The Gender Investment Gap over the Life-Cycle

Annika Bacher*

March 21, 2023

Click here for the most recent version.

Abstract

Single women hold less risky financial portfolios than single men. This paper analyzes the determinants of the "gender investment gap" based on a structural life-cycle framework. The model is able to rationalize the investment gap without introducing gender heterogeneity in preferences (e.g. in risk aversion). Rather, lower income levels and larger household sizes of single women are the main determinants for explaining the gap. Importantly, expectations about future realizations of both variables (that cannot easily be controlled for in regressions) drive most of the investment differences for young households whereas heterogeneity in observable characteristics explains the gap later in life.

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

*BI Norwegian Business School: annika.bacher@bi.no. I am indebted to my advisors, Russell Cooper and Thomas Crossley, for their continuous guidance and support. Moreover, I thank my discussants Francisco Gomes and Oliwia Komada 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, the University of Wisconsin-Madison, the 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, and the 2021 VfS Annual Meeting for comments and suggestions that helped to greatly improve this work. All remaining errors are mine. 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 and diversification gains, 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. Generally, differences in investment behavior can arise due to differences in circumstances (such as income profiles, number of household members, etc.) or due to differences in unobservable characteristics such as preferences. In fact, there exists a large empirical literature documenting higher degrees of risk aversion for women with regard to financial choices (see for example Eckel and Grossman (2008), Croson and Gneezy (2009), or Charness and Gneezy (2012) for a review) which would be a natural candidate explanation for a lower female equity share.[1]

However, by being the first paper to analyze the question through the lens of a structural model, I show that such a framework is able to match the empirical gender investment gap *without* introducing gender heterogeneity in preferences. It is rational for a single woman who has the same level of risk aversion as a single man and who shares the same observable characteristics to invest less risky because she expects to earn less in future periods and because she is more likely to have children living with her in the future. Consequently, reduced form regressions that control for household observable characteristics but do not take into account these expectations fail to fully explain the empirical gender investment gap.

In the following, I first document life-cycle profiles of asset holdings and portfolio choices for single men, single women, and couples using survey data on US 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

[1] See the "Related Literature" section for a more detailed discussion on how I relate to these papers.

toward risky assets. All differences are statistically different from zero, even after controlling for a wide range of observable characteristics that have previously been shown to affect investment behavior. The unexplained part of the gender investment gap is largest among young households and declines over the life-cycle.

Next, to uncover which factors explain 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 levels (i.e. in the deterministic part of their income profiles), their income risk (i.e. the stochastic part of their income profiles), the number of individuals who live in their household (e.g. children), their marital transitions probabilities, the (expected) characteristics of their partner in the event of marriage as well as their survival probabilities and out-of-pocket medical expenditures during retirement. In contrast, I restrict preference parameters to be identical across all types of households.

I estimate and 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 wealth holdings and equity shares for both single men and single women. By means of counterfactual exercises, I show that heterogeneity in income levels and in the average number of household members (household sizes) are the most important determinants of the gender investment gap.

With regard to the income level channel, not only differences in current income levels are important to explain this result but also the fact that single women expect to earn less than their male counterparts in future periods. Merton (1969, 1971) shows that 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 financial wealth.[2] Hence, even if a man and a woman

[2] This results holds if the correlation between labor income (human capital) and asset returns is small because then, human capital acts as a substitute for the safe asset. Cocco (2005) and Davis and Willen

have the same level of financial wealth (and the same current period income), it is optimal for the woman to invest less risky if she expects to earn less in future periods, that is, if she is endowed with less human capital.

In addition, larger female household sizes – which arise mainly through a higher likelihood of having children living in the same household – are an important determinant in explaining the observed gap. Again, not only current household sizes affect savings and equity shares (through different consumption needs) but larger expected household sizes act as a future consumption commitment that makes single women more vulnerable to financial shocks. As a result, they reduce financial risk-taking.

I then decompose the gender investment gap into a *composition* and into a *policy* effect. The composition effect explains how much of the gap arises through 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 decision rules for equity shares conditional on the state vector, that is by (expected) differences in future state variables.

Early in life, single men and single women are still relatively similar in terms of observable characteristics and most of the gender investment gap is driven by the policy effect. Single women expect on average to earn less and to have larger household sizes than single men in the future, making it optimal for them to invest less risky, even if they share the same (current) observable characteristics. However, as households age, there are fewer periods left to form expectations over. Therefore, from around age 50 onward, the gender investment gap is mostly driven by differences in observable characteristics, i.e. in the distribution of individuals across the state space. As a result, reduced form estimates that do not control for expectations have less predictive power in explaining the gender investment gap early in life when these expectations are most important, which is in line with my empirical results.

(2014) confirm that the correlation between stock return and idiosyncratic labor income shocks in the data is close to zero.

Lastly, I provide direct empirical support of gender differences in expectations that align with the predictions from the structural model. By complementing the analysis with data from the New York Fed Survey of Consumer Expectations, I show that single women expect lower future earnings than single men, after controlling for current earnings, and that this gender gap is more pronounced among young households. In addition, single women are more likely to expect to live with additional household members who are not prospective spouses.

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 that they choose less risky portfolio compositions in retirement accounts.[3] 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). Finally, Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (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 for instance 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.[4] If differences are purely preference driven, both men and women behave 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 consumption commitments and subsequently

[3] See e.g. Sunden and Surette (1998), Agnew, Balduzzi, and Sunden (2003), Arano, Parker, and Terry (2010), Säve-Söderbergh (2012) on retirement accounts and Andersen, Marx, Nielsen, and Vesterlund (2021), Girshina, Bach, Sodini, and Team (2021), Goldsmith-Pinkham and Shue (2023) on gender gaps in real estate markets.

[4] Such policies may include for example 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).

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. In general, there is consensus that women invest less risky than men. Jianakoplos and Bernasek (1998) document lower equity shares among single women than among single men in US data. Sunden and Surette (1998) and Agnew et al. (2003) show that women in the US choose lower equity allocations in retirement saving plans. Arano et al. (2010) cannot confirm significant gender differences in retirement accounts for US 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 include stocks less frequently in their pension contribution plan, they do allocate a smaller share into risky assets. Almenberg and Dreber (2015) and Thörnqvist and Olafsson (2019) 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 stock market participation rates. Moreover, several papers document that women earn lower returns in real estate markets (Andersen et al., 2021, Girshina et al., 2021, 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, this paper relates 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 sight, my results seem to contradict this literature because my model can replicate the gender investment gap without having to introduce heterogeneity in risk aversion. However, consistent with prior experimental and survey evidence, single women in my framework also behave observationally differently than single men, conditional on state variables, that is, conditional on *current* observable characteristics. The structural analysis then reveals that heterogeneity in expectations about future income levels 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 results.

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) was the first 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, Maurer, and Mitchell (2015) extend the analysis by incorporating endogenous labor supply and realistically calibrated social security benefit claiming. Christiansen, Joensen, and Rangvid (2015) empirically address the heterogeneous impact of family shocks on portfolio choices across gender using an administrative panel dataset from Denmark. Similar to Love (2010) for the US, their findings suggest that the fraction of risky assets in women's portfolios increases after marriage whereas it declines after a divorce. For men, this relationship points in the opposite direction. Along the same lines, Bertocchi, Brunetti, and Torricelli (2011) find in an empirical framework that the marital gap of stock holdings in Italy is larger for women than for men. 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 pattern of household finances (for a literature review see Poterba and Samwick (2001) and Gomes (2020)). Life-cycle models of portfolio choice typically predict the optimal equity share to be increasing in the ratio of the present value of human capital over current financial wealth (Merton, 1969, 1971, Viceira, 2001). Consequently, it should be optimal for young investors to allocate 100% of their financial wealth into stocks and to decrease the equity share as they age. In contrast, we observe only limited stock market participation and (conditional) equity shares, especially for young investors, in 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), and cyclicality of labor income (Catherine, 2022). However, so far little focus has been on marital transition risk as an additional source of financial uncertainty that limits the propensity of (young) investors to take risk in the stock market.

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 shows 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

The following section first describes the data and the sample selection criteria. Next, I provide empirical evidence on portfolio choices of 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 financial choices of households. Throughout the analysis, I restrict the sample to individuals between 30 and 65 years. The SCF is a triennial repeated cross-sectional survey sponsored by the Federal Reserve Board. It is carried out at the household level but collects individual demographic characteristics and income variables as well as detailed information on joint asset holdings of the household.

For income variables 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 US running from 1968 until today.[5] Besides the core sample, the PSID oversamples low-income families (the 'SEO' sample) and immigrant families (the 'immigrant' sample).

I combine two datasets for my analysis because I need both detailed portfolio choice information as well as panel data on household income (to estimate the income processes). Unfortunately, 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 the comparability of the sample across datasets, I show in Appendix A.1 that life-cycle profiles of variables that are available in both datasets look very similar.[6]

[5] 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 were collected annually until 1997 and afterwards every two years.

[6] Combining multiple datasets to estimate structural models is not uncommon in the literature. For instance,

Moreover, to ensure the representativeness of the US 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. All financial variables are converted into 2007 dollars using the CPI-U.

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 total, the PSID sample consists of 100,907 individual-year observations (82,705 for couples, 7,057 for single men, and 11,145 for single women) that correspond to 2,091 unique single women, 1,624 unique single men and 11,376 individuals who live in couples. The data drawn from the SCF includes information on 23,496 individuals in couples, 4,088 single men, and 6,155 single women.

2.2 Life-Cycle Profiles of Portfolio Allocation

Financial assets are defined as overall wealth net of housing assets and debt. 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.[7]

Figure 1a displays the life-cycle profiles of equity shares for single men, single women, and couples.[8] The equity share combines the extensive margin (whether or not the household owns any risky assets) with the intensive margin (conditional on holding risky assets, what portfolio share is allocated to them). Figure 1b and Figure 1c separately plot the stock market participation rate (only the extensive margin) and the conditional risky share (only the intensive margin), respectively. The gender differences are statistically different from zero, in particular during young age, as displayed by the confidence bands and by the corresponding

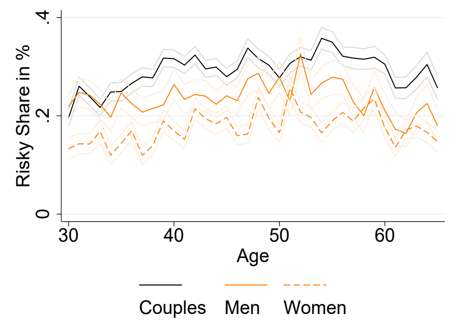
Cooper and Zhu (2016) combine the PSID and SCF to estimate the effect of education on stock market investment. Borella, De Nardi, and Yang (2023) use the PSID together with the Health and Retirement Study (HRS) to study lifetime outcomes during working life and retirement.

