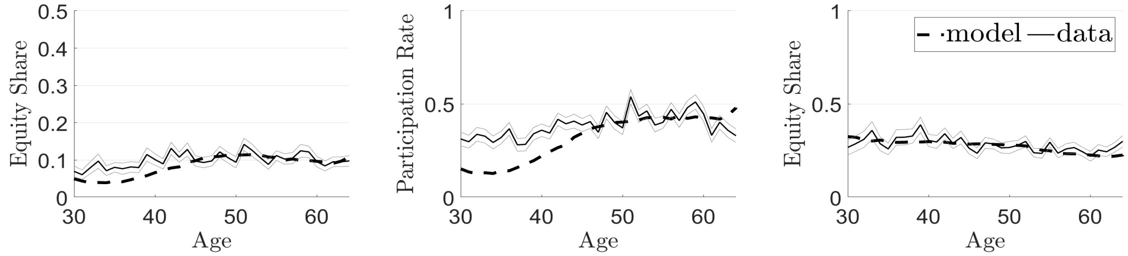
5.2 Model Fit

Life-Cycle Profiles of Household Finances. Figure 3 shows that the model matches well the life-cycle profiles of equity shares, participation rates, and conditional risky shares for both single men and single women. Hence, the model is able to capture the gender investment gap without introducing preference heterogeneity across households. In addition, it closely replicates the evolution of net worth for single men and single women until age 50 (Figure 4). It undershoots the asset accumulation of single men during older ages. Appendix D.2 further discusses the model fit for couple households.

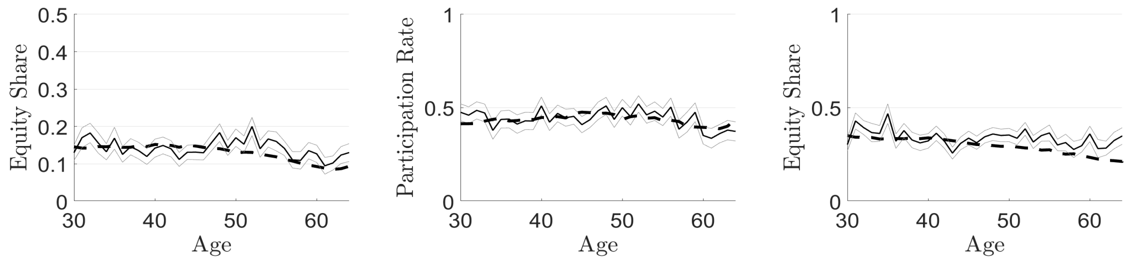
Simulated Regressions. To compare the reduced form regressions from Section 2.3 with the model, Table 4 replicates the regressions from Table 1 with model simulated data. All of these coefficients are untargeted in the calibration exercise.

The model over-predicts the effect of gender on the equity share (Columns (1) and (2) in Table 4), meaning that the absolute values for the coefficient "single woman" and its interaction term with age are larger than in the data. Nevertheless, the model simulated data replicates the negative, but increasing marginal effect ("ME") of being a woman on the equity share over the life-cycle. Thus, reduced form regressions that control for household heterogeneity fail to fully explain the gender investment gap, in particular among young households, even if the underlying data generating process assumes homogeneous preferences across men and women.

Figure 3: Model Fit of Investment Patterns (Singles)



(a) Women: Equity Share (b) Women: Participation (c) Women: Cond. Share



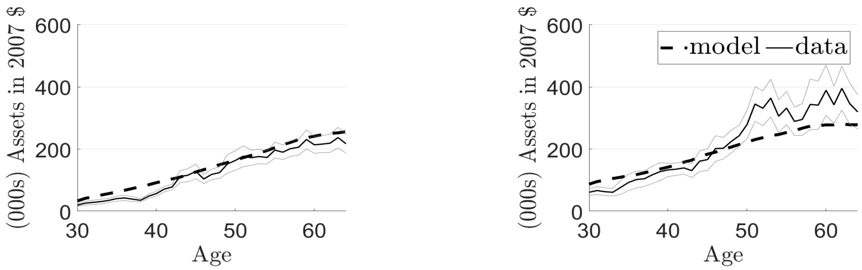
(d) Men: Equity Share (e) Men: Participation (f) Men: Cond. Share

Notes: Figure 3 plots the model fit of equity shares, participation rates, and conditional risky shares for single women and single men. The solid lines show the data (including 95% confidence bands, as plotted in Figure 1) whereas the dashed lines display the simulated life-cycle profiles generated from the model.

Columns (3) and (4) report the estimated gender effect on stock market participation rates. In line with the data, the model predicts the marginal effect of being a woman to be negative and increasing in age. In both data and model, the marginal effect of being a woman on the participation rate accounts for the majority of the overall marginal effect on the equity share. When considering the conditional risky share (Columns (5) and (6)), the model produces a negative baseline effect of being a woman (in line with the data, albeit non significant), and a positive interaction term (opposed to a negative one in the data). Hence, even though the model matches the marginal effect of being a woman on the conditional equity for young households, it predicts, in contrast to the data, an increasing trend over the life-cycle.

As in the SCF, the coefficient for income is positive across all specifications on model simulated data. The model also replicates the positive coefficient for safe assets with regard to the equity share and participation rate, and a negative coefficient in the specification for the conditional risky share. All else equal, higher net worth helps households to pay the stock

Figure 4: Model Fit of Net Worth (Singles)



(a) Single Women (b) Single Men

Notes: Figure 4 plots the model fit of asset accumulation (expressed as net worth) for single women and single men. The solid lines show the data (including 95% confidence bands, as plotted in Figure 1) whereas the dashed lines display the simulated life-cycle profiles generated from the model. Data lines are smoothed using a three-point moving average.

market participation cost. However, conditional on participating, the optimal risky share is slightly decreasing in household wealth, explaining the negative coefficient for safe assets in Column (5).

With model simulated data, the coefficient for age on the unconditional equity share is positive, that of its squared term negative, and that of its cubic term positive which is in contrast to the SCF. The coefficient for high education is positive throughout all six specifications. In the model, education increases deterministic income, making highly educated households more willing (and able) to participate in risky asset markets.

6 Understanding the Mechanisms

As Table 4 shows, reduced form regressions that control for household heterogeneity fail to fully explain the gender investment gap, even if the underlying data generating process assumes homogeneous preference parameters across men and women. To better understand what drives gender heterogeneity in equity shares, this section analyzes gender specific model elements more carefully and explains how they affect households' investment choices.

Table 4: Regression Coefficients & Marginal Effects – Data vs. Model Simulations

	(1)	(2)	(3)	(4)	(5)	(6)
	Equity Share		Participation Rate		Conditional Share	
	Model	Data	Model	Data	Model	Data
single woman	-0.5337◇	-0.1671◇	-0.4328◇	-0.1363◇	-0.0771◇	-0.0195
	(0.0346)	(0.0123)	(0.0366)	(0.0116)	(0.0132)	(0.0170)
single woman*age	0.0092◇	0.0026◇	0.0074◇	0.0026◇	0.0012◇	-0.0008◇
	(0.0007)	(0.0002)	(0.0008)	(0.0002)	(0.0003)	(0.0004)
age	0.1634◇	-0.0483◇	-0.0072	-0.0917◇	0.0663◇	0.0307◇
	(0.0273)	(0.0165)	(0.0296)	(0.0147)	(0.0097)	(0.0118)
$age^2 * 100$	-0.3784◇	0.1188◇	-0.0319	0.2123◇	-0.1517◇	-0.0591◇
	(0.0586)	(0.0356)	(0.0641)	(0.0324)	(0.0206)	(0.0242)
$age^3 * 10000$	0.2741◇	-0.0968◇	0.0460	-0.1611◇	0.1099◇	0.0394◇
	(0.0409)	(0.0249)	(0.0450)	(0.0231)	(0.0143)	(0.0163)
high education	0.0374◇	0.2469◇	0.0235◇	0.2585◇	0.0288◇	0.0547◇
	(0.0074)	(0.0019)	(0.0082)	(0.0018)	(0.0027)	(0.0032)
no. of HH members		-0.0558◇		-0.0407◇		-0.0234◇
		(0.0020)		(0.0008)		(0.0012)
non-asset income	0.0418◇	0.0288◇	0.0390◇	0.0261◇	0.0060◇	0.0038◇
	(0.0036)	(0.0006)	(0.0040)	(0.0004)	(0.0012)	(0.0006)
safe assets	0.1226◇	0.0218◇	0.1848◇	0.0147◇	-0.0574◇	-0.0380◇
	(0.0039)	(0.0012)	(0.0039)	(0.0003)	(0.0013)	(0.0006)
constant	-4.1235◇	-0.1261	-1.6514◇	1.1751◇	0.0384	0.1063
	(0.4178)	(0.2398)	(0.4446)	(0.2160)	(0.1482)	(0.1882)
Observations	10,240	10,240	10,240	10,240	3,900	4,285
Year FE	No	Yes	No	Yes	No	Yes
ME at age 30	-0.2585◇	-0.0892◇	-0.2110◇	-0.0590◇	-0.0418◇	-0.0440◇
	(0.0148)	(0.0064)	(0.0152)	(0.0048)	(0.0058)	(0.0064)
ME at mean age (47)	-0.1026◇	-0.0441◇	-0.0853◇	-0.0143◇	-0.0217◇	-0.0593◇
	(0.0074)	(0.0035)	(0.0081)	(0.0018)	(0.0027)	(0.0026)
ME at age 65	0.0626◇	0.0016	0.0477◇	0.0311◇	-0.0005	-0.0727◇
	(0.0138)	(0.0032)	(0.0159)	(0.0040)	(0.0048)	(0.0072)

Notes: Estimations are based on regressions on the sample of individuals who live in households with no spouse present. Columns "Model" are model simulations, whereas columns "Data" refer to data from the SCF waves 1989 until 2016. Equity Share = Unconditional risky share. Participation Rate = Stock market participation rate. Conditional Share = Conditional risky share. *single woman* is a dummy indicating that the household head is a woman. *high education* is a dummy equal to one if the household head has more than 12 years of education. *safe assets* refers to all net worth that is not invested in risky assets. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, ◇ $p<0.05$.

6.1 Decomposing the Gender Investment Gap

First, I decompose the gender gaps in equity shares and wealth levels along the dimensions of gender heterogeneity in the model, that is along deterministic income and its individual components, stochastic income, marital transition probabilities, the expected characteristics of the partner in the event of marriage (the "marriage market": Π), the distribution of individuals across education levels, initial wealth holdings, the average number of household members (captured by the equivalence scale η), as well as medical expenses and survival probabilities during retirement. In all cases, I replace the female value with that of men and study the resulting gender gaps in asset holdings and equity shares. Table 5 displays the results. The column "Model" reports the gender gap in the respective counterfactual whereas the column "% explained" indicates how much of the baseline gap can be explained through that particular channel.