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

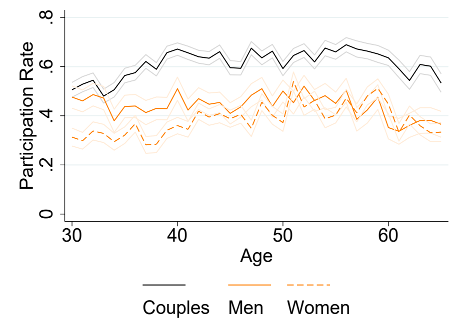
[8] To account for cohort effects, Appendix A.2 replicates Figure 1 for individuals born within a relatively short time-frame.

regression coefficients in Table 1.

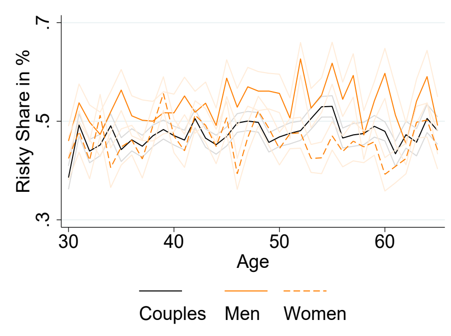
Figure 1: Life-Cycle Patterns of Household Finances (Data)



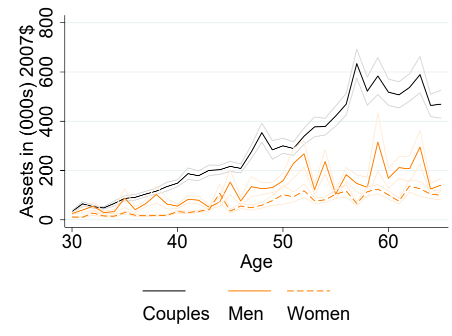
(a) Equity Share



(b) Participation Rate



(c) Conditional Risky Share



(d) Financial Assets

Notes: Figure 1 plots the life-cycle profiles of the equity share, stock market participation rate, conditional risky share, and financial assets 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.

Figure 1a shows that the equity share of single women is lower than that of single men during their entire working life. On average, the equity share of single women is around 6%-points lower than that of men which – given an average male equity share of 23.79% – corresponds to being 26.90% lower and roughly remains constant over the life-cycle. In contrast, the observed gender gap in stock market participation rates (Figure 1b) converges toward the entry to retirement.

Furthermore, the black solid line in Figure 1a shows that couples have on average a higher equity share than singles which is mainly driven by the extensive margin (see the black solid lines in Figures 1b and 1c, respectively). This finding is partly mechanical as couples are composed of two individuals for whom I compute the joint probability of participation. If I randomly draw a single man and a single woman and compute the likelihood that at least one of them holds risky assets (conditional on age), the participation rate of such a "generated couple" closely aligns with the one of couples in the data.

Finally, Figure 1d confirms that single women accumulate less wealth than single men also in absolute terms. This gap is often referred to as the "gender wealth gap". Throughout their working life, the gap in financial wealth is on average $59,280 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 circumstances or due to differences in 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. In particular, 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 the literature has shown to be important predictors of portfolio choices. Following Christelis, Georgarakos, and Haliassos (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.[9] Column (3) furthermore includes the inverse hyperbolic sine transformation of the households' safe financial assets, whereas Column (4) additionally controls for occupation and industry fixed effects. The corresponding marginal effects of being a single woman along with their standard errors at various ages are reported in the last three rows of Table 1.[10]

The coefficient for being a single woman is negative (and statistically significant) across all

[9] 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.

[10] Appendix A.4 reports the corresponding specifications separately for the participation rate and the conditional risky share.

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. When considering the marginal effect of being a single woman ("ME"), I find a negative and significant gender effect across all four columns. However, as individuals age, this "negative" effect of being a woman on the equity share becomes smaller.

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 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 to account for factors that cannot be easily controlled for in reduced form specifications, such as expectations and risk exposure.

2.4 On the (Non-)Presence of Housing

The focus of this paper is on liquid financial wealth which is why I abstract from housing. However, housing constitutes a large share of households' portfolios and affects stock market behavior.[11] For the purpose of the current analysis, abstracting from housing is a problem if either 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, I would expect the life-cycle profiles of housing variables to closely follow those in Figures 1a to 1c. Figure 2 displays singles' life-cycle profiles of homeownership rates, housing wealth (henceforth: "HW"), and

[11] 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, Sandor, and Szeidl (2017) or Paz-Pardo (2021) to name a few.

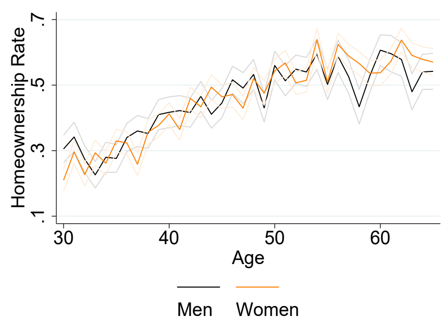
Table 1: Regression Coefficients & Marginal Effects – Equity Shares of Singles

	(1) Equity Share	(2) Equity Share	(3) Equity Share	(4) Equity Share
single woman	-0.3911$^{\diamond}$	-0.2295$^{\diamond}$	-0.1113$^{\diamond}$	-0.0734$^{\diamond}$
	(0.0103)	(0.0095)	(0.0110)	(0.0029)
single woman*age	0.0057$^{\diamond}$	0.0032$^{\diamond}$	0.0012$^{\diamond}$	0.0003$^{\diamond}$
	(0.0002)	(0.0002)	(0.0002)	(0.0001)
age	-0.1579$^{\diamond}$	-0.0886$^{\diamond}$	-0.0630	-0.0579
	(0.0269)	(0.0265)	(0.0324)	(0.0355)
$age^2 * 100$	0.3854$^{\diamond}$	0.2295$^{\diamond}$	0.1636$^{\diamond}$	0.1496
	(0.0581)	(0.0574)	(0.0702)	(0.0770)
$age^3 * 10000$	-0.3014$^{\diamond}$	-0.1851$^{\diamond}$	-0.1374$^{\diamond}$	-0.1187$^{\diamond}$
	(0.0407)	(0.0402)	(0.0493)	(0.0538)
high education		0.3796$^{\diamond}$	0.1951$^{\diamond}$	0.1236$^{\diamond}$
		(0.0037)	(0.0042)	(0.0063)
number of HH members		-0.0727$^{\diamond}$	-0.0528$^{\diamond}$	-0.0493$^{\diamond}$
		(0.0015)	(0.0021)	(0.0029)
non-asset income		0.0478$^{\diamond}$	0.0307$^{\diamond}$	0.0295$^{\diamond}$
		(0.0009)	(0.0007)	(0.0007)
safe assets			0.0866$^{\diamond}$	0.0574$^{\diamond}$
			(0.0009)	(0.0012)
constant	2.0325$^{\diamond}$	0.2078	-0.5683	-0.4294
	(0.4000)	(0.3914)	(0.4748)	(0.5199)
Observations	10,243	10,239	10,239	7,606
Year FE	No	Yes	Yes	Yes
Industry & Occupation FE	No	No	No	Yes
ME for women at age 30	-0.2191$^{\diamond}$	-0.1334$^{\diamond}$	-0.0746$^{\diamond}$	-0.0632$^{\diamond}$
	(0.0048)	(0.0038)	(0.0045)	(0.0015)
ME for women at mean age (47)	-0.1195$^{\diamond}$	-0.0778$^{\diamond}$	-0.0533$^{\diamond}$	-0.0578$^{\diamond}$
	(0.0031)	(0.0028)	(0.0027)	(0.0032)
ME for women at age 65	-0.0184$^{\diamond}$	-0.0213$^{\diamond}$	-0.0317$^{\diamond}$	-0.0513$^{\diamond}$
	(0.0047)	(0.0055)	(0.0053)	(0.0055)

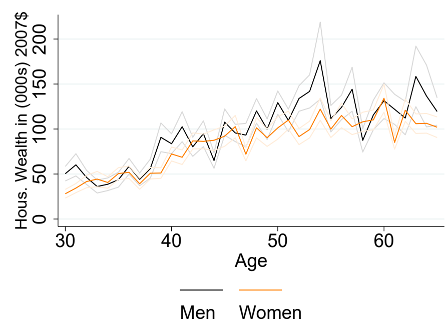
Notes: Estimations are based on Tobit regressions on the sample of individuals who live in households with no spouse present. Source: SCF waves 1989 until 2016. Equity Share = Unconditional risky share. *single woman* is a dummy indicating that the household head is a woman. *high education* is a dummy equal to one if the household head has more than 12 years of education. *safe assets* refers to safe liquid assets. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

housing wealth-to-income ratio (henceforth: "HI"). For all three variables, I do not find any significant differences between men and women despite significant gender gaps in (financial) portfolio choices both along the extensive and intensive margin. Moreover, the life-cycle patterns for housing variables are different than those of portfolio choices: neither housing graph displays a relatively flat life-cycle profile (as for the equity share and for the conditional risky share), nor a converging gender gap (as for the stock market participation rate).

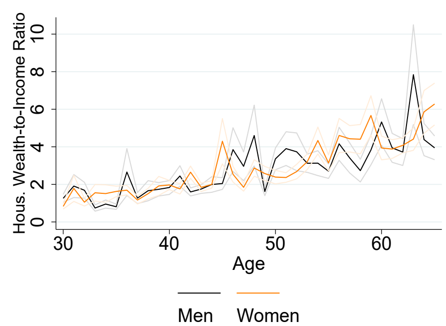
Figure 2: Life-Cycle Profiles of Housing Patterns (Singles)



(a) Homeownership Rate



(b) (Gross) Housing Wealth



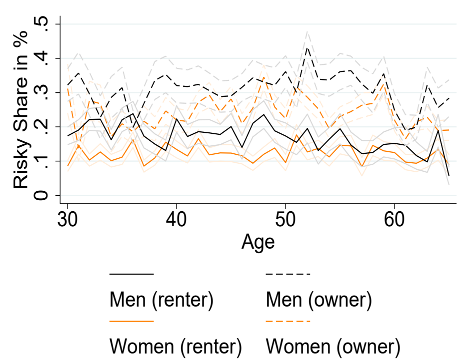
(c) Housing Wealth-to-Income Ratio

Notes: Figure 2 plots life-cycle profiles of the homeownership rate, gross housing wealth, and the housing wealth-to-income ratio for single men and single women, including 95% confidence intervals. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF).

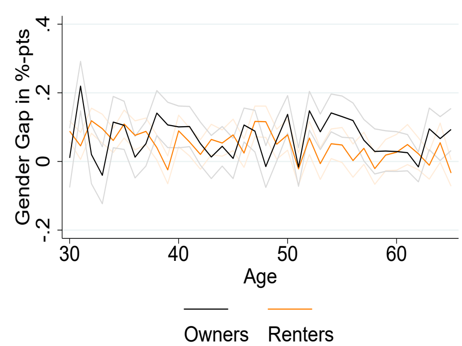
Second, if housing is an important driver of the gender investment gap itself, the gender gap in equity shares should differ by housing tenure. To test for this possibility, Figure 3a plots the equity share of single homeowners and single non-homeowners (renters) over their life-cycle, separately by gender.

In line with previous literature, equity shares differ by housing tenure, in that homeowners hold on average more risky portfolios (i.e. the dashed lines lie above the solid lines in Figure 3a). However, *gender differences* in equity shares (i.e. the gap between black and orange lines) are very similar for homeowners and renters. If anything, the gender gap in equity shares is slightly larger among homeowners toward the end of the life-cycle. To further illustrate this finding, Figure 3b plots the gender investment gap for renters and owners by age. Both lines behave very similarly and are not statistically significant different from one another, reassuring that housing choices do not differently affect portfolio choices of single men and single women.

Figure 3: Gender Gaps in Equity Shares by Housing Tenure



(a) Equity Shares



(b) Gender Investment Gap

Notes: Figure 3a plots singles' life-cycle profiles of equity shares by gender and by housing tenure. Figure 3b plots the gender gap in equity shares for homeowners and renters, respectively. The gender gap in Figure 3b is defined as the average equity share of single men minus the average equity share of single women at the respective age, i.e. the difference between black and orange lines in Figure 3a. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF).

3 A Life-Cycle Model of Portfolio Choice

In this section, I develop a stochastic life-cycle model with women and men (denote gender by $i = \{f, m\}$) who live either as singles ($\mathcal{S}$) or as a married couple ($\mathcal{M}$). 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 age 65 and die deterministically at 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 (i.e. having completed high school) or not in the data.

During working age, households are subject to uninsurable labor income shocks that depend on their gender and on their marital status. 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). Couples decide jointly on the level of consumption ($c_{\mathcal{M}}$) and how much to save in both types of assets ($a_{\mathcal{M}}^s, a_{\mathcal{M}}^r$). Moreover, singles face an exogenous marriage probability each period that depends on their gender, age, and education. Likewise, couples face an exogenous divorce probability that varies by age and by both spouses' education.

During retirement, agents face age- and gender-dependent medical expenditures and are subject to longevity risk that depends on their age j. Upon dying, agents value leaving bequests. As during working age, they can live either as singles or couples, however, 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 to account for increased medical expenditures in the year before death as well as for bequests to non-spousal heirs. As before, households have a portfolio choice between a safe asset and a 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:

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

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. In the event of death, individuals derive utility from leaving bequests according to:

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

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 hence, values leaving bequests in the case of his or her own death.

3.2 Dynamics

Asset Returns. The safe asset pays a time-invariant return r_s. The return of the risky asset is drawn from the distribution $r_r \sim N(\mu_r, \sigma_r^2)$ that is assumed to be i.i.d. and for which $\mu_r > r_s$. As in Fagereng, Gottlieb, and Guiso (2017), I allow for the possibility of stock market crashes and augment the return of the risky asset by a "disaster" state. With probability $(1 - p_{tail})$ the return is drawn from the above normal distribution and with probability p_{tail} a tail event $r_{tail} < \underline{r}_r$ materializes.[12] Short-selling and borrowing are not allowed.

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

$$y_{ij} = \bar{y}_i \theta_i \xi_{ij} \tilde{y}_{ij}$$

The term $\bar{y}_i$ denotes a constant, θ_i is the (exogenous) education premium, and ξ_{ij} stands for an age-specific component. The term $\tilde{y}_{ij}$ represents the stochastic component of income and consists of a transitory and a persistent shock:

$$\tilde{y}_{ij} = z_{ij} + \epsilon_{\tilde{y}ij} \quad \text{with} \quad z_{i,j+1} = \rho_{zi} z_{ij} + \nu_{zij}$$

[12] The term $\underline{r}_r$ denotes the lowest possible realization of the risky asset return that is drawn from the (discretized) normal distribution.

where $\epsilon_{\tilde{y}ij}$ and ν_{zij} are independent zero mean random shocks with variances $\sigma^2_{\tilde{y}i}$ and σ^2_{zi} respectively. The parameter $\rho_{zi} \in (0, 1]$ captures the persistence of shock ν_{zi}.

Within couples, the transitory shocks $\epsilon_{\tilde{y}fj}$ and $\epsilon_{\tilde{y}mj}$ are assumed to be correlated (with $\rho_{\sigma_{\tilde{y}f},\sigma_{\tilde{y}m}} = 0.3$) as spouses live in the same area and are likely to work in similar industries and are thus subject to correlated labor market shocks.[13] In contrast, following Cocco (2005) and Davis and Willen (2014), labor income shocks are uncorrelated with realizations of the stock return.

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 because medical expenditures are strictly increasing in age, deterministic medical expenditures impose a source of risk in the sense that agents are uncertain whether or not they live until a certain age and have to pay the corresponding medical bills. This modeling choice is motivated by De Nardi, French, and Jones (2010) who show that the main sources of risk during retirement are not fluctuations of medical expenditures around its mean but rather their age-dependent level combined with longevity risk.

3.3 Stock Market Participation Cost

Agents have to pay a fixed cost S^F each period if they choose to invest part of their savings in the risky asset. As in 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 flow costs do not require introducing stock holdings as a state variable.

[13] 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 US data of 0.22 for the age group 25-34, 0.36 for ages 35-44, and 0.42 for couples above 45 years old.

3.4 Marriage and Divorce

Single individuals get married with an exogenous probability that depends on their gender, age, and education. Denote this marriage probability by $\mu(i, j, \theta)$. Conditional on meeting a partner, the probability of meeting a partner with education θ_p and income shock realization $\tilde{y}_p$ is:

$$\Pi(.) = \Pi(\theta_p, \tilde{y}_p | \theta_i, \tilde{y}_i)$$

Both partners always 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 as we observe it in the data. Couples face an exogenous divorce probability each period that depends on age and the education of each spouse $\lambda(j, \theta_f, \theta_m)$. Upon divorce, assets are split equally between spouses and 10% of assets are destroyed to account for legal fees of divorce and general costs of asset splitting between spouses.[14] There are no alimony payments.

3.5 Timing

At the beginning of period t, agents learn their current productivity state(s), their stock market return as well as their marital status. Thus, agents start period t with a given amount of savings that depends on their decisions in period $t-1$, their marital status, and the realization of the asset return state. After observing all shock realizations, agents decide on how much to consume and how much to save in both the risky and the safe asset. When investing part of their endowment in the risky asset, they have to pay S^F in the current period t.

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

3.6 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, all during working age and during 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, De Nardi, and Yang, 2020). Moreover, because the stock market participation cost has to be paid per period and given the i.i.d. nature of the return process for the risky asset, I can combine safe and risky assets into one "asset cash-in-hand" state variable: $a = (1 + r_r)a_r + (1 + r_s)a_s$.

Singles – Working Age. The state variables of a single agent are her gender i, age j, education θ, asset cash-in-hand a, and her current income realization $\tilde{y}$. The corresponding value function reads as:

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

subject to:

$$a'_r + a'_s + c = y(j, \theta_i, \tilde{y}_i) + a - \mathbb{1}_{a'_r > 0} S^F \quad \text{with} \quad a = (1 + r_r)a_r + (1 + r_s)a_s$$

and:

$$\tilde{y} = z + \epsilon_{\tilde{y}} \quad \text{with} \quad z' = \rho_z z + \nu_z \quad \text{and} \quad \epsilon_{\tilde{y}} \sim N(0, \sigma^2_{\tilde{y}}), \nu_z \sim N(0, \sigma^2_z)$$

$$r_r = \begin{cases} \mathcal{N} \sim (\mu_r, \sigma^2_r) & \text{with probability } (1 - p_{\text{tail}}) \\ r_{tail} < \underline{\mathrm{r}}_r & \text{with probability } p_{\text{tail}} \end{cases} \quad \text{with} \quad \mu_r > r_s \quad \text{and} \quad \tilde{y} \perp r_r$$

where η_{ij} denotes an equivalence parameter that controls for changing family size over the life-cycle. $\hat{V}^C$ expresses the value of individual i of getting married to partner p. Single individuals take the expected value over future productivity realizations and asset returns when staying single whereas they form expectations over future productivity realizations, asset returns, and their specific partner in case of getting married.

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

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

subject to:

$$a_r' + a_s' + c = pen_s(\hat{y}) + a - m_{ij} - \mathbb{1}_{a_r'>0}S^F \quad \text{with} \quad a = (1+r_r)a_r + (1+r_s)a_s$$

and:

$$r_r = \begin{cases} \mathcal{N} \sim (\mu_r, \sigma_r^2) & \text{with probability } (1-p_{\text{tail}}) \\ r_{tail} < \underline{\mathrm{r}}_r & \text{with probability } p_{\text{tail}} \end{cases} \quad \text{with} \quad \mu_r > r_s \quad \text{and} \quad \tilde{y} \perp r_r$$

where ψ_{ij} and m_{ij} denote age- and gender-dependent survival probability and medical expenditures, respectively. Retired singles take the expected value over their next-period asset return as well as their likelihood of survival.