Decomposing the Gap in Wealth Levels. The upper panel of Table 5 shows that differences in the deterministic part of income, stochastic part of income, and household sizes explain the largest fraction of the wealth gap between single men and single women.

Overall, heterogeneity in deterministic income explains around 50% of the "gender wealth gap". However, this result masks substantial heterogeneity among the individual components of deterministic income. Only replacing the constant $\bar{y}$ of single women with its male counterpart reverses the gender gap in asset holdings. In contrast, adjusting only the age-specific component ξ or – to a lesser extent – the education premium θ further widens the gap. When estimating the deterministic income component on PSID data (see Table 11 in Appendix C.2), I find a negative coefficient of being a single woman, as opposed to a positive interaction term of being a single woman with age and education. Consequently, when replacing the female constant $\bar{y}$ with that of single men, it will be multiplied by the larger female age and education specific interaction term, resulting in the reversed gender wealth gap. In contrast, adjusting the age-specific term ξ not only affects the slope of single women's deterministic income but also reduces their lifetime earnings and asset holdings.

Next, I perform two additional exercises to isolate how the level vs. slope of deterministic income affect the gender wealth gap. First, to analyze the importance of gender heterogeneity in the slope, I again change the age-component ξ of single women to their male counterpart but readjust the female constant $\bar{y}$ such that – on average – single women have the same deterministic income as in the baseline. Hence, I fix the level of female deterministic income but adjust its slope to be equal to that of single men. Second, to analyze the importance of gender heterogeneity in the level of deterministic income, I increase single women's deterministic income at each age by the average difference between men and women. As a result, single women will have on average the same deterministic income as single men while their slope remains equal to the baseline. Table 5 shows that adjusting the slope accounts for 10% of the gap, whereas adjusting the level accounts for 38%. That is, level differences in deterministic income are a more important determinant of gender differences in net worth than heterogeneity in its slope.

In addition, the income process of single women exhibits a smaller overall variance than that of single men (see Table 2). Therefore, assigning single women the stochastic part of the male income process increases female precautionary savings. This channel in isolation explains 54% of the overall wealth gap.

Gender differences in household sizes further explain 44%. By the specification of the utility function (Equation 1), larger household sizes of single women (mainly through the presence of children) reduce their effective discount factor, increasing consumption relative to savings. This modeling feature reflects in a parsimonious way that more household members increase consumption needs and lower single women's ability to save.

Increasing single women's marriage probability to that of single men, assuming that women marry the same partners (in terms of income realization, education, and wealth), or assigning single women the male medical expenses, survival probability, and male fraction of highly educated individuals is quantitatively less important for explaining gender heterogeneity in asset holdings. Simulating the model under the assumption that both single men and single women start from the same (male) wealth level reduces the wealth gap early in life but has little effect on asset holdings during old age.

Decomposing the Gap in Equity Shares. Similar to asset holdings, differences in deterministic income and household sizes explain the largest fraction of the gender gap in equity shares (see lower panel of Table 5). In contrast, heterogeneity in stochastic income contributes to widening the gap.

Assigning single women the male deterministic income explains around 51% of the gender investment gap. Further decomposing this mechanism into the contribution of slope vs. level (as explained above) reveals that gender heterogeneity in the slope of deterministic income accounts for 12% of the gap, whereas level differences account for 40%.

On average, single men's deterministic income is higher than that of single women. Hence, when adjusting the female level to its male counterpart, single women's lifetime income increases, making them more likely to invest in risky assets.[18] However, as visible from Figure 2b, the life-cycle path of single men's deterministic income is more 'front-loaded', in that single women start from a lower level but catch up to men as they age. Consequently, when equating the slope of female deterministic income to that of single men, the equity share of single women increases especially early in life, as they now start from a higher level. Participating in the stock market during young ages has the advantage to earn higher returns in expectation, thereby accumulating more wealth, and, as a result, becoming more likely to also participate when old. Thus, despite single women having the same lifetime income as in the baseline, equating their slope to that of single men increases the average female equity share. Hence, it is not only gender heterogeneity in the level of deterministic income *per se*

[18] Higher deterministic income may affect the equity share by shifting the distribution of households or by altering policy functions conditional on state variables. See Section 6.2 for a more detailed discussion on the relative importance of both channels.

Table 5: Decomposition Results

Gap in Asset Holdings in (000s) 2007 $	Model	% explained	Data
Baseline	48.31		91.00[a]
Male deterministic income	24.64	48.99%	
only: constant ($\bar{y}$)	-106.71	320.89%	
age component (ξ)	95.47	-97.62%	
education premium (θ)	57.95	-19.96%	
Male deterministic income: slope effect	43.31	10.35%	
Male deterministic income: level effect	30.05	37.81%	
Male stochastic income	22.44	53.55%	
Male household size	27.14	43.83%	
Male marriage probability	43.51	9.93%	
Male marriage market	45.02	6.82%	
Male education distribution	44.55	7.78%	
Male medical expenses	48.54	-0.47%	
Male survival probability	49.04	-1.51%	
Male initial wealth	33.11	31.47%	
Gap in Equity Share in % - points	**Model**	**% explained**	**Data**
Baseline	4.31		4.24
Male deterministic income	2.09	51.39%	
only: constant ($\bar{y}$)	-6.34	247.12%	
age component (ξ)	8.45	-96.32%	
education premium (θ)	4.97	-15.50%	
Male deterministic income: slope effect	3.81	11.64%	
Male deterministic income: level effect	2.57	40.30%	
Male stochastic income	5.72	-32.87%	
Male household size	-1.17	127.21%	
Male marriage probability	3.52	18.30%	
Male marriage market	3.02	29.80%	
Male education distribution	4.15	3.68%	
Male medical expenses	4.25	1.21%	
Male survival probability	4.50	-4.38%	
Male initial wealth	1.88	56.33%	

Notes: Table 5 shows the results of the decomposition exercise. The column "Model" reports the average gender gap in the respective counterfactual. The column "% explained" indicates how much of the baseline gap can be explained through that channel. All values refer to averages over the life-cycle.

[a] The lower wealth gap in the model compared to the data mainly arises from the model's inability to explain the fast asset accumulation of single men beyond age 50 (see Figure 4). When replicating the decomposition exercise on the sample of households below 50 (i.e., for which the model matches the data better), the relative importance of the respective channels hardly changes.

that matters for the gender investment gap, but also its distribution over the life-cycle.

Changing the constant $\bar{y}$ of single women to its male value scales the entire female deterministic income path upward. That is, single women start from a higher level and their deterministic income increases faster than that of single men (due to the positive interaction term of being a single woman with age, see Table 11). In response, the gender investment gap reverses quite substantially. In contrast, adjusting the age component ξ widens the gender gap in equity shares. Single women start from a lower level and their deterministic income grows slower, reducing female lifetime income.

Eliminating heterogeneity in household sizes reverses the gender investment gap. Reducing female household sizes to that of single men increases single women's effective discount factor. As a result, they accumulate more wealth and are more likely to invest in the risky asset. In addition, they choose a higher equity share conditional on state variables, because they understand that their household sizes will also be lower in future periods. I explain this mechanism more carefully in Section 6.2.

When single women face the same stochastic income as single men, the gender gap in equity shares increases by 33%. Since the income process of single men has a higher overall variance than that of single women (see Table 2), assigning single women the male stochastic income component lowers their willingness to take financial risk, despite them accumulating more wealth on average. Consequently, adjusting *overall* income (i.e., both the deterministic and stochastic component) of single women to that of single men has a smaller impact (in absolute terms) on the gender investment gap than changing each component individually, as the two elements partly offset each other. However, note that I control for *overall* income in empirical regressions (Table 1), alluding to why controlling for income does not close (or, at least, does not narrow more strongly) the gender investment gap in the data, despite heterogeneity in income being an important driver of differences in equity shares between single men and single women.

Altering the marriage market explains 30% of the gender investment gap which is partly driven by higher female asset holdings because single women expect their partner to be less wealthy, thereby decreasing the financial returns to marriage. In addition, the distribution of prospective spouses in terms of their stochastic income realization becomes less dispersed which encourage risk-taking. Finally, when single women start from the same wealth level as single men, the gender investment gap substantially declines among young households. Hence, already at age 30, the wealth gap between single women and single men is sufficiently large to prevent single women from participating in the stock market.

6.2 Composition vs. Policy Effect

In the model, the equity share for individual s, α_s, is determined by the policy function $\phi(X_s)$ which maps the individual's state variables X_s into the optimal equity share. In turn, aggregate portfolio allocations are determined by individual policy functions and the distribution of individuals across the state space: $\frac{1}{S}\sum_{s=1}^{S}\alpha_s = \frac{1}{S}\sum_{s=1}^{S}\phi(X_s)$.

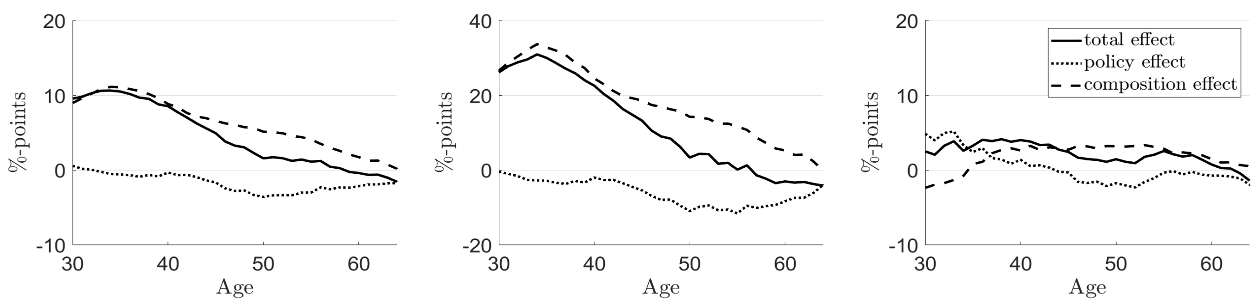
Thus, gender differences in aggregate investment patterns can arise either because the distribution of individuals across the state space differs ("composition effect") or because of gender heterogeneity in policy functions at any given point in the state space ("policy effect"). The objective of this section is to quantify the relative importance of each effect on the gender investment gap along the life-cycle.

To do so, I decompose average investment differences between single men (m) and single women (f) in the model at every age j according to:

$$\frac{1}{F}\sum_{f=1}^{F}\phi(X_f;f) - \frac{1}{M}\sum_{m=1}^{M}\phi(X_m;m) \approx \underbrace{\left[\frac{1}{F}\sum_{f=1}^{F}\phi(X_f;f) - \frac{1}{F}\sum_{f=1}^{F}\phi(X_f;m)\right]}_{\text{Policy Effect}} + \underbrace{\left[\frac{1}{F}\sum_{f=1}^{F}\phi(X_f;m) - \frac{1}{M}\sum_{m=1}^{M}\phi(X_m;m)\right]}_{\text{Composition Effect}} \tag{19}$$

The first difference on the right-hand side is the policy effect (i.e., fixing the vector of state variables and letting the policy functions for the equity share differ) and the second difference is the composition effect (i.e., fixing the policy functions and letting the vector of state variables differ).[19] Figure 5a plots how many percentage points of the gender gap in equity shares can be explained by either effect. Figures 5b and 5c repeat the same exercise separately for the extensive margin (participation rates) and intensive margin (conditional risky shares).