Couples – Working Age. The state variables of a couple that consists of a woman f and a man m can be summarized by their age j, education of both spouses θ_f, θ_m, their joint asset holdings a, as well as both productivity realizations $\tilde{y}_f, \tilde{y}_m$. The corresponding value function reads as:

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

subject to:

$$a'_r + a'_s + c = \sum_{i=f,m} y(j, \theta_i, \tilde{y}_i) + a - \mathbb{1}_{a'_r > 0} S^F \quad \text{with} \quad a = (1 + r_r)a_r + (1 + r_s)a_s$$

and:

$$\tilde{y}_i = z_i + \epsilon_{\tilde{y}_i} \quad \text{with} \quad z'_i = \rho_{zi} z_i + \nu_{zi} \quad \text{and} \quad \epsilon_{\tilde{y}_i} \sim N(0, \sigma^2_{\tilde{y}_i}), \nu_z \sim N(0, \sigma^2_{zi}) \quad \text{for} \quad i = \{f, m\}$$

$$\begin{pmatrix} \epsilon_{\tilde{y}f} \\ \epsilon_{\tilde{y}m} \end{pmatrix} \sim \begin{pmatrix} \sigma^2_{\tilde{y}f} & \rho_{\sigma_{\tilde{y}f}, \sigma_{\tilde{y}m}} \\ \rho_{\sigma_{\tilde{y}f}, \sigma_{\tilde{y}m}} & \sigma^2_{\tilde{y}m} \end{pmatrix}$$

$$r_r = \begin{cases} \mathcal{N} \sim (\mu_r, \sigma^2_r) & \text{with probability } (1 - p_{\text{tail}}) \\ r_{tail} < \underline{\mathrm{r}}_r & \text{with probability } p_{\text{tail}} \end{cases} \quad \text{with} \quad \mu_r > r_s \quad \text{and} \quad \tilde{y} \perp r_r$$

Couples take the expected value of both partners' future productivity realizations and joint asset returns when staying married as well as the respective individual's productivity realization and asset returns when getting divorced.

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

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

subject to:

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

and:

$$
r_r = \begin{cases} \mathcal{N}\sim(\mu_r,\sigma_r^2) & \text{with probability } (1-p_{\text{tail}}) \\ r_{tail} < \underline{\mathrm{r}}_r & \text{with probability } p_{\text{tail}} \end{cases} \quad \text{with} \quad \mu_r > r_s \quad \text{and} \quad \tilde{y} \perp r_r
$$

Retired couples take the expected value over their joint asset returns as well as the individual survival probabilities of both spouses.

Value to an individual of becoming a couple. The value of an individual in a couple is the relevant object when computing the value of single i for getting married to partner p, i.e. the present discounted value of the individual's utility in the event of marriage (Borella et al., 2020). In this context, variables denoted with a $\hat{hat}$ indicate 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.

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

4 Estimation & Calibration

I estimate and calibrate the model in a two-step strategy following Gourinchas and Parker (2002) and Cagetti (2003). First, I estimate all parameters that can be cleanly identified directly 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.

First stage parameters include initial distributions, parameters related to medical expenditures, the labor income process, survival probabilities, and asset returns. I borrow the parameters for the bequest motive (ω, L) from Cooper and Zhu (2016). Second stage parameters include the discount factor β, the coefficient of relative risk aversion γ, and the stock market participation cost S^F. I collect the second stage parameters in the vector $\Theta = \{\beta, \gamma, S^F\}$.

4.1 First Stage Estimation

Income Profiles. Figure 4 shows the life-cycle profiles of the deterministic income component by gender and by marital status from the PSID. Income is expressed as annual income out of labor earnings (including labor 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, and transfers and add half of that from other household members.[15] Lastly, 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

[15] In some waves, the PSID does not separately report transfer income or social security benefits 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.

regress the inverse hyperbolic sine of income for an individual of gender i at age j,

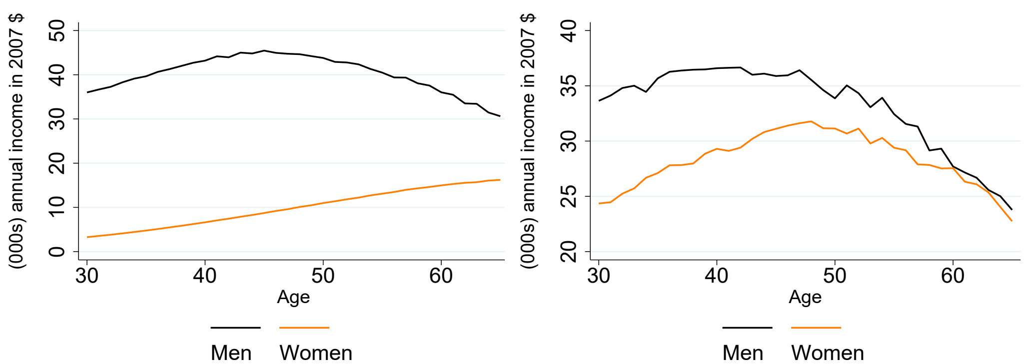
$$income_{ij} = \alpha + \beta_1 age_{ij} + \beta_2 age_{ij}^2 + \beta_3 woman_i * age_{ij} + \delta_i + u_{ij}$$

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

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

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

Figure 4: Life-Cycle Profiles of the Deterministic Income Component



(a) Couples (b) Singles

Notes: Figure 4 plots the life-cycle profiles of the deterministic part of labor income by gender and marital status. Data is from the waves 1989 until 2016 of the Panel Study of Income Dynamics (PSID).

The coefficients from these income equations (reported in Table 11 in Appendix B.2) inform me about the deterministic component of the income process in the model. Note that parts of the estimated age gradients are driven by variation in hours worked and by 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 observed life-cycle patterns in overall income for single households (Figure 4b).

I estimate the parameters governing the stochastic component of the income process using the minimum distance estimator from Guvenen (2009).[16] Table 2 summarizes the results. My point estimates imply a slightly less persistent income process for single women than for single men (i.e. when comparing the first two columns), whereas the variance of both the persistent shock σ_z^2 and the transitory shock $\sigma_{\tilde{y}}^2$ is lower for single women. The overall variance of single women's income process is lower than single men's, which may for example arise due to single women sorting into more stable occupations (Bertrand, 2011). In contrast, the income process of married women exhibits a much higher variance of the transitory shock $\sigma_{\tilde{y}}^2$ than that of singles and married men. When solving the model, I discretize the labor income shock using the Rouwenhorst method (Rouwenhorst, 1995).

Table 2: Estimation Results – Stochastic Income Process

Parameter	Men	Women	Men	Women
	Singles		Couples	
ρ_z	0.9522	0.9341	0.9392	0.9270
	(0.0079)	(0.0099)	(0.0051)	(0.0043)
σ_z^2	0.0912	0.0867	0.0826	0.1616
	(0.0122)	(0.0121)	(0.0056)	(0.0090)
$\sigma_{\tilde{y}}^2$	0.1681	0.1558	0.1391	0.2854
	(0.0354)	(0.0224)	(0.0117)	(0.0174)

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

Marital Transitions. Marital transitions are defined as the likelihood of getting married (respectively divorced) within the next period conditional on not being married (respectively being married) in the current period. More specifically, I estimate the following logit function,

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

separately for couples and singles:

$$\xi_{i,j+1} = \frac{exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta^s})}{1 + exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta^s})}$$

where $\xi_{i,j+1}$ denotes the probability for household i in period j of being married (respectively divorced) next period. As explanatory variables ($\boldsymbol{x}$), I include age, schooling, and a dummy for waves after 1997 to account for the switch from annual to biannual frequency in the PSID.[17] Table 12 in Appendix B.3 reports the corresponding regression coefficients ($\boldsymbol{\beta^s}$).

I find that the likelihood of both marriage and divorce declines over the life-cycle. At any given age, single women are less likely than single men to get married within the next year. The probability of marriage is increasing in education whereas divorce becomes less likely if both spouses have more than 12 years of schooling. Finally, I estimate the matching of spouses in terms of income, assets, and education (summarized by the term Π) non-parametrically directly from the PSID.

Out-of-Pocket Medical Expenditures. I borrow the parameters describing medical expenditures by age and gender from Borella et al. (2020). The authors estimate deterministic out-of-pocket medical expenditures profiles with data from the HRS separately for men and women. They estimate 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 US Social Security Administration.[18] The death probability at age j is defined as the probability to die within the next year conditional on having survived up to age j. I compute the inverse of those probabilities and work with average values between the years 1990, 2000, and 2010, corresponding to the sample period of my study. For couples, if the husband dies, the surviving wife keeps 60% of the household's assets, whereas a surviving husband keeps

[17] For couples, age refers to the household head.
[18] All tables available under this link.

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.[19]

Asset Returns. The annual return rate of the risk-free asset is 2%, taken from Catherine (2022). The return of the risky asset is drawn with a 98% probability from a normal distribution with mean 6% and a variance of $Var(\tilde{R}(s)) = \sigma_r^2 = (0.1758)^2$. The latter reflects the variance of the annual total return index of the S&P 500 from 1989 until 2016. With a 2% probability, a disaster state realizes that results in the loss of 40% of all assets, both values that Barro (2009) empirically estimates from historical US data on stock market crashes. Hence, the overall equity premium is 3.12%. When simulating the model for a large set of individuals over their life-cycle, I simulate the return of the risky asset as an aggregate shock.

Pension Payments. Pension payments are flat and assumed to be 60% of the income during the last year of work. Couples receive a common pension that is 1.7 times higher than that of single men.

Equivalence Scales. To compute the equivalence scales η, I first 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.

Initial Conditions. The initial distribution over asset holdings in the model is chosen such that it mimics the distribution of wealth across individuals at age 30 in the SCF. Similarly, 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. Finally, the initial distribution of couples and singles is set equal to PSID data for individuals at age 30.

[19] I choose these values because Jones, De Nardi, French, McGee, and Rodgers (2020) document that households who experienced the death of one spouse in the last years hold around 30% less assets than couples who did not experience a death, and that surviving single men are on average wealthier than surviving single women.

4.2 Second Stage Calibration

I borrow the bequest parameters from Cooper and Zhu (2016) which results in $L = 0.128$ and $\omega = 0.73$. Taking the parameters from the first stage as given, I calibrate the remaining structural parameters $\Theta = \{\beta, \gamma, S^F\}$ using the Simulated Method of Moments. The exercise is to find $\hat{\Theta}$ that solves the following optimization problem:

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

where W represents a weighing matrix, M^d 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 wealth levels to identify the discount factor β. Moreover, once households cross the threshold of stock market participation, the participation cost S^F becomes irrelevant for their decision on how much to 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 across participating households, that is, in the conditional risky share. The stock market participation cost S^F serves as the target 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 HH types), with only three parameters.

The Weighting Matrix W. I first calibrate the second-stage parameters by using a slightly modified identity matrix ($W = \mathcal{I}$). In particular, given the paper's focus on equity share, I place less weight (30%) on asset profiles than on the participation rate and the conditional equity share. In a second run, I 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

($W = \frac{1}{\mathcal{V}}$). This approach follows Cooper and Zhu (2016) and is in contrast to papers that use the standard variance-covariance matrix (e.g. Cagetti (2003) or Alan (2006)). In the current set-up, different moments are based on different sample sizes: while participation rates and wealth levels include all observations, the conditional risky share only includes stock market participants. Hence, I could only estimate covariances for the restricted sample of stockholders which is not necessarily more informative than the diagonal matrix.