Figure 5: Composition vs. Policy Effect – Aggregate



(a) Equity Share (b) Participation Rate (c) Conditional Share

Notes: Figure 5 decomposes aggregate gender gaps in (unconditional) equity shares, participation rates, and conditional risky shares into a composition and policy effect. The composition effect (dashed line) shows what part of the overall gap can be explained through gender differences in the sample composition whereas the policy effect (dotted line) shows what part of the gap can be explained by differences in policy functions, conditional on state variables.

The composition effect explains the majority of gender differences in unconditional equity shares early in the life-cycle. That is, most of the gender investment gap among young households arises from differences in the sample composition. Beyond age 40, the policy effect even has a negative impact on gender differences in equity shares, meaning that single women take *more* financial risk than single men. When further separating the extensive from the intensive margin, I find that compositional differences primarily affect participation rates. In contrast, the gender gap in conditional risky shares among young households is mostly explained by heterogeneity in policy functions (Figure 5c).

However, as discussed in Section 6.1, individual model elements affect the gender investment gap in opposite directions. Thus, Figure 5 may mask substantial heterogeneity about the

[19] I thank an anonymous referee for suggesting this illustrative way of separating the composition from the policy effect.

relative importance of these individual channels. To analyze such heterogeneity, Figure 6 further decomposes the policy effect into its contribution arising from deterministic income, household sizes, and stochastic income.[20] The solid lines denote the total gender investment gap, the dashed lines the composition effect, the dash-dotted lines indicate how much of the policy effect can be accounted for by the respective channel, and the dotted lines illustrate the importance of all remaining factors (as listed in Table 5) for the policy effect.[21]

In Figure 6a, the dash-dotted line (i.e., the part of the policy effect can be explained by deterministic income) is slightly above the dotted line (i.e., the part of the policy effect can be explained by all other channels), meaning that lower female deterministic income contributes to single women taking less financial risk (conditional on state variables). When separately considering the extensive and intensive margin (see Figures 16 and 17 in Appendix D.3), I find that lower deterministic income affects female policy functions mainly with regard to the participation decision and, to a larger extent, the conditional risky share.

How does deterministic income affect individuals' policy functions? When deciding on the optimal equity share, forward-looking agents take into account both their contemporaneous state variables (e.g. their current level of wealth and income) as well as all future variables. That is, conditional on current-period variables, policy functions differ between single men and single women because of heterogeneity in future outcomes. As long as the risky asset return is only mildly correlated to income shocks (that is, as long as labor income is rather bond than stock-like), a higher human capital endowment (i.e., higher expected discounted earnings) increases households' willingness to take financial risk. Hence, conditional on state variables, it is optimal for single women to choose a lower equity share because their de-

[20] To do that, I decompose the policy effect according to: $\frac{1}{F}\sum_{f=1}^{F}\phi(X_f;f) - \frac{1}{F}\sum_{f=1}^{F}\phi(X_f;m) \approx \left[\frac{1}{F}\sum_{f=1}^{F}\phi(X_f;f) - \frac{1}{F}\sum_{f=1}^{F}\phi(X_f;cf)\right] + \left[\frac{1}{F}\sum_{f=1}^{F}\phi(X_f;cf) - \frac{1}{F}\sum_{f=1}^{F}\phi(X_f;m)\right]$, where cf denotes the respective counterfactual model simulation. Accordingly, the first term on the right hand side is the part of the policy effect that can be explained by the respective counterfactual(s) and the second term is the part that can be explained by all remaining factors.

[21] Figures 16 and 17 in Appendix D.3 document the corresponding results for participation rates and conditional risky shares.

terministic income is in expectation lower than that of single men (see Figure 2), especially early in the life-cycle.

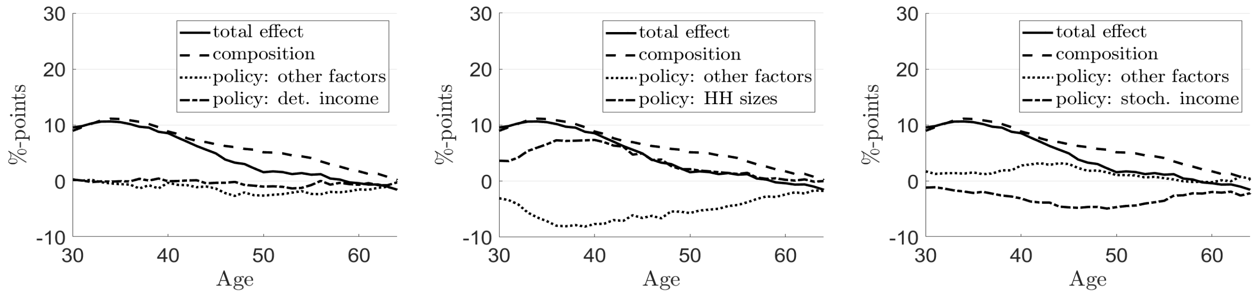
Moreover, gender heterogeneity in household sizes affects the gender investment gap through its impact on policy functions (Figure 6b). As shown in Figure 14, female household sizes are larger than those of single men during their entire working life. In the model, larger household sizes enter through demographic shifters in the utility function (Equation 1), lowering single women's effective discount factor. A smaller discount factor decreases the present value of future labor income and as a result, the optimal female equity share. That is, forward-looking single women understand that their optimal consumption will be higher than that of single men (conditional on state variables) also in future periods, reducing their willingness to take financial risk.[22]

Finally, as shown in Figure 6c, gender heterogeneity in the stochastic part income lowers the total policy effect. The income process of single women is characterized by a smaller variance than that of single men (Table 2), making single women more willing to take financial risk (conditional on state variables). The increased willingness to invest in the stock market affects both their decision to participate, as well as their optimal risky share conditional on participating (see Figures 16 and 17).

Thus, on aggregate it seems that heterogeneity in the sample composition (that is, in contemporaneous characteristics) is most important to explain the gender investment gap. However, further analysis reveals that gender heterogeneity in the main determinants of the gap – deterministic income, household sizes, and stochastic income – all contribute to policy functions for the risky share to differ across single men and single women, albeit in different directions. Conditional on their position in the state space, policy functions across households differ because of heterogeneity in future outcomes, something one cannot as easily control for in

[22] Note that part of the policy effect arises from heterogeneity in household sizes during the current period, as a man and woman with the same state variables differ in both contemporaneous and future household sizes. However, the contribution of current-period household sizes to the policy effect is small when compared to the impact of future outcomes.

Figure 6: Further Decomposition of Policy Effect



(a) Deterministic Income (b) Household Sizes (c) Stochastic Income

Notes: Figure 6 decomposes the aggregate gender gaps in (unconditional) equity shares, participation rates, and conditional risky shares into a composition and policy effect along the life-cycle. The composition effect (dashed line) shows what part of the overall gap can be explained through gender differences in the sample composition whereas the policy effect (dotted line) shows what part of the gap can be explained by differences in policy functions, conditional on state variables.

reduced form regressions.

6.3 Additional Evidence

The decomposition exercise in Section 6.1 revealed that heterogeneity in deterministic income and household sizes are important determinants of the gender investment gap. Moreover, as shown in Section 6.2, contemporaneous and future outcomes of both variables matter for current period investment choices.

The purpose of this section is to provide further evidence of these channels in the data. First, I show that controlling for proxies of deterministic income (as opposed to overall income) helps to reduce the empirical gender investment gap. Second, I provide direct survey evidence on singles' perceptions about future household sizes and (overall) income.

Controlling for Deterministic Income Proxies. According to the structural model, controlling for observable characteristics that proxy for 'deterministic income' should narrow empirical gender differences in investment choices. In fact, including occupation and industry fixed effects (that arguably contain information about the deterministic part of income) in the regressions from Table 1 shrinks the baseline gender effect from 17%-points to 12%-points (Columns (3) and (4) in Table 1). In addition, the estimated age gradient of being a single woman becomes smaller. Thus, including occupation and industry fixed effects helps to reduce the unexplained part of the investment gap in particular among young households for whom differences in deterministic income are largest (see Figure 2).

Direct Survey Evidence. Neither the PSID nor SCF contains information on individuals' perceptions about future income or household sizes. Therefore, I complement the analysis with data from the New York Fed Survey of Consumer Expectations (SCE).[23] The SCE asks respondents about their expected annual earnings in four months and the expected number of individuals living in the same household one year from the time of the interview. However, annual earnings refer to overall income, that is both deterministic and stochastic income in the model. Thus, one should not regard this exercise as a direct mapping from model to data but rather as suggestive evidence whether single women predict their income to be lower than that of single men in future periods. Conditional on the same current overall income, single women should be on average more likely to have a positive transitory income shock (because their deterministic income is lower) and thus, more likely to predict their earnings to be below those of single men in the next period.

I regress expected annual earnings of single households on a gender dummy while controlling for current earnings, education, age, year- and region-fixed effects (Column (1) in Table 6). In Column (2), I additionally control for the inverse hyperbolic sine of financial wealth. In line with the proposed model mechanism, single women expect their future earnings to be $9 - 16\%$ lower than that of single men, depending on the specific set of included control variables. When splitting the sample by age (Columns (3) and (4)), I find that gender differences in income expectations are larger among younger singles, when gender differences in deterministic income are largest (Figure 2).[24]

Furthermore, Figure 7a plots the distribution of expected household sizes for all single house-

[23] See Appendix A.3 for details about this dataset and variable construction.

[24] The negative coefficient for being a single woman does not mean that women necessarily expect their income to decrease which could be at odds with their steep income slope early in life (Figure 2). The negative coefficient rather states that they expect to earn less *relative to single men*.

Table 6: Expected Earnings of Single Households – SCE Data

	(1) log(exp. earnings)	(2) log(exp. earnings)	(3) log(exp. earnings) Age $\leq$ 45	(4) log(exp. earnings) Age $>$ 45
single woman	-0.1619$^{\diamond}$	-0.0922$^{\diamond}$	-0.2394$^{\diamond}$	-0.0915$^{\diamond}$
	(0.0376)	(0.0464)	(0.0548)	(0.0445)
log(current earnings)	0.6556$^{\diamond}$	0.5792$^{\diamond}$	0.5712$^{\diamond}$	0.7927$^{\diamond}$
	(0.0502)	(0.0670)	(0.0648)	(0.0550)
high education	0.2617$^{\diamond}$	0.2132$^{\diamond}$	0.3591$^{\diamond}$	0.1429$^{\diamond}$
	(0.0365)	(0.0481)	(0.0524)	(0.0382)
$\mathbb{1}$. age $>$ 40	-0.0207	-0.0186		
	(0.0421)	(0.0460)		
financial wealth		0.0246$^{\diamond}$		
		(0.0095)		
constant	3.5953$^{\diamond}$	4.2170$^{\diamond}$	4.4481$^{\diamond}$	2.1708$^{\diamond}$
	(0.5231)	(0.6941)	(0.6874)	(0.5859)
Observations	3,009	1,697	1,774	1,235
Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes

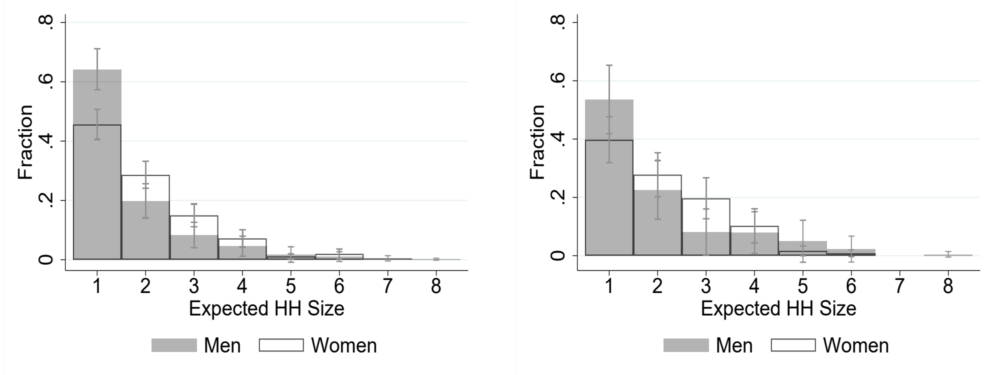
Notes: Estimations are based on linear regressions on the sample of individuals who live in households with no spouse present. Data is from the Labor Market Module of the Survey of Consumer Expectations (SCE), waves 2014 until 2019. *single woman* is a dummy indicating that the household head is a woman. *high education* is a dummy equal to one if the household head has more than 12 years of schooling. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

holds whereas Figure 7b restricts the sample to households below age 40. In both cases, single women assign a lower probability to living alone than single men. On the other hand, they are more likely to expect to live with one or more additional household members who are not prospective spouses. These differences largely reflect heterogeneity in the current number of household members. Hence, Figure 7 rather ensures that singles (correctly) predict household sizes to be persistent. Again, this finding aligns with the model where single women have, conditional on age, both larger current and future household sizes.

7 Robustness Checks & Discussion of Assumptions

In this section, I perform several robustness checks and discuss certain assumptions of the structural framework. First, I revisit the presence of child support and spousal maintenance

Figure 7: Expected Household Sizes of Singles – SCE Data



(a) All Households (b) Households below age 40

Notes: Figure 7 plots the histogram of expected household sizes (= expected number of household members) for single households. Figure 7a plots the distribution for the entire sample, whereas Figure 7b restricts the sample to households below age 40. Data is from the Household Finance Module of the Survey of Consumer Expectations (SCE), waves 2014 until 2019.

payments. Next, I test the sensitivity of the model results with regard to potential gender heterogeneity in bequests. Finally, I show that my results are robust to variations in exogenously set parameters.

Alimony Payments and Child Support. In the US, official regulations determine the amount of alimony payments and child support following a divorce.[25] Typically, the non-custodial parent (or, in case of joint custody, the partner living primarily with the child) receives regular payments from either their ex-partner or a government agency to economically support the child. Spousal support, in contrast, is more restricted and depends on factors such as the length of the marriage, relative income of spouses, and their future financial prospects. In both cases, these transfers may alter the income path of singles in the form of a redistribution from single men to single women because women are more likely to be granted custody and tend to earn less than their (former) husbands.

Yet, I abstract from introducing these payments explicitly for two reasons. First, it has been shown that compliance with such laws tends to be low (Del Boca and Flinn 1995; Case et al. 2003). Second, introducing alimony payments and child support requires an additional state

[25] Spousal and child support in the U.S. are governed by state laws. A comprehensive overview of individual regulations can be found here.

variable that keeps track of the individual's marital history. I, however, account for alimony payments and child support empirically by including them in the income measure. Moreover, I test the sensitivity of the model results with regard to the asset allocation upon divorce and solve a counterfactual version in which the wife receives 65% and the husband 35% of the couple's assets following a divorce (instead of the 50-50 splitting rule assumed in the benchmark). In a parsimonious way, one could think of all alimony and child support claims being paid in a lump-sum transfer directly after divorce instead of being spread out across multiple years.

In response, single men hold on aggregate fewer assets, whereas single women are slightly richer than in the baseline framework. Panel I of Table 7 reports that the resulting gender gap in asset holdings shrinks from on average $48,000 to $28,500. Consequently, the gender investment gap shrinks from 4.31%-points to 2.65%-points. However, when performing the decomposition analysis of Section 6.1 on the model with this modified asset splitting rule, the relative importance of the most important factors contributing to the gender investment gap changes little when compared to the baseline. That is, the main results of the paper are robust to this modification.

Bequests. It may be that both bequests given and bequests received differ by gender, which in turn alters expectations about income and wealth outcomes. To test for this possibility in the data, I exploit a module in the SCF that collects information on whether its respondents have ever received an inheritance or substantial gift and if they expect one in the future. Panel I of Table 8 lists the distribution of these (expected) inheritances by family type. Single men and single women are equally likely to have ever received an inheritance. However, conditional on having inherited something, men receive on average more.[26] Moreover, single women are less likely to expect an inheritance in the future.

[26] For both measures, I restrict the sample to households above 55 years because Bauluz and Meyer (2024) document that most households in the U.S. inherit wealth when they are between 50 and 60 years old. The results remain qualitatively unchanged when considering the entire sample.

Table 7: Robustness Checks

Gender Gap in ...	Asset Holdings (in 000s $)	Equity Shares (in %-points)
Baseline	48.31	4.31
Panel I: Alimony & Child Support		
Modified splitting rule after divorce	28.55	2.65
Panel II: Bequests		
Variation in bequest parameters	48.42	4.29
Lump-sum transfer at age 55	48.41	4.17
Panel III: Exogenous Parameters		
Higher correlation of income shocks within couples	48.50	3.93
No asset drop after death of spouse	48.10	4.22
Reduced medical expenses for couples	47.90	4.18
Higher pension payments for couples	47.79	4.05

Notes: Table 7 reports the average gender gap in asset holdings and equity shares in the baseline model, as well as in alternative versions: Panel I performs robustness with regard to alimony payments and child support, Panel II with regard to bequests, and Panel III with regard to exogenously set model parameters. "Gender Gap" describes the difference of the respective variable between single men and single women, averaged over working age. The gender gap in asset holdings is expressed in (000s) 2007$. The gender gap in equity shares is expressed in percentage points.

Building on this evidence, I perform a robustness check where I introduce a lump-sum transfer to all households at age 55 in the model. Single men receive $9,143, single women $7,988, and couples $16,977. These values reflect the empirical amount of received bequests times the probability of having received something (as shown in the first two rows of Panel I in Table 8). In response to introducing these transfers, households accumulate less wealth in the years leading up to age 55 and hold on average more wealth afterward. However, given that the received amount is quite similar for single men and single women, the gender gap in asset holdings and equity shares hardly changes when compared to the benchmark (see Panel II of Table 7).

Furthermore, the SCF contains information on whether households perceive leaving bequests as important and if they expect to leave a "sizable estate" to others. Panel II of Table 8 shows that around half of all households consider leaving an inheritance as important, regardless of their family type. Hence, the assumption of homogeneous bequest parameters across all

Table 8: Distribution of (expected) Bequests by Family Type (Data)

	Couples	Singles	
		Men	Women
Panel I: Bequests received			
% received inheritance (> 55 yrs)	30.23 (0.36)	25.61 (0.73)	26.25 (0.55)
$ amount received in (000s) (> 55 yrs)	56.16 (1.52)	35.70 (2.61)	30.43 (1.78)
% expect inheritance	17.05 (0.14)	14.87 (0.28)	10.24 (0.19)
Panel II: Bequests given			
% perceive bequests as important	51.81 (0.18)	52.25 (0.41)	53.39 (0.31)
% expect to give bequest	56.06 (0.18)	51.61 (0.40)	41.75 (0.31)

Notes: Table 8 reports the distribution of (expected) bequests across family types. "% received inheritance (> 55 yrs)" indicates the fraction of HHs above age 55 who have received an inheritance or substantial gift, whereas "$ amount received in (000s) (> 55 yrs)" indicates the average amount. "% expect inheritance" denotes the fraction of households who expect to receive an inheritance in the future. "% perceive bequests as important" is the fraction who considers leaving something behind as important, and "% expect to give bequest" is the fraction who expects to do so. Standard errors are in parentheses. Data is from the Survey of Consumer Finances (SCF), waves 1989-2016.

household types is supported empirically. In addition, as I model bequests to be a luxury good, the share of bequeathed wealth is increasing in households' asset holdings, and will therefore be on average highest for couples, followed by single men, and then single women. Again, I confirm this pattern in the data: whereas 42% of single women expect to leave something behind, 52% of single men and 56% of couples do. Finally, Panel II of Table 7 shows that my results are robust to modifying the exact parameter values of the bequest motive from 0.128 to 12.8 (L) and from 0.73 to 73 (ω).

Variation in exogenously set parameters. Panel III of Table 7 compares the gender gap in equity shares and asset holdings of the baseline model to alternative versions in which I test the robustness with regard to exogenously set parameters. I change – one-by-one – the correlation of transitory income shocks within couples from 0.3 to 0.9, I assume that assets remain constant whenever one spouse dies, that couples only pay 80% of medical expenses to account for informal care arrangements across partners, and that pension payments of couples are twice as large as that of single men (instead of 1.7 times). In all cases, the gender investment gap changes little when compared to the baseline framework.

8 Conclusion

This paper studies the gender investment gap through the lens of a structural life-cycle framework. First, I provide empirical evidence that single women are less likely to participate in the stock market than single men and that they allocate a smaller share of their net worth toward risky assets. This gap is largest among young households and remains statistically significant in reduced form regressions after controlling for a wide range of observable characteristics that have been shown to affect investment behavior.

A life-cycle model of portfolio choice that restricts preferences to be equal across men and women is able to replicate the empirical gap. Counterfactual simulations reveal that higher male deterministic income and fewer household members are the main determinants for explaining the gender investment gap. In contrast, gender heterogeneity in the stochastic part of income contributes to widening the gap.

Importantly, both contemporaneous and future outcomes of deterministic income and household sizes affect current-period investment choices. Because labor income is only mildly correlated to stock returns, a higher human capital endowment (i.e., higher future deterministic income) increases single men's optimal equity share. Similarly, lower future household sizes raise the present value of their future labor income, increasing financial risk-taking. Hence, in line with the empirical evidence, reduced form regressions on model simulated data that do not take into account gender heterogeneity in future realizations fail to fully explain the gender investment gap, in particular during young age.