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 a lower discount factor, a slightly lower coefficient of risk aversion, and lower participation costs. When using the inverse variance matrix ($W = \frac{1}{\mathcal{V}}$), life-cycle profiles of conditional equity shares receive the most weight, followed by participation rates, and lastly asset moments (that receive even less weight than when using the modified identity matrix). Hence, deviations from the asset profile have the least impact on the objective function, resulting in a larger β. As a consequence, a higher stock market participation cost is required to match empirical participation rates. Finally, a slightly higher coefficient of risk aversion is necessary to match conditional equity shares.

The calibration that uses the inverse variance matrix finds an annual stock market participation cost of \$575. With regard to the coefficient of risk aversion, my estimates suggest $\gamma = 4.6$ which is at the lower end of estimates introduced by previous papers of portfolio choice with a per-period participation cost. In contrast, my participation costs are rather high. Fagereng et al. (2017) estimate an annual stock market participation cost of \$69 but also introduce a very high degree of risk aversion of $\gamma = 11$. Catherine (2022) estimates a CRRA coefficient of $\gamma = 8$ and an annual stock market participation cost of \$400 but he includes housing wealth in the analysis.[20] The value for β (0.685) is relatively low but still in the range of values found in previous studies. For example, Cooper and Zhu (2016) estimate a discount factor of 0.869, Fagereng et al. (2017) of 0.77, and Catherine (2022) of 0.88. On the other hand, Bonaparte, Cooper, and Zhu (2012) estimate – very similar to my findings – a value for β of 0.686.

One reason why I find rather low values for the discount factor and high values for the stock market participation cost is that I exclude housing wealth and target the life-cycle profile of financial wealth instead of net worth. In addition, the possibility of divorce generates high-asset single households (who got divorced) that are absent in models with only one generic household type. Moreover, my model is able to match equity shares with relatively low degrees of risk aversion because marriage and divorce introduce an additional dimension of financial risk that lowers households' willingness to invest in the risky asset.

Table 3: 2nd Stage Parameters

W	β	γ	S^F
$\mathcal{I}$	0.651	4.388	\$534
$\frac{1}{\mathcal{V}}$	0.685	4.600	\$575

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

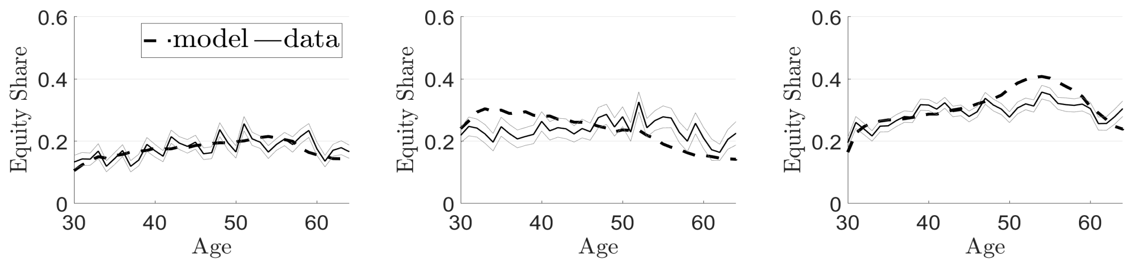
5.2 Model Fit

Figure 5 shows that the model matches well the life-cycle profiles of equity shares for single men, single women, and couples. Importantly, it is able to capture the gender investment gap without introducing preference heterogeneity by gender (e.g. in risk aversion). In addition, the model replicates the evolution of wealth (Figure 6) for single men and single women. For

[20] Catherine (2022) addresses the puzzle that life-cycle models of portfolio choice either require a very high degree of risk aversion (typically in combination with a very low discount factor) or a very high stock market participation cost to match the data by introducing cyclical skewness in labor earnings. To make his results comparable to mine, the listed values refer to the case when he estimates his model without cyclical skewness.

both, the asset accumulation profiles are slightly too concave when compared to the data: whereas it overpredicts asset accumulation early in life, it undershoots the increase in asset accumulation leading up to retirement. Appendix C.2 further discusses the model fit for participation rates, conditional risky shares, and couple households.

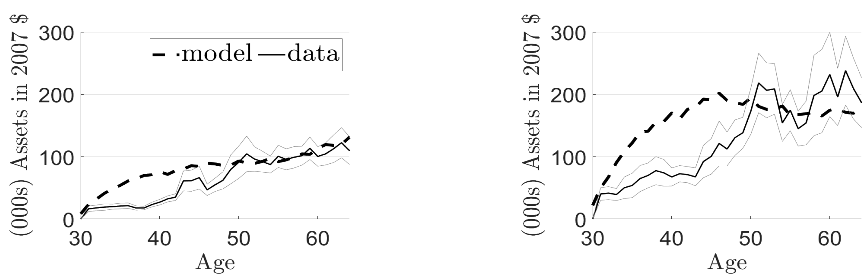
Figure 5: Model Fit of Equity Shares



(a) Single Women (b) Single Men (c) Couples

Notes: Figure 5 plots the model fit of equity shares for single women, single men, and couples. The solid lines show the data (including 95% confidence bands, as plotted in Figure 1) whereas the dashed line display the simulated life-cycle profiles generated from the model.

Figure 6: Model Fit of Asset Accumulation (Singles)



(a) Single Women (b) Single Men

Notes: Figure 6 plots the model fit of asset accumulation 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.

5.3 Simulated Regressions

To compare the reduced form regressions from Section 2.3 with the model, Table 4 replicates the same regressions on generated data from model simulations. Note that all of these coefficients are untargeted in the calibration exercise.

The model slightly over-predicts the effect of gender on the equity share (Columns (1) and (2)), meaning that the absolute values for the coefficient "single woman" and for its interaction term with age are larger than in the data. However, the 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. Hence, it appears that either factors that cannot easily be controlled for in reduced form regressions (such as expectations) or non-linearities account for the residual part of the gap (which I will explore in more detail in Section 6).

Columns (3) and (4) report the estimated gender effect on the stock market participation rate. In line with the data, the model predicts the marginal effect of being a woman to be negative and to be increasing in age. In both specifications, the marginal effect of being a woman on the participation rate accounts for around 65% of the marginal effect on the overall equity share at age 30, and the effect becomes smaller as households age. When considering the conditional risky share (Columns (5) and (6)), the model produces a negative baseline effect of being a woman (as opposed non-statistically significant coefficient in the data), and a negative interaction term (as opposed to a positive one in the data). As a result, 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.

In addition, the model predicts the equity share to be increasing in income, which is mainly driven by the participation rate. All else equal, higher income helps households to pay the stock market participation cost. However, conditional on participating, the optimal risky share is increasing in the ratio of human capital over financial wealth, which renders the effect of income on the conditional risky share ambiguous: on the one hand, higher income today increases human capital (since the income shock is persistent), but on the other hand, it also increases the chance that the household has accumulated more financial wealth.

With model simulated data, the coefficient of age on the equity share is negative, that of its squared term positive, and that of its cubic term negative. In the model, age is a strong predictor for asset accumulation, which displays a concave pattern that is slightly increasing toward the end of working life (see Figure 6).

Finally, the education coefficient in the model is positive for the equity share and conditional risky share whereas it is negative for the participation rate. After controlling for income, education affects households in the model through its impact on marriage probabilities and expected income levels. For example, more educated households are more likely to get married, which in turn reduces their precautionary saving motive and makes them less likely to cross the participation threshold. On the other hand, education increases expected future income, and thus increases households' optimal risky share.

6 Understanding the Mechanisms

6.1 Decomposing the Gender Investment Gap

In this section, I decompose the gender gap in equity shares and wealth levels along the dimensions of gender heterogeneity in the model, that is along income levels (i.e. the deterministic part of the income process), income risk (i.e. the stochastic part of the income process), 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, differences in 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 shows the results. The column "Model" reports the gender investment gap in the respective counterfactual whereas the column "% explained" indicates how much of the baseline gap can be explained through the respective channel.

Decomposing the Gap in Wealth Levels. The upper panel of Table 5 shows that

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	-0.5206$^\diamond$	-0.2295$^\diamond$	-0.3061$^\diamond$	-0.1775$^\diamond$	-0.1019$^\diamond$	0.0239
	(0.0423)	(0.0095)	(0.0355)	(0.0096)	(0.0167)	(0.0222)
single woman*age	0.0082$^\diamond$	0.0032$^\diamond$	0.0043$^\diamond$	0.0031$^\diamond$	0.0027$^\diamond$	-0.0019$^\diamond$
	(0.0008)	(0.0002)	(0.0007)	(0.0002)	(0.0003)	(0.0005)
age	0.1403$^\diamond$	-0.0886$^\diamond$	0.1088$^\diamond$	-0.0926$^\diamond$	-0.0518$^\diamond$	0.0502$^\diamond$
	(0.0331)	(0.0265)	(0.0280)	(0.0141)	(0.0128)	(0.0254)
$age^2 * 100$	-0.2972$^\diamond$	0.2295$^\diamond$	-0.2233$^\diamond$	0.2228$^\diamond$	0.0832$^\diamond$	-0.0961
	(0.0703)	(0.0574)	(0.0597)	(0.0308)	(0.0271)	(0.0541)
$age^3 * 10000$	0.2004$^\diamond$	-0.1851$^\diamond$	0.1523$^\diamond$	-0.1712$^\diamond$	-0.0520$^\diamond$	0.0611
	(0.0487)	(0.0402)	(0.0414)	(0.0218)	(0.0187)	(0.0347)
high education	0.0164	0.3796$^\diamond$	-0.0135	0.2750$^\diamond$	0.0470$^\diamond$	0.0027
	(0.0090)	(0.0037)	(0.0077)	(0.0020)	(0.0035)	(0.0029)
no. of HH members		-0.0727$^\diamond$		-0.0417$^\diamond$		-0.0165$^\diamond$
		(0.0015)		(0.0007)		(0.0024)
non-asset income	0.3410$^\diamond$	0.0478$^\diamond$	0.3069$^\diamond$	0.0325$^\diamond$	0.0021	0.0030$^\diamond$
	(0.0053)	(0.0009)	(0.0036)	(0.0005)	(0.0021)	(0.0006)
constant	-5.6509$^\diamond$	0.2078	-4.4896$^\diamond$	1.1514$^\diamond$	1.5608$^\diamond$	-0.4361
	(0.5104)	(0.3914)	(0.4275)	(0.2079)	(0.1997)	(0.3910)
Observations	10,239	10,239	10,239	10,239	4,323	4,521
Year FE	No	Yes	No	Yes	No	Yes
ME at age 30	-0.2732$^\diamond$	-0.1334$^\diamond$	-0.1763$^\diamond$	-0.0846$^\diamond$	-0.0211$^\diamond$	-0.0340$^\diamond$
	(0.0184)	(0.0038)	(0.0154)	(0.0040)	(0.0074)	(0.0085)
ME at mean age (47)	-0.1329$^\diamond$	-0.0778$^\diamond$	-0.1027$^\diamond$	-0.0308$^\diamond$	0.0247$^\diamond$	-0.0691$^\diamond$
	(0.0088)	(0.0028)	(0.0075)	(0.0021)	(0.0034)	(0.0014)
ME at age 65	0.0156	-0.0213$^\diamond$	-0.0248	0.0239$^\diamond$	0.0732$^\diamond$	-0.1014$^\diamond$
	(0.0155)	(0.0055)	(0.0133)	(0.0041)	(0.0058)	(0.0078)