References

Agnew, J., P. Balduzzi, and A. Sunden (2003). Portfolio choice and trading in a large 401 (k) plan. *American Economic Review 93*(1), 193–215.

Alan, S. (2006). Entry costs and stock market participation over the life cycle. *Review of Economic Dynamics 9*(4), 588–611.

Almenberg, J. and A. Dreber (2015). Gender, stock market participation and financial literacy. *Economics Letters 137*, 140–142.

Ameriks, J. and S. P. Zeldes (2004). How do household portfolio shares vary with age. *Working Paper, The Vanguard Group*.

Andersen, S., J. Marx, K. M. Nielsen, and L. Vesterlund (2021). Gender differences in negotiation: Evidence from real estate transactions. *The Economic Journal 131*(638), 2304–2332.

Arano, K., C. Parker, and R. Terry (2010). Gender-based risk aversion and retirement asset allocation. *Economic Inquiry 48*(1), 147–155.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Bauluz, L. and T. Meyer (2024). The wealth of generations. *Working Paper, CUNEF Universidad*.

Bertocchi, G., M. Brunetti, and C. Torricelli (2011). Marriage and other risky assets: A portfolio approach. *Journal of Banking & Finance 35*(11), 2902–2915.

Bertrand, M. (2011). New perspectives on gender. In *Handbook of Labor Economics*, Volume 4, pp. 1543–1590. Elsevier.

Borella, M., M. De Nardi, and F. Yang (2020). The lost ones: The opportunities and outcomes of white, non-college-educated americans born in the 1960s. *NBER Macroeconomics Annual 34*(1), 67–115.

Borella, M., M. De Nardi, and F. Yang (2023). Are marriage-related taxes and social security benefits holding back female labour supply? *The Review of Economic Studies 90*(1), 102–131.

Cagetti, M. (2003). Wealth accumulation over the life cycle and precautionary savings. *Journal of Business & Economic Statistics 21*(3), 339–353.

Case, A. C., I.-F. Lin, and S. S. McLanahan (2003). Explaining trends in child support: Economic, demographic, and policy effects. *Demography 40*(1), 171–189.

Catherine, S. (2022). Countercyclical labor income risk and portfolio choices over the life cycle. *The Review of Financial Studies 35*(9), 4016–4054.

Charness, G. and U. Gneezy (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization 83*(1), 50–58.

Chetty, R., L. Sandor, and A. Szeidl (2017). The effect of housing on portfolio choice. *The Journal of Finance 72*(3), 1171–1212.

Christelis, D., D. Georgarakos, and M. Haliassos (2013). Differences in portfolios across countries: Economic environment versus household characteristics. *The Review of Economics and Statistics 95*(1), 220–236.

Christiansen, C., J. S. Joensen, and J. Rangvid (2015). Understanding the effects of marriage and divorce on financial investments: The role of background risk sharing. *Economic Inquiry 53*(1), 431–447.

Cocco, J. F. (2005). Portfolio choice in the presence of housing. *The Review of Financial Studies 18*(2), 535–567.

Cooper, R. and G. Zhu (2016). Household finance over the life-cycle: What does education contribute? *Review of Economic Dynamics 20*, 63–89.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–74.

Cubeddu, L. and J.-V. Ríos-Rull (2003). Families as shocks. *Journal of the European Economic Association 1*(2-3), 671–682.

De Nardi, M. (2004). Wealth inequality and intergenerational links. *The Review of Economic Studies 71*(3), 743–768.

Del Boca, D. and C. J. Flinn (1995). Rationalizing child-support decisions. *The American Economic Review*, 1241–1262.

DeLeire, T. and H. Levy (2004). Worker sorting and the risk of death on the job. *Journal of Labor Economics 22*(4), 925–953.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Eckel, C. C. and P. J. Grossman (2008). Men, women and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results 1*, 1061–1073.

Fagereng, A., C. Gottlieb, and L. Guiso (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance 72*(2), 705–750.

Flavin, M. and T. Yamashita (2011). Owner-occupied housing: Life-cycle implications for the household portfolio. *American Economic Review 101*(3), 609–14.

Girshina, A., L. Bach, P. Sodini, and M. Team (2022). Soft negotiators or modest builders? why women earn lower real estate returns. *Working Paper, Stockholm School of Economics*.

Goldsmith-Pinkham, P. and K. Shue (2023). The gender gap in housing returns. *The Journal of Finance 78*(2), 1097–1145.

Gomes, F. (2020). Portfolio choice over the life cycle: A survey. *Annual Review of Financial Economics 12*, 277–304.

Gomes, F. and A. Michaelides (2005). Optimal life-cycle asset allocation: Understanding the empirical evidence. *The Journal of Finance 60*(2), 869–904.

Gourinchas, P.-O. and J. A. Parker (2002). Consumption over the life cycle. *Econometrica 70*(1), 47–89.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic Dynamics 12*(1), 58–79.

Guvenen, F. and S. Ozkan (2021). Tiktak: A multistart global optimization algorithm 1.0. Orcid: https://orcid.org/0000-0001-6364-9606, accessed on 2024-01-23.

Hubener, A., R. Maurer, and O. S. Mitchell (2015). How family status and social security claiming options shape optimal life cycle portfolios. *The Review of Financial Studies 29*(4), 937–978.

Huggett, M. and G. Kaplan (2016). How large is the stock component of human capital? *Review of Economic Dynamics 22*, 21–51.

Jianakoplos, N. A. and A. Bernasek (1998). Are women more risk averse? *Economic Inquiry 36*(4), 620–630.

Jones, J. B., M. De Nardi, E. French, R. McGee, and R. Rodgers (2020). Medical spending, bequests, and asset dynamics around the time of death. *NBER WP No 26879*.

Ke, D. (2018). Cross-country differences in household stock market participation: The role of gender norms. In *AEA Papers and Proceedings*, Volume 108, pp. 159–62.

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. *The Review of Financial Studies 23*(1), 385–432.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature 52*(1), 5–44.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 247–257.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory 3*(4), 373–413.

Neelakantan, U. and Y. Chang (2010). Gender differences in wealth at retirement. *American Economic Review: Papers & Proceedings 100*(2), 362–67.

Panel Study of Income Dynamics (2021). Produced and distributed by the survey research center, institute for social research, university of michigan, ann arbor, mi (2021). *Dataset*.

Paz-Pardo, G. (2024). Homeownership and portfolio choice over the generations. *American Economic Journal: Macroeconomics 16*(1), 207–237.

Poterba, J. M. and A. Samwick (2001). Household portfolio allocation over the life cycle. In *Aging Issues in the United States and Japan*, pp. 65–104. University of Chicago Press.

Rouwenhorst, K. G. (1995). Asset pricing implications of equilibrium business cycle models. *Frontiers of Business Cycle Research 1*, 294–330.

Säve-Söderbergh, J. (2012). Self-directed pensions: Gender, risk, and portfolio choices. *The Scandinavian Journal of Economics 114*(3), 705–728.

Shen, J. (2024). Countercyclical risks, consumption, and portfolio choice: Theory and evidence. *Management Science 70*(5), 2862–2881.

Sunden, A. E. and B. J. Surette (1998). Gender differences in the allocation of assets in retirement savings plans. *The American Economic Review 88*(2), 207–211.

Vissing-Jorgensen, A. (2002). Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. *NBER WP No 8884*.

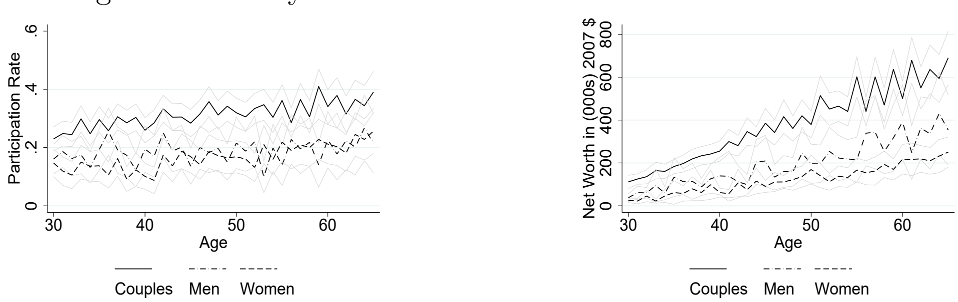
Yao, R. and H. H. Zhang (2005). Optimal consumption and portfolio choices with risky housing and borrowing constraints. *The Review of Financial Studies 18*(1), 197–239.

A Data Appendix

A.1 Comparability of PSID and SCF

Figure 8 plots the life-cycle profiles of stock market participation rates (excluding stocks held through retirement accounts) and net worth by family type in the PSID, as these are measures that are available in both the PSID and SCF. The life-cycle profiles of stock market participation rates without stocks held through retirement accounts in the PSID (Figure 8a) look very similar to those in the SCF (Figure 9b). Most importantly, I can replicate the converging gender investment gap over the life-cycle. When comparing net worth in the PSID (Figure 8b) to the SCF (Figure 1d), I find that wealth levels in the PSID are lower than in the SCF. However, in both datasets, couples and single men have higher net worth than single women, especially as they approach retirement.

Figure 8: Life-Cycle Patterns of Household Finances – PSID



(a) Participation Rate

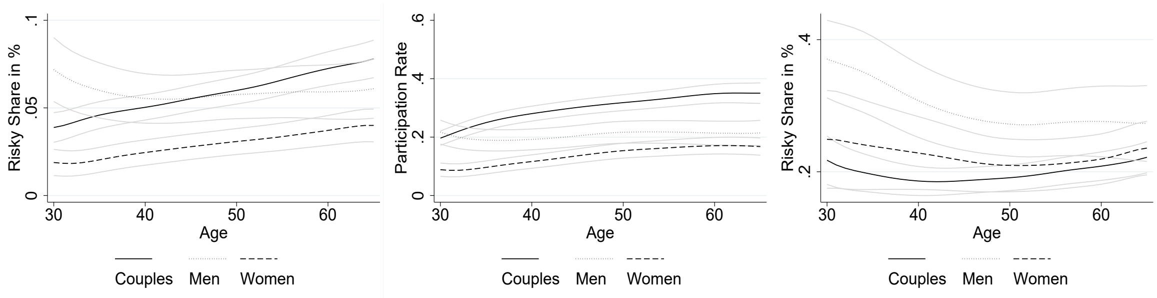
(b) Net Worth

Notes: Figure 8 plots the life-cycle profiles of stock market participation rates and net worth for singles and couples, including 95% confidence intervals. Data is from the waves 1989 until 2017 of the Panel Study of Income Dynamics (PSID). Risky assets are defined as direct stock holdings, excluding stocks held through retirement accounts. Net worth is defined as total assets minus total debt.

A.2 Supplementary Figures

Excluding Retirement Accounts. If single men are more likely to hold retirement accounts than single women, and if individuals, regardless of gender, tend to invest retirement savings riskier than other types of wealth, the gender investment gap could reflect gender heterogeneity in the labor market rather than in investment choices. Figure 9 therefore plots the life-cycle profiles of equity shares, stock market participation rates, and conditional risky shares based on a tighter definition of risky assets that excludes savings held through retirement accounts. The gender gap in equity shares (Figure 9a) remains statistically significant, alleviating concerns that investment differences across gender are mainly driven through savings that are linked to certain types of jobs.

Figure 9: Life-Cycle Patterns of Household Finances – Excluding Retirement Accounts



(a) Equity Share (b) Participation Rate (c) Conditional Risky Share

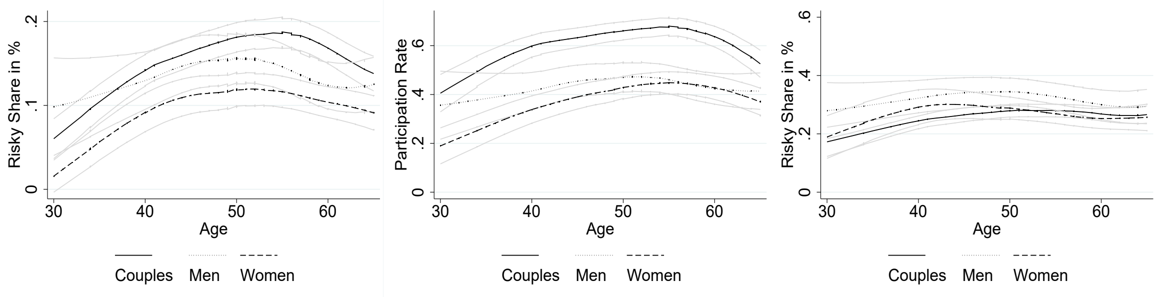
Notes: Figure 9 plots the life-cycle profiles of equity shares, stock market participation rates, and conditional risky shares for singles and couples, including 95% confidence intervals. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF). Risky assets are defined as direct stock holdings, corporate and foreign bonds, and the fraction of mutual funds that include the former. All figures are smoothed to increase readability.

Cohort Effects. Heterogeneity in the gender investment gap across different ages could be driven by cohort-specific investment behavior (Ameriks and Zeldes 2004). Therefore, Figure 10 plots the empirical patterns from Figure 1 when restricting the sample to individuals who were born within a relatively short time frame (1945-1960). All three graphs look qualitatively very similar to the baseline, with larger standard errors due to the reduced sample size.

A.3 New York Fed Survey of Consumer Expectations

In Section 6.3, I work with the Labor Market Survey and Household Finance module from the New York Fed Survey of Consumer Expectations (SCE). The SCE is a nationally representative online survey of around 1,300 private households in the U.S. running from 2013 until today. While respondents are interviewed monthly, questions on topical modules are included less frequently. The Labor Market Survey has been collected in March, July, and November since 2014. The Household Finance module has been included each August from 2014 until 2019. I work with all available waves from these two topical modules up until

Figure 10: Life-Cycle Patterns of Household Finances – One Cohort



(a) Equity Share (b) Participation Rate (c) Conditional Risky Share

Notes: Figure 10 plots the life-cycle profiles of equity shares, stock market participation rates, and conditional risky shares for singles and couples, including 95% confidence intervals. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF) for individuals who are born between 1945-1960. Risky assets are defined as direct stock holdings, corporate and foreign bonds, the fraction of mutual funds that include the former, and the fraction of retirement accounts which is invested in stocks. All figures are smoothed to increase readability.

2019.

I merge both modules to the core data which includes demographic characteristics such as gender and marital status. In the Labor Market module, respondents report their current annual earnings (before taxes) and expected annual earnings in four months, which serves as my dependent variable in Table 6. Annual earnings include bonuses, overtime pay, tips, and commissions. In the household finance module, respondents indicate the number of household members they expect to live with 12 months from the time of the interview (Figure 7). In this figure, I exclude one observation that reports to have more than 20 household members.

In addition, the household finance module contains information about financial wealth, which I use as a control variable in Table 6. Financial wealth is defined as the current value of savings and investments, such as checking and savings accounts, CDs, stocks, bonds, mutual funds, and treasury bonds. It excludes retirement accounts. Note that households are never asked in the same month about both their expected earnings and financial wealth. Therefore, I assign the reported amount of financial wealth to the months in which households report expected earnings and only include households who answer to both modules. All financial variables are expressed in June 2014 Dollars. In line with the main analysis, I further restrict the sample to single households of working age.

A.4 Regression Coefficients and Marginal Effects

Table 9: Regression Coefficients & Marginal Effects – Participation Rates of Singles

	(1) SMP	(2) SMP	(3) SMP	(4) SMP
single woman	-0.2879$^\diamond$	-0.1793$^\diamond$	-0.1363$^\diamond$	-0.1133$^\diamond$
	(0.0103)	(0.0096)	(0.0116)	(0.0115)
single woman * age	0.0049$^\diamond$	0.0031$^\diamond$	0.0026$^\diamond$	0.0019$^\diamond$
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
age	-0.1368$^\diamond$	-0.0931$^\diamond$	-0.0917$^\diamond$	-0.0916$^\diamond$
	(0.0140)	(0.0140)	(0.0147)	(0.0220)
$age^2 * 100$	0.3218$^\diamond$	0.2240$^\diamond$	0.2123$^\diamond$	0.2108$^\diamond$
	(0.0305)	(0.0305)	(0.0324)	(0.0489)
$age^3 * 10000$	-0.2452$^\diamond$	-0.1720$^\diamond$	-0.1611$^\diamond$	-0.1541$^\diamond$
	(0.0216)	(0.0216)	(0.0231)	(0.0350)
high education		0.2764$^\diamond$	0.2585$^\diamond$	0.1609$^\diamond$
		(0.0020)	(0.0018)	(0.0028)
no. of HH members		-0.0419$^\diamond$	-0.0407$^\diamond$	-0.0422$^\diamond$
		(0.0007)	(0.0008)	(0.0016)
non-asset income		0.0292$^\diamond$	0.0261$^\diamond$	0.0252$^\diamond$
		(0.0005)	(0.0004)	(0.0003)
safe assets			0.0147$^\diamond$	0.0117$^\diamond$
			(0.0003)	(0.0003)
constant	2.3350$^\diamond$	1.1982$^\diamond$	1.1751$^\diamond$	1.1868$^\diamond$
	(0.2061)	(0.2070)	(0.2160)	(0.3195)
Observations	10,244	10,240	10,240	7,607
Year FE	No	Yes	Yes	Yes
Industry & Occupation FE	No	No	No	Yes
ME for women at age 30	-0.1417$^\diamond$	-0.0854$^\diamond$	-0.0590$^\diamond$	-0.0563$^\diamond$
	(0.0045)	(0.0041)	(0.0048)	(0.0058)
ME for women at mean age (47)	-0.0571$^\diamond$	-0.0310$^\diamond$	-0.0143$^\diamond$	-0.0260$^\diamond$
	(0.0024)	(0.0021)	(0.0018)	(0.0031)
ME for women at age 65	0.0289$^\diamond$	0.0242$^\diamond$	0.0311$^\diamond$	0.0102$^\diamond$
	(0.0041)	(0.0040)	(0.0040)	(0.0027)

Notes: Estimations are based on OLS on the sample of individuals who live in households with no spouse present. Source: SCF waves 1989 until 2016. SMP = Stock Market Participation. *single woman* is a dummy indicating that the household head is a woman. *high education* indicates that the household head has more than 12 years of education. *safe assets* refers to all net worth that is not invested in risky assets. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

Table 10: Regression Coefficients & Marginal Effects – Conditional Risky Share of Singles

	(1) Cond. Share	(2) Cond. Share	(3) Cond. Share	(4) Cond. Share
single woman	-0.0120$^\diamond$	-0.0019	-0.0195	-0.0113
	(0.0034)	(0.0067)	(0.0170)	(0.0168)
single woman * age	-0.0009$^\diamond$	-0.0009$^\diamond$	-0.0008$^\diamond$	-0.0011$^\diamond$
	(0.0001)	(0.0002)	(0.0004)	(0.0004)
age	0.0272	0.0429$^\diamond$	0.0307$^\diamond$	0.0393$^\diamond$
	(0.0145)	(0.0147)	(0.0118)	(0.0105)
$age^2 * 100$	-0.0626$^\diamond$	-0.0971$^\diamond$	-0.0591$^\diamond$	-0.0775$^\diamond$
	(0.0293)	(0.0294)	(0.0242)	(0.0213)
$age^3 * 10000$	0.0444$^\diamond$	0.0686$^\diamond$	0.0394$^\diamond$	0.0532$^\diamond$
	(0.0193)	(0.0193)	(0.0163)	(0.0142)
high education		0.0288$^\diamond$	0.0547$^\diamond$	0.0337$^\diamond$
		(0.0022)	(0.0032)	(0.0023)
no. of HH members		-0.0237$^\diamond$	-0.0234$^\diamond$	-0.0224$^\diamond$
		(0.0017)	(0.0012)	(0.0016)
non-asset income		0.0039$^\diamond$	0.0038$^\diamond$	0.0063$^\diamond$
		(0.0006)	(0.0006)	(0.0009)
safe assets			-0.0380$^\diamond$	-0.0389$^\diamond$
			(0.0006)	(0.0007)
constant	-0.0178	-0.3769	0.1063	-0.0875
	(0.2336)	(0.2332)	(0.1882)	(0.1670)
Observations	4,285	4,285	4,285	3,774
Year FE	No	Yes	Yes	Yes
Industry & Occupation FE	No	No	No	Yes
ME for women at age 30	-0.0400$^\diamond$	-0.0295$^\diamond$	-0.0440$^\diamond$	-0.0444$^\diamond$
	(0.0016)	(0.0028)	(0.0064)	(0.0068)
ME for women at mean age	-0.0574$^\diamond$	-0.0467$^\diamond$	-0.0593$^\diamond$	-0.0639$^\diamond$
	(0.0031)	(0.0029)	(0.0026)	(0.0031)
ME for women at age 65	-0.0726$^\diamond$	-0.0617$^\diamond$	-0.0727$^\diamond$	-0.0830$^\diamond$
	(0.0049)	(0.0049)	(0.0072)	(0.0070)

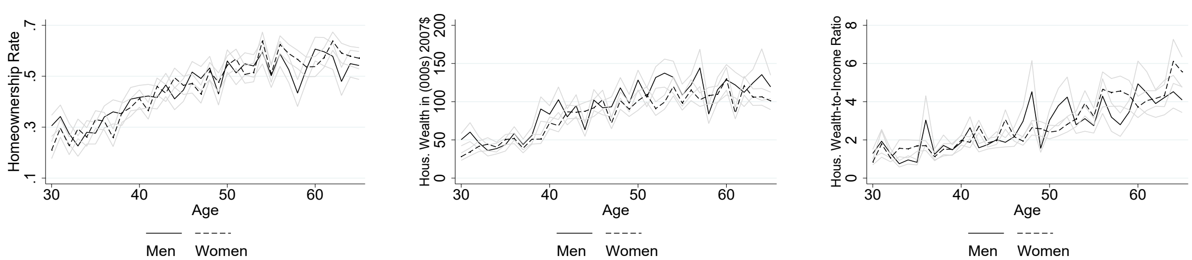
Notes: Estimations are based on OLS on the sample of individuals who live in households with no spouse present. Source: SCF waves 1989 until 2016. Cond. Share = risky share conditional on participation. *single woman* is a dummy indicating that the household head is a woman. *high education* indicates that the household head has more than 12 years of education. *safe assets* refers to all net worth that is not invested in risky assets. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

B Housing

The focus of this paper is on risky financial assets which is why I abstract from modeling housing explicitly. However, housing constitutes a large share of households' portfolios and affects stock market behavior.[27] For the current analysis, abstracting from housing imposes a problem if housing choices directly map into portfolio behavior (and hence, the gender investment gap is in fact a gender housing gap) or if housing differentially affects portfolio choices by gender, i.e., if housing is an important driver of the gender investment gap itself.