Notes: Estimations are based on linear 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. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

differences in income levels, income risk, and household sizes explain the largest fraction of the wealth gap between single men and single women. Lower income levels of women translate into lower asset holdings, explaining 28.84% of the "gender wealth gap". At the same time, the income process of single women exhibits a smaller overall variance than that of single men. Therefore, assigning single women the stochastic part of the male income process increases their precautionary savings motive. This channel in isolation explains 31.83% of the gap. Gender differences in household sizes further explain 30.44% of the gap because larger female household sizes (mainly through the presence of children) act as a consumption commitment and lower single women's ability to save.

The remaining channels are quantitatively less important for explaining gender heterogeneity in asset holdings. Assigning women the (lower) male marriage probabilities slightly reduces the gender wealth gap which arises mainly through a composition effect of more never married single women in the sample who are – on average – richer than divorcees. Assigning women the male partner's characteristics in the event of marriage (marriage market) increases their incentive to save as they now expect their prospective partners to be less wealthy and less educated at the time of marriage.

In contrast, assigning women the male medical expenses, male survival probability, and the male fraction of highly educated individuals hardly alters the gender wealth gap. Simulating the model under the assumption that both single men and single women start from the same (male) wealth level at age 30 reduces the wealth gap, in particular early in life.

Decomposing the Gap in Equity Shares. Similar to asset holdings, differences in income levels and household sizes explain the largest fraction of the gender gap in equity shares.

When single women receive the male income level (i.e. the male deterministic income component), their lifetime income increases (see Figure 4b). Hence, the simulated sample is composed of richer women who are more likely to cross the participation threshold of risky asset holdings. In addition to this compositional shift, single women are also more willing to invest in the risky asset conditional on their state variables.[21] Because of the bond-like nature of labor income, a higher human capital endowment (i.e. higher expected income in future periods) increases the willingness of single women to invest in the risky asset. Therefore, controlling for current income in reduced-form regressions is not sufficient to explain the overall effect of income on investment patterns. In addition to higher income today, portfolio choices are also affected by (expectations over) future income.

A similar mechanism applies when lowering female household sizes to that of single men. First, the sample composition shifts toward richer women because smaller household sizes decrease per-period consumption. Second, the policy functions, or decision rules, for the optimal equity share become more "risky". Assigning single women the male household sizes not only decreases their consumption today but also consumption in future periods, making them less vulnerable to financial shocks and thus increasing their willingness to invest in the risky asset. Quantitatively, assigning single women the male income level reduces the gender investment gap by 55.11%, whereas eliminating heterogeneity in household sizes narrows the gap by 71.57%.

In contrast, when single women face the same income risk as single men, the gender gap in equity shares widens by 25.67%. Because the income process of single men has a higher overall variance than that of single women (see Table 2), assigning single women the stochastic part of the male income process increases their exposure to income risk and lowers their willingness to take additional risk in the financial market.

Furthermore, assigning women the male marriage probability and altering the marriage market decreases the gender investment gap. In contrast, lowering the fraction of highly educated women to the level of single men widens the gender gap in equity shares. Less educated women have a lower income, reducing aggregate female investment in the risky asset.

[21] See Section 6.2 for a more detailed discussion on the relative importance of each mechanism over the life-cycle.

Table 5: Decomposition Results

Gap in Asset Holdings in (000s) 2007 $	Model	% explained	Data
Baseline	74.48		61.28
Male income level	53.00	28.84%	
Male income risk	50.77	31.83%	
Male HH size	51.81	30.44%	
Male marriage probability	72.76	2.30%	
Male marriage market	69.64	6.49%	
Male education distribution	73.79	0.93%	
Male medical expenses	73.14	1.80%	
Male survival probability	72.85	2.19%	
Male initial wealth	70.94	4.75%	
Gap in Equity Share in % - points	**Model**	**% explained**	**Data**
Baseline	5.99		5.93
Male income level	2.69	55.11%	
Male income risk	7.52	-25.67%	
Male HH size	1.70	71.57%	
Male marriage probability	5.46	8.78%	
Male marriage market	5.43	9.26%	
Male education distribution	6.33	-5.79%	
Male medical expenses	5.65	5.64%	
Male survival probability	5.79	3.28%	
Male initial wealth	5.53	7.60%	

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.

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:

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

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).[22] Figure 7a reports the relative importance of each effect as a fraction of the gender investment gap, whereas Figure 7b plots their absolute importance, that is how many percentage points of gender differences in equity shares can be explained by both effects.[23]

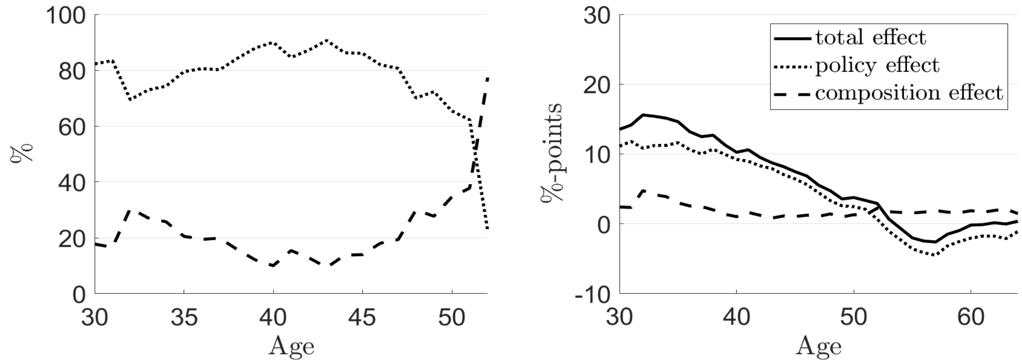
The policy effect explains the majority of gender differences in aggregate equity shares early in the life-cycle. That is, lower equity shares for relatively young single women arise from differences in policy functions rather than from differences in the sample composition. Figure 7b further reveals that the absolute gender difference in equity shares that can be explained by the composition effect is rather stable over the life-cycle. However, given that the overall gap declines as households age, the composition effect becomes relatively more important.

Conditional on current state variables, policy functions differ between single men and single women because of heterogeneous expectations over future variables, such as for example

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

[23] In Figure 7a, I drop some years prior to retirement during which the overall difference between female and male equity shares is small (as plotted in Figure 7b), which results in extremely large values when reporting the relative importance of each channel.

Figure 7: Composition vs. Policy Effect



(a) Relative Importance (b) Absolute Importance

Notes: Figure 7 decomposes the aggregate gender investment gap into a composition and a 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. Figure 7a plots the relative importance of each effect in percent. Figure 7b plots the absolute importance of each effect in percentage points, displaying also the total gender investment gap (solid line). In Figure 7a, both lines mechanically add up to 100 at every age.

income levels. These expectations are more important among young households who have more periods ahead of them to form expectations over. The older one gets, however, the fewer periods are left in which single women can form different expectations than single men, reducing the importance of the policy effect. In line with this finding, reduced form regressions that do not control for expectations fail to explain the gender investment gap, in particular among young households (see Table 4).

Beyond age 52, the policy effect even has a negative impact on the gender investment gap, meaning that the policy functions of single women are *more* risky than those of single men. However, the composition effect, i.e. gender heterogeneity in the distribution of individuals across the state space, counteracts this pattern.

Next, I further decompose the policy effect to understand the importance of gender heterogeneity in income levels, household sizes, and income risk for explaining the policy effect.[24]

[24] 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 is the part of the policy effect that can be explained by the respective channel(s) and the second term is the part that can be explained by all remaining factors.

Figure 8a plots how much of the total policy effect can be explained by gender differences in income levels and household sizes (relative to all other channels considered in Table 5), whereas Figure 8b illustrates how much of the policy effect arises from heterogeneity in income risk. In both figures, the solid lines denote the overall (or total) policy effect, the dashed lines indicate how much of it can be explained by the respective channel(s), and the dotted lines illustrate the importance of all remaining factors.

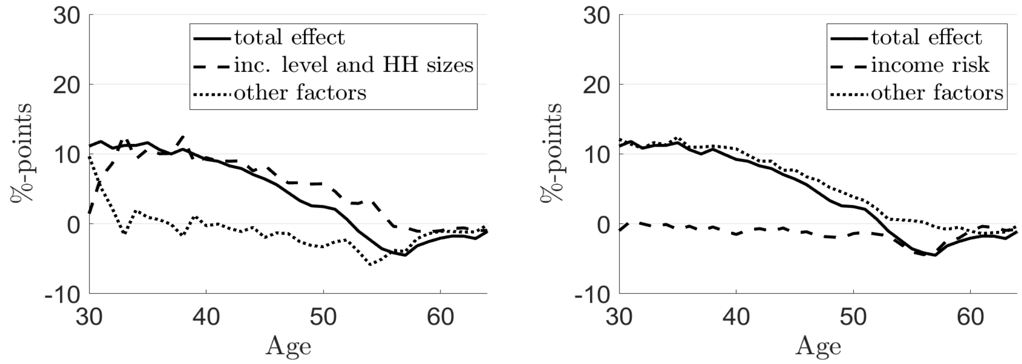
At the beginning of the life-cycle, gender differences in policy functions can be almost entirely attributed to heterogeneity in household sizes and income levels (Figure 8a). However, as individuals age, average household sizes and income levels of single women and single men converge, explaining the declining importance of both factors on the policy effect, as indicated by the negative slope of the dashed line in Figure 8a. Importantly, this pattern results in the overall policy effect (solid line) to decrease as well. Moreover, from age 45 onward, gender heterogeneity in household sizes and income levels over-explain the total policy effect, meaning that the gender variation in policy functions would be even larger if all remaining factors were identical across single men and single women.