To explore whether either of these issues is present in the data, I conduct two exercises: first, if portfolio choices are a direct mapping of housing decisions, the life-cycle profiles of housing variables should closely follow those in Figures 1a to 1c. Figure 11 displays singles' life-cycle profiles of homeownership rates, housing wealth (henceforth: "HW"), and housing wealth-to-income ratio (henceforth: "HI"). Housing refers to the primary residence. For all three variables, I do not find any significant differences between single men and single women despite significant gender gaps in equity shares both along the extensive and intensive margin.

Figure 11: Life-Cycle Profiles of Housing Patterns (Singles)



(a) Homeownership Rate (b) (Gross) Housing Wealth (c) Housing Wealth-to-Income

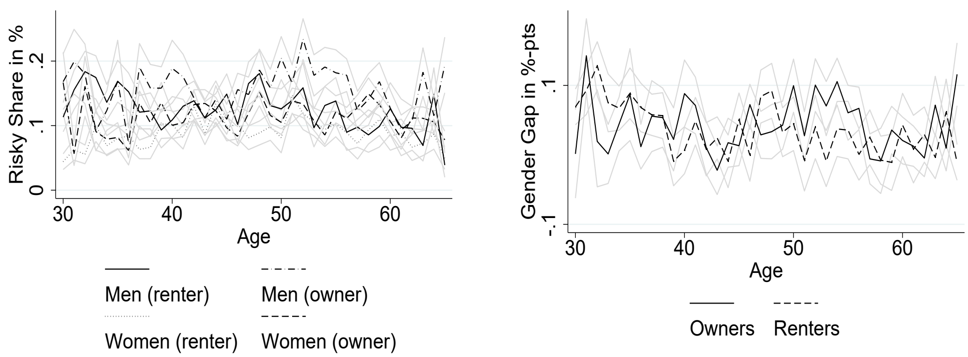
Notes: Figure 11 plots life-cycle profiles of the homeownership rate, gross housing wealth, and housing wealth-to-income ratio for single men and single women, including 95% confidence intervals. Housing refers to the primary residence. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF).

Second, if housing is an important driver of the gender investment gap itself, the gender gap in equity shares should differ by housing tenure. To test for this possibility, Figure 12a

[27] Two of the first papers to introduce housing in a model of portfolio choice were Cocco (2005) and Yao and Zhang (2005). Since then, there has been a large and ongoing literature on housing and portfolio choices, see for example Flavin and Yamashita (2011), Chetty et al. (2017), or Paz-Pardo (2024).

plots the equity share of single homeowners and single renters over their life-cycle, separately by gender. I find that gender *differences* in equity shares are very similar for homeowners and renters. To further illustrate this finding, Figure 12b plots the gender investment gap for renters and owners by age. Both lines are not statistically significant different from one another, reassuring that housing does not differently affect portfolio choices of single men and single women.

Figure 12: Gender Gaps in Equity Shares by Housing Tenure



(a) Equity Shares

(b) Gender Investment Gap

Notes: Figure 12a plots singles' life-cycle profiles of equity shares by gender and housing tenure. Figure 12b plots the gender gap in equity shares for homeowners and renters, respectively. The gender gap in Figure 12b is defined as the average equity share of single men minus the average equity share of single women at the respective age. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF).

C Model Estimation – First Stage

C.1 Income Process Estimation – Stochastic Component

I estimate the stochastic component of the income process by the minimum distance estimator as in Guvenen (2009). I assume that the unexplained part of the income process (that is, the residual term w_{ij} from Equation (3)) to follow a persistent-transitory process which depends on the aggregate state Ω:

$$\tilde{y}_{ij} = z_{ij} + \epsilon_{ij} \quad \text{with} \quad z_{ij} = \begin{cases} \mu_z^{\text{boom}} + \rho_i z_{i,j-1} + \nu_{ij} & \text{if } \Omega = b \text{ with probability } (1 - p_{\text{rec}}) \\ \mu_z^{\text{rec}} + \rho_i z_{i,j-1} + \nu_{ij} & \text{if } \Omega = r \text{ with probability } p_{\text{rec}} \end{cases} \tag{20}$$

Abstracting from the state-dependent mean μ_z, I need to estimate three parameters which are allowed to vary by gender and marital status: the persistence parameter ρ, the variance of the persistent shock σ_ν^2, and the variance of the transitory shock σ_ϵ^2. To do so, I minimize the distance between the covariance-variance matrices of the income process in the data and their theoretical counterparts under the assumption that $Var(z_{-1}) = 0$. Moreover, because the PSID collects data every two years after 1997 while the model is written in annual frequency, I linearly interpolate income for individuals that I observe in two consecutive waves for the missing year in which no PSID data was collected. I run four different estimations for married men, married women, single men, and single women. Table 2 displays the results.

Human Capital Returns. The mean of the persistent part of (stochastic) income μ_z depends on the aggregate state and hence, governs the correlation between stock market returns and income realizations. I set its value during recessions μ_z^{rec} to be small enough such that the correlation between human capital and stock returns is positive but does not exceed values from previous literature.[28] As in Huggett and Kaplan (2016), I define the return to human capital as next period's human capital value plus net earnings (i.e., the dividend of human capital), divided by the current period's value of human capital: $R_{j+1}^h = \frac{v_{j+1}^n + e_{j+1}}{v_j^n}$.

Again following Huggett and Kaplan (2016), the value of human capital v^n is defined as households' expected discounted net earnings, with discounting done using the household's stochastic discount factor. Thus, the value of human capital for a household with gender i, age j, education θ, asset cash-on-hand a, and stochastic income realization $\tilde{y}$ reads as:[29]

$$v(i, j, \theta, a, \tilde{y}) = \mathbb{E}_{k=j+1}^{J} m_{j,k} e_k(i, j, \theta, a, \tilde{y}) \tag{21}$$

where $e_k(i, j, \theta, a, \tilde{y})$ denote expected net earnings in period k. The term $m_{j,k}$ is the house-

[28] The value for μ_z^{boom} is then automatically determined through the stationary condition for the income process: $\mu_z^{\text{boom}} = \left(\frac{-p_{\text{rec}}}{1-p_{\text{rec}}}\right) \mu_z^{\text{rec}}$.

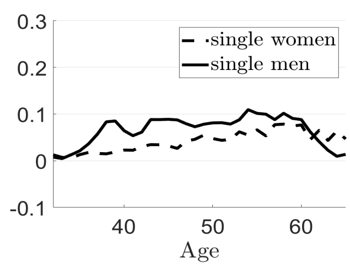
[29] I assume for simplicity that singles take expectations only over the states when they remain single when computing human capital values.

hold's stochastic discount factor. Let X^j summarize household's state space in period j. I then define the stochastic discount factor $m_{j,k}$ as:

$$m_{j,k} = \beta P(X^k|X^j)\frac{\partial U(c^*)/\partial c_k(X^k)}{\partial U(c^*)/\partial c_z(X^j)} \tag{22}$$

The operator $P(X^k|X^j)$ denotes the likelihood to end up in state X^k in period k, given that the household is in state X^j today. Figure 13 plots the correlation between human capital and stock returns for single men and single women over the life-cycle. While the correlation is positive, it is below the estimates in Huggett and Kaplan (2016) who find values of around 0.3 to 0.5. However, there are several key model ingredients that differ between their and my framework. First, Huggett and Kaplan (2016) do not impose any fixed cost of stock market participation. Second, while they consider the sample of (both married and unmarried) males, I restrict the analysis to singles who are generally more exposed to labor income risk (e.g. because they cannot pool individual income streams). Finally, Huggett and Kaplan (2016) introduce pension payments that depend on the aggregate component of earnings, whereas these two factors are uncorrelated in my framework. All those elements contribute to agents being less willing to invest in the risky asset and hence, a smaller correlation between human capital and stock returns is needed to match empirical equity shares.

Figure 13: Correlation: Human Capital and Stock Returns



Notes: Figure 13 plots the correlation between human capital and stock returns by age for single men and single women.

C.2 Income Process – Deterministic Component

Table 11: Regression Coefficients of the Income Estimation (Deterministic Component)

	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage
	Couples		Singles	
high educ.		0.5528$^\diamond$		0.4480$^\diamond$
		(0.0174)		(0.0259)
woman		-3.9134$^\diamond$		-0.7109$^\diamond$
		(0.0328)		(0.0245)
woman*high educ.		-0.3021$^\diamond$		0.1160$^\diamond$
		(0.0421)		(0.0321)
age	0.0798$^\diamond$		0.0592$^\diamond$	
	(0.0165)		(0.0115)	
$age^2 * 100$	-0.0884$^\diamond$		-0.0728$^\diamond$	
	(0.0173)		(0.0113)	
age*woman	0.0523$^\diamond$		0.0110$^\diamond$	
	(0.0046)		(0.0038)	
constant	7.5465$^\diamond$	1.7196$^\diamond$	9.6060$^\diamond$	0.1375$^\diamond$
	(0.3838)	(0.0136)	(0.2728)	(0.0200)
Observations	76,926	76,926	17,482	17,482
Number of unique indiv.	10,705		3,622	

Notes: Estimations are based on (fixed-effect) OLS regressions from PSID Data, waves 1989-2017. Corresponding Figure is Figure 2 in the main text. Dependent variable of first stage: inverse hyperbolic sine transformation of annual non-asset income (labor income, social security income, and transfers). In waves where social security income and transfers are not available separately for head and spouse, I use the combined social security and transfer income and assign it 50-50 to both spouses. For singles, I add labor income, social security benefits, and transfers from other household members. For couples, I split income from other household members 50-50 between spouses. Dependent variable of second stage: fixed effects plus residual from first stage. *high educ.* is a dummy equal to one if the individual has more than 12 years of schooling; *woman* is a dummy indicating if the individual is a woman; Robust standard errors in parentheses, $\diamond$ p<0.05.