However, as shown in Figure 8b, gender heterogeneity in income risk is one main factor that lowers the total policy effect, in particular during the second half of working life. The overall variance of the income process is lower for single women than for single men (see Table 2). As a result, the income risk channel contributes to the policy functions of women being more risky than those of men and thus, having a negative impact on the total policy effect.

6.3 Additional Evidence

In this section, I provide further evidence of gender heterogeneity in expectations about future income levels and household sizes. First, I present direct survey evidence on gender differences in expectations. Next, I interpret the empirical findings from Section 2 through the lens of the model.

Figure 8: Further Decomposition of Policy Effect



(a) Income Levels & Household Sizes (b) Income Risk

Notes: Figure 8a illustrates the importance of gender heterogeneity in income levels and household sizes for the policy effect, whereas Figure 8b illustrates the importance of income risk. In both figures, the solid lines refer to the overall policy effect (see Figure 7b), the dashed lines show the fraction of the total effect that can be explained by the respective channel(s), and the dotted line reports the importance of all remaining channels (as listed in Table 5).

Direct Survey Evidence. Neither the PSID nor the SCF contain information about individual expectations. Therefore, to empirically verify whether single men and single women indeed differ in their expectations about future income levels and the number of household members, I complement the analysis with data from the New York Fed Survey of Consumer Expectations (SCE).[25]

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. To understand whether income expectations differ by gender, 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 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. Again, this finding aligns with model predictions in that gender differences in expectations are most important early in life.

[25] See Appendix A.3 for further details about the data set and variable construction.

Table 6: Expected Earnings of Single Households – SCE Data

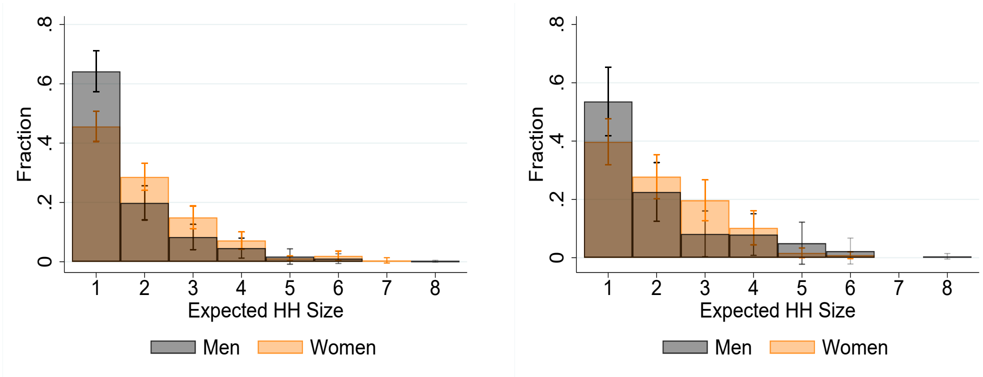
	(1) log(exp. earnings)	(2) log(exp. earnings)	(3) log(exp. earnings) Only age ≤ 45	(4) log(exp. earnings) Only age > 45
single woman	-0.1619$^{\diamond}$	-0.0922$^{\diamond}$	-0.2394$^{\diamond}$	-0.0915$^{\diamond}$
	(0.0376)	(0.0464)	(0.0548)	(0.0445)
log(current earnings)	0.6556$^{\diamond}$	0.5792$^{\diamond}$	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 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.

Furthermore, Figure 9a plots the distribution of expected household sizes for all single households whereas Figure 9b 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. Note, however, that these differences largely reflect heterogeneity in the current number of household members. Hence, Figure 9 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 expected household sizes.

Reconciling with Reduced Form Results. If expectations about future income levels

Figure 9: Expected Household Sizes of Singles – SCE Data



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

Notes: Figure 9 plots the histogram of expected household sizes (= expected number of household members) for single households. Figure 9a plots the distribution for the entire sample, whereas Figure 9b restricts the sample to households below age 40. Data is from the waves 2014 until 2019 of the Survey of Consumer Expectations (SCE).

are an important determinant of the gender investment gap, controlling for observable characteristics that proxy for these expectations should narrow observed gender differences in investment choices. In fact, including occupation and industry fixed effects (that arguably contain information about future income levels) in the reduced form regressions from Table 1 shrinks the baseline gender effect from around 11%-points to 7%-points (i.e. when comparing Columns (3) and (4)). In addition, the estimated age gradient of being a 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.

7 Robustness Checks & Discussion of Assumptions

In this section, I perform several robustness checks and discuss certain assumptions regarding the structural framework. First, I revisit the presence of child support and spousal maintenance payments. Next, I test the sensitivity of the model results with regard to potential gender heterogeneity in bequests given and bequests received. 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.[26] 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 from 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 (as women are more likely to be granted custody and because married women tend to earn less than their husbands).

Yet, I abstract from introducing these payments explicitly in the baseline model for two reasons. First, it has been shown that compliance with such laws tends to be low (Del Boca and Flinn, 1995, Case, Lin, and McLanahan, 2003). Second, introducing alimony payments and child support requires an additional state variable that keeps track of the individual's marital history. Note, however, that I account for alimony payments and child support empirically by including them in the income measure. Moreover, I test the sensitivity of my results with regard to the asset allocation upon divorce in the model. In particular, I 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 \$74,470 to \$56,260, i.e. by approximately 25%. Consequently, the gender investment gap shrinks as well by around 25% (from 5.99%-points to 4.52%-points). However, when performing the decomposition analysis of Section 6 on the model with this modified asset splitting rule, I find that the most important factors

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

contributing to the observed gaps remain gender differences in income levels, household sizes, and income risk. That is, the main results of the paper are robust to this modification.

Table 7: Robustness Checks

Gender Gap in ...	Asset Holdings (in 000s $)	Equity Shares (in %-points)
Baseline	74.48	5.99
Panel I: Alimony & Child Support		
Modified splitting rule after divorce	56.26	4.52
Panel II: Bequests		
Variation in bequest parameters	74.40	5.89
Lump-sum transfer at age 55	74.06	5.81
Panel III: Exogenous Parameters		
Higher correlation of income shocks within couples	74.02	5.83
No asset drop after death of spouse	74.35	5.95
Reduced medical expenses for couples	74.33	5.91
Higher pension payments for couples	74.13	5.84

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.

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 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 do not differ in the likelihood of ever having received an inheritance. However, conditional on having inherited something, men receive on average more.[27] Moreover, single women are less likely to expect an inheritance in the future.

Building on this evidence, I perform a robustness check in the model where I introduce a

[27] For both measures, I restrict the sample to households above 55 years because Bauluz and Meyer (2022) document that most households in the US inherit wealth when they are between 50 and 60 years old. The results remain qualitatively unchanged when considering the entire sample.

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

	Couples	Singles	
		Men	Women
Panel I: Bequests received			
% received inheritance (> 55 yrs)	27.28 (3.43)	22.97 (6.99)	22.60 (5.20)
$ amount received in (000s) (> 55 yrs)	56.16 (1.52)	35.70 (2.61)	30.43 (1.78)
% expect inheritance	17.05 (1.39)	14.88 (2.83)	10.24 (1.86)
Panel II: Bequests given			
% perceive bequests as important	51.81 (1.84)	52.23 (4.06)	53.39 (3.08)
% expect to give bequest	56.06 (1.85)	51.63 (4.02)	41.75 (3.07)

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, whereas "$ amount received in (000s) (> 55 yrs)" indicates the average amount of that inheritance. "% 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.

lump-sum transfer to all households at age 55 that varies by family type. In particular, I impose that single men receive $8,199, single women $6,876, and couples $15,322. These values reflect the empirical amount of received bequests times the probability of having received something (i.e. 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 afterwards. However, given that the received amount is quite similar for single men and single women (in particular as a fraction of their average wealth), the gender gap in both 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 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. In particular, 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 gaps in asset holdings and equity shares remain almost unchanged when compared to the baseline framework.

8 Conclusion

This paper studies the determinants of 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 hold risky assets than single men and that they allocate a smaller share of their financial wealth to risky assets. This gap remains statistically significant in reduced form regressions after controlling for a wide range of observable characteristics that have been shown to affect stock market behavior. The unexplained part of the gender investment gap is largest for young households and declines over the life-cycle.

In contrast, a life-cycle model of portfolio choice that restricts preferences to be equal across single men and single women is able to replicate the empirical gap. Counterfactual simulations reveal that higher male income levels and fewer household members (e.g. children) of single men are the main determinants for explaining the gender investment gap.

Importantly, both contemporaneous income levels and household sizes as well as their *expected path* matter for current-period investment behavior. Because of the bond-like nature of labor income, a higher human capital endowment increases an agent's optimal equity share for any given level of wealth. Similarly, lower expected household sizes reduce future consumption needs and increase financial risk-taking already in the current period. For young households, the impact of these expectations on their portfolio choices dominates gender heterogeneity in observed characteristics. Hence, in line with the empirical evidence, reduced form regressions that do not take into account households' expectations fail to explain the empirical gender investment gap, especially early in life.

References

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

Alan, S. (2006): "Entry costs and stock market participation over the life cycle," *Review of Economic Dynamics*, 9, 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," *Mimeo.*

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

Arano, K., C. Parker, and R. Terry (2010): "Gender-based risk aversion and retirement asset allocation," *Economic Inquiry*, 48, 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, 261–292.

Barro, R. J. (2009): "Rare disasters, asset prices, and welfare costs," *American Economic Review*, 99, 243–64.

Bauluz, L. and T. Meyer (2022): "The Wealth of Generations," *Mimeo.*

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

Bertrand, M. (2011): "New perspectives on gender," in *Handbook of Labor Economics*, Elsevier, vol. 4, 1543–1590.

Bonaparte, Y., R. Cooper, and G. Zhu (2012): "Consumption smoothing and portfolio rebalancing: The effects of adjustment costs," *Journal of Monetary Economics*, 59, 751–768.

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, 67–115.

——— (2023): "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labour Supply?" *The Review of Economic Studies*, 90, 102–131.

Cagetti, M. (2003): "Wealth accumulation over the life cycle and precautionary savings," *Journal of Business & Economic Statistics*, 21, 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, 171–189.

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

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

Chetty, R., L. Sandor, and A. Szeidl (2017): "The effect of housing on portfolio choice," *The Journal of Finance*, 72, 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, 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, 431–447.