C.3 Marriage and Divorce Probabilities

Table 12: Regression Coefficients for Marriage and Divorce Probabilities

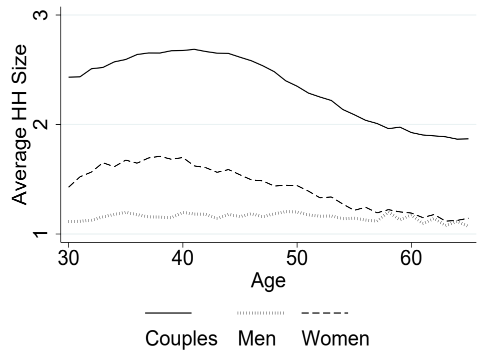
	(1) Marriage Prob.	(2) Divorce Prob.
woman	-0.3203$^{\diamond}$	
	(0.0739)	
age	-0.0588$^{\diamond}$	-0.0442$^{\diamond}$
	(0.0038)	(0.0054)
$\mathbb{1} > 1997$	0.3843$^{\diamond}$	0.5821$^{\diamond}$
	(0.0666)	(0.0958)
high educ. (head)	0.0896	-0.4061$^{\diamond}$
	(0.0727)	(0.1148)
high educ. (spouse)		-0.0826
		(0.1109)
constant	0.1202	-1.5753$^{\diamond}$
	(0.1586)	(0.2502)
Observations	13,331	18,179

Notes: Estimations are based on Logit regressions from PSID Data, waves 1989-2017. Dependent variable: likelihood of getting married (resp. divorced) within the next wave, conditional on not being married (resp. being married) today. The age of a couple refers to the household head. For education within a couple, head refers to the husband and spouse refers to the wife. Singles are always labeled as heads. *high educ.* is a dummy equal to one if the individual has more than 12 years of schooling; *woman* is a dummy indicating if the individual is a woman; $\mathbb{1} > 1997$ indicates observations that were interviewed after 1997 to account for changing frequency of the PSID. Standard errors are clustered on individual level. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

C.4 Equivalence Scales

Figure 14 plots the predicted household sizes by family type over the life-cycle. To obtain these values, I compute the average number of household members by age and family type from the PSID, assigning a weight of 1 to the first adult member, a weight of 0.7 to all further adult members, and a weight of 0.5 to each child.

Figure 14: Life-Cycle Profiles of Average Household Sizes



Notes: Figure 14 plots average household sizes for single men, single women, and couples by age. The first adult member in the household receives a weight of 1, all further adults a weight of 0.7, and each child a weight of 0.5.

D Model Results

D.1 Solution Method & Simulation

For a given set of parameters, I solve the model using backward iteration. Agents die with certainty in the terminal period (T) and I can directly solve for their optimal consumption/saving combination for each point in the state space via grid search in period T. Having found the optimal choices in period T, I iterate one period backward and solve for the optimal choices in period T-1 and so forth. During retirement, I solve the problem independently for couples, single men, and single women. During working age, I need to take into account that individuals may switch marital status and hence, the continuation value of couples depends on the solution of the single problem (and vice versa).

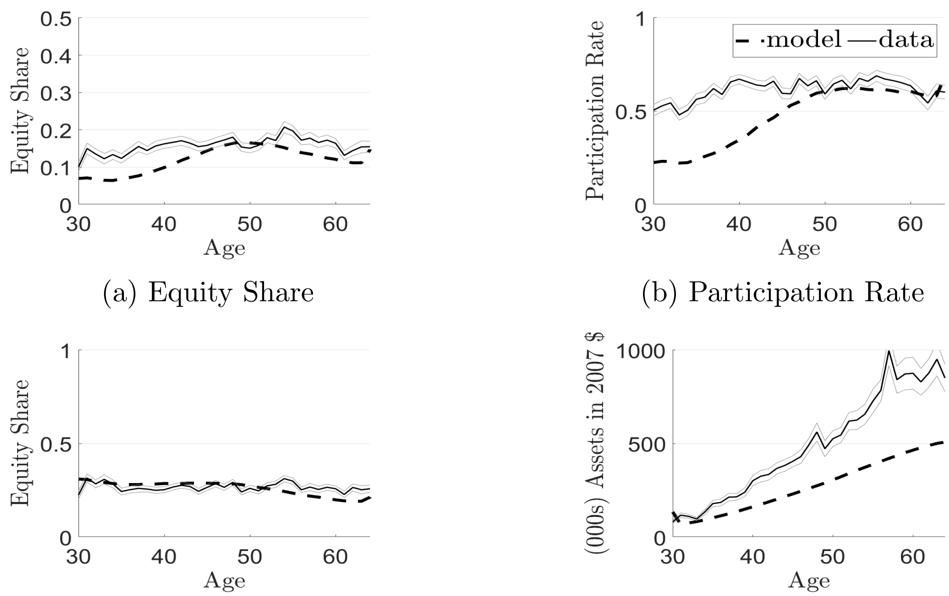
After having solved for the policy functions, I simulate the model for a large number of individuals over their life-cycle. At age 30, I assign each individual an initial level of wealth, education, and marital status. Next, I simulate a chain of the aggregate state, marital transitions, labor income shocks, and asset return realizations and assign each individual a certain chain of these shock processes. I simulate the model for 25,000 men and 25,000 women who may switch marital status throughout their working life. Once a single gets married, his or her partner is assigned from "outside" the model. Likewise, if the couple gets divorced, that partner again disappears from the simulation. Lastly, I construct the moments for each simulation, compute the objective function using the weighting matrix, and find the statistic of the model fit $\mathcal{L}$. Given the computational complexity of the structural framework, I can only repeat this procedure for a limited number of runs. Therefore, I use the TikTak global optimization algorithm (Guvenen and Ozkan 2021) to find a sequence of 70 sobol points at which I evaluate the model. Next, I further improve the model fit by searching within a narrow bound of parameters around the before found minimum.

D.2 Further Results on Model Fit

Figure 15 reports the model fit for couple households, which are left entirely untargeted in the calibration exercise. While the model matches the equity share of couples over the life-cycle well (Figure 15a), it underpredicts their participation rate early in life. Moreover, it misses the fast asset accumulation of couples.

The faster empirical asset accumulation can arise from various sources that go beyond the focus of my paper and are therefore absent in the model. For example, couples have on average more children than singles which may induce them to save for their children's education expenses. In addition, couples are more likely to be homeowners, generating high saving rates. Regardless of the underlying mechanism, the results of my paper do not depend on the model's ability to match the life-cycle profiles of couples: in a previous version, I improved the fit for couples by allowing their preference parameters (discount factor, coefficient of risk aversion, and stock market participation cost) to differ from those of singles, which did not alter the main results.

Figure 15: Model Fit of Couples



(a) Equity Share

(b) Participation Rate

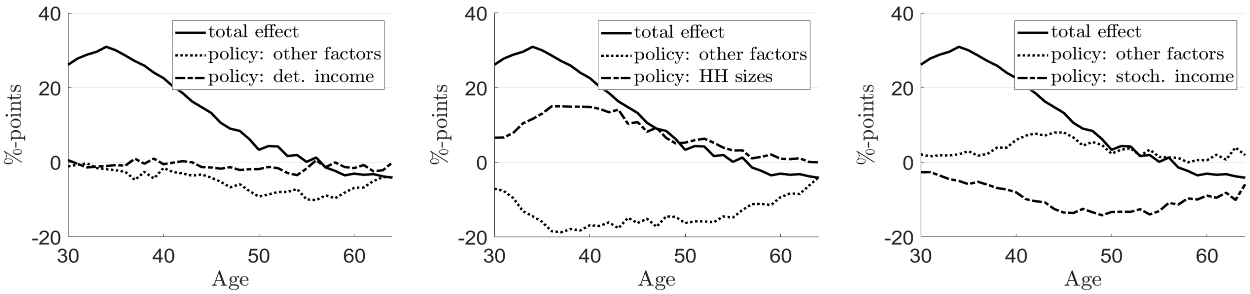
(c) Conditional Risky Share

(d) Net Worth

Notes: Figure 15 plots the model fit for the unconditional equity share, participation rate, conditional risky share, and net worth for couple households. The solid lines show the data (including 95% confidence bands, as plotted in Figure 1) whereas the dashed lines display the simulated life-cycle profiles generated from the model.

D.3 Composition vs. Policy Effect – Further Results

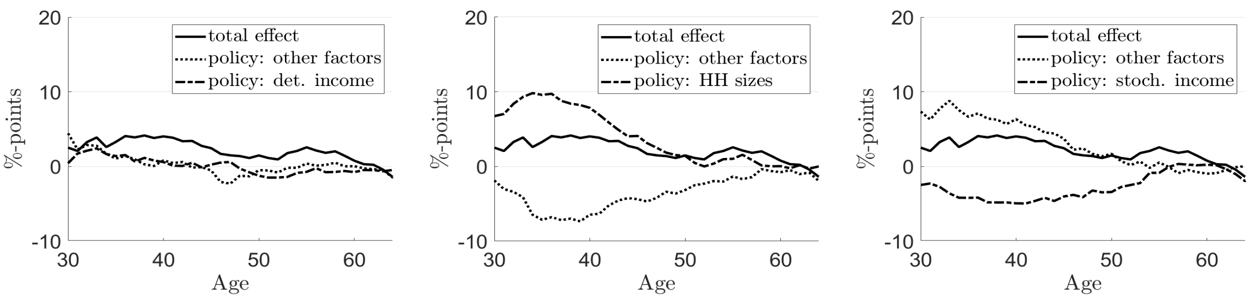
Figure 16: Decomposition of Policy Effect – Participation Rates



(a) Deterministic Income (b) Household Sizes (c) Stochastic Income

Notes: Figure 16 illustrates the importance of gender heterogeneity in deterministic income, household sizes, and stochastic income for the policy effect in participation rates. The solid lines refer to the total gender investment gap, the dashed-dotted lines show the part of the policy effect that can be attributed to the respective channel, whereas the dotted lines report the importance of all remaining channels (as listed in Table 5).

Figure 17: Decomposition of Policy Effect – Conditional Risky Shares



(a) Deterministic Income (b) Household Sizes (c) Stochastic Income

Notes: Figure 17 illustrates the importance of gender heterogeneity in deterministic income, household sizes, and stochastic income for the policy effect in conditional risky shares. The solid lines refer to the total gender investment gap, the dashed-dotted lines show the part of the policy effect that can be attributed to the respective channel, whereas the dotted lines report the importance of all remaining channels (as listed in Table 5).