Cocco, J. F. (2005): "Portfolio choice in the presence of housing," *The Review of Financial Studies*, 18, 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, 448–74.

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

Davis, S. J. and P. Willen (2014): "Occupation-level income shocks and asset returns: Their covariance and implications for portfolio choice," *The Quarterly Journal of Finance*, 3, 1350011.

De Nardi, M., E. French, and J. B. Jones (2010): "Why do the elderly save? The role of medical expenses," *Journal of Political Economy*, 118, 39–75.

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, 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, 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, 705–750.

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

Girshina, A., L. Bach, P. Sodini, and M. Team (2021): "Why Women Earn Lower Real Estate Returns," *Mimeo*.

Goldsmith-Pinkham, P. and K. Shue (2023): "The Gender Gap in Housing Returns," *The Journal of Finance, forthcoming*.

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, 869–904.

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

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

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, 937–978.

Jianakoplos, N. A. and A. Bernasek (1998): "Are women more risk averse?" *Economic Inquiry*, 36, 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*, vol. 108, 159–62.

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

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

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

——— (1971): "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3, 373–413.

Neelakantan, U. and Y. Chang (2010): "Gender differences in wealth at retirement," *American Economic Review: Papers & Proceedings*, 100, 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. (2021): "Homeownership and Portfolio Choice over the Generations," *ECB Working Paper No 2522*.

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

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, 705–728.

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

Thörnqvist, T. and A. Olafsson (2019): "Bargaining over risk: The impact of decision power on household portfolios," *Mimeo*.

Viceira, L. M. (2001): "Optimal portfolio choice for long-horizon investors with nontradable labor income," *The Journal of Finance*, 56, 433–470.

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, 197–239.

A Data Appendix

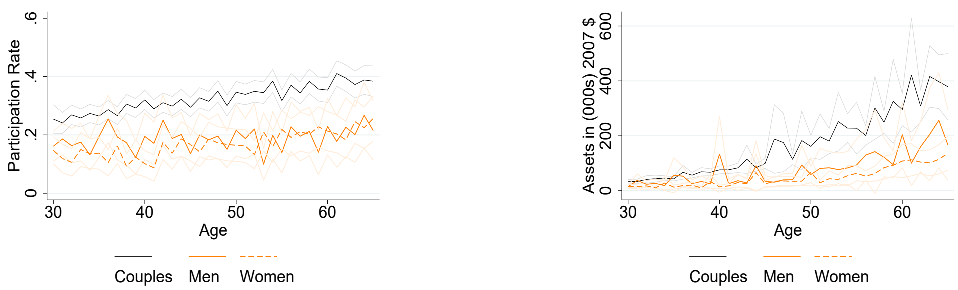
A.1 Comparability of PSID and SCF

Figure 10 plots the life-cycle profiles of stock market participation rates (excluding stocks held through retirement accounts) and financial wealth holdings by family type in the PSID, as these are measures that are available in both the PSID and the SCF. The life-cycle profiles of stock market participation rates without stocks held through retirement accounts in the PSID (Figure 10a) look very similar to those in the SCF (Figure 11b). Most importantly, I can replicate the converging gender investment gap over the life-cycle. When comparing financial asset holdings in the PSID (Figure 10b) to the SCF (Figure 1d), I find that wealth levels in the PSID are lower than in the SCF. However, couples and single men hold more financial wealth than single women in both datasets, especially as they approach retirement.

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

Figure 10: Life-Cycle Patterns of Household Finances – PSID

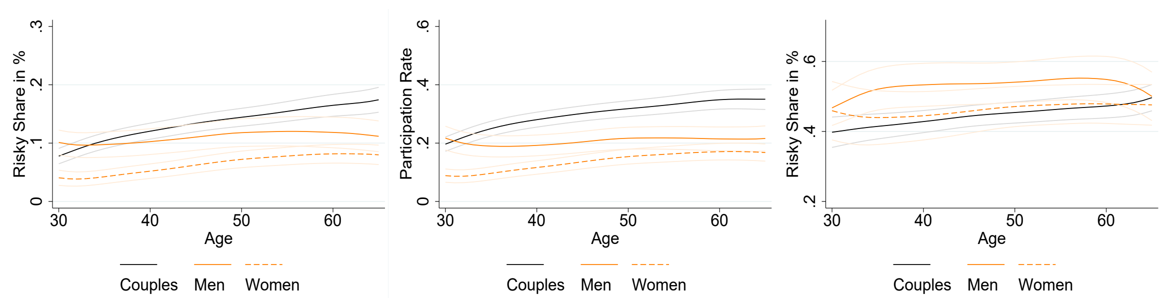


(a) Participation Rate (b) Financial Assets

Notes: Figure 10 plots the life-cycle profiles of stock market participation rates and financial wealth holdings 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. Financial wealth combines both safe and risky assets.

heterogeneity in the labor market rather than in investment choices. Figure 11 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 11a) remains statistically significant, alleviating concerns that investment differences across gender are mainly driven through savings that are linked to certain types of jobs.

Figure 11: Life-Cycle Patterns of Household Finances – Excluding Retirement Accounts

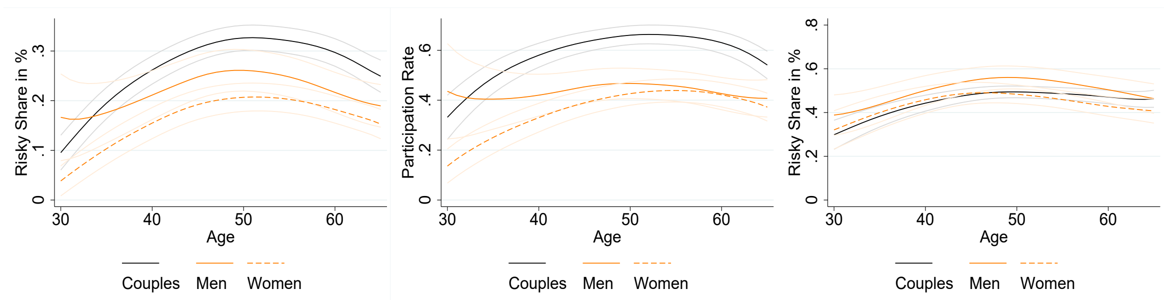


(a) Equity Share (b) Participation Rate (c) Conditional Risky Share

Notes: Figure 11 plots the life-cycle profiles of the equity share, 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. In contrast to Figure 1, financial assets do not include wealth held through retirement accounts. 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. For example, the converging gender gap in participation rates might reflect that older cohorts display a smaller gender gap in participation rates than younger ones. In fact, Ameriks and Zeldes (2004) point out that empirical patterns of portfolio allocation look very different depending on whether one controls for time or cohort effects. Therefore, Figure 12 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, but with larger standard errors due to the reduced sample size.

Figure 12: Life-Cycle Patterns of Household Finances – One Cohort



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

Notes: Figure 12 plots the life-cycle profiles of the equity share, 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.

A.3 New York Fed Survey of Consumer Expectations

To uncover expectations of households, I work with the Labor Market Survey and the 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 US 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 2019.[28]

I merge both modules to the core data which includes demographic characteristics such as

[28] I exclude the Covid period to avoid that my results are driven by responses during highly uncertain times.

gender and marital status. In the Labor Market module, respondents additionally report their current annual earnings as well as their expected annual earnings in four months, which serves as my dependent variable in Table 6. In the household finance module, respondents indicate the number of household members they expect to live with them 12 months from the time of the interview (Figure 9).

In addition, the household finance module contains information about financial wealth, which I use as a control variable in Table 6. Note, however, that households are never asked in the same month about both their expected earnings and financial wealth. Therefore, I extrapolate financial wealth to those months where households report expected earnings and only include households who answer to both modules. All variables are expressed in June 2014 Dollars.

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	-0.2868$^\diamond$	-0.1775$^\diamond$	-0.0853$^\diamond$	-0.0649$^\diamond$
	(0.0103)	(0.0096)	(0.0107)	(0.0097)
single woman * age	0.0049$^\diamond$	0.0031$^\diamond$	0.0016$^\diamond$	0.0012$^\diamond$
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
age	-0.1362$^\diamond$	-0.0926$^\diamond$	-0.0770$^\diamond$	-0.0861$^\diamond$
	(0.0140)	(0.0141)	(0.0167)	(0.0246)
$age^2 * 100$	0.3205$^\diamond$	0.2228$^\diamond$	0.1821$^\diamond$	0.1968$^\diamond$
	(0.0305)	(0.0308)	(0.0365)	(0.0543)
$age^3 * 10000$	-0.2442$^\diamond$	-0.1712$^\diamond$	-0.1414$^\diamond$	-0.1449$^\diamond$
	(0.0216)	(0.0218)	(0.0259)	(0.0387)
high education		0.2750$^\diamond$	0.1236$^\diamond$	0.0821$^\diamond$
		(0.0020)	(0.0025)	(0.0032)
no. of HH members		-0.0417$^\diamond$	-0.0259$^\diamond$	-0.0290$^\diamond$
		(0.0007)	(0.0007)	(0.0014)
non-asset income		0.0325$^\diamond$	0.0199$^\diamond$	0.0205$^\diamond$
		(0.0005)	(0.0003)	(0.0003)
safe assets			0.0550$^\diamond$	0.0543$^\diamond$
			(0.0003)	(0.0003)
constant	2.3263$^\diamond$	1.1514$^\diamond$	0.7493$^\diamond$	0.8537$^\diamond$
	(0.2062)	(0.2079)	(0.2465)	(0.3602)
Observations	10,243	10,239	10,239	7,606
Year FE	No	Yes	Yes	Yes
Industry & Occupation FE	No	No	No	Yes
ME for women at age 30	-0.1414$^\diamond$	-0.0846$^\diamond$	-0.0379$^\diamond$	-0.0297$^\diamond$
	(0.0046)	(0.0040)	(0.0042)	(0.0046)
ME for women at mean age (47)	-0.0572$^\diamond$	-0.0308$^\diamond$	-0.0105$^\diamond$	-0.0109$^\diamond$
	(0.0024)	(0.0021)	(0.0020)	(0.0024)
ME for women at age 65	0.0282$^\diamond$	0.0239$^\diamond$	0.0173$^\diamond$	0.0114$^\diamond$
	(0.0041)	(0.0041)	(0.0046)	(0.0028)

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 safe liquid 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	0.0174	0.0239	-0.0279	-0.0175
	(0.0198)	(0.0222)	(0.0185)	(0.0165)
single woman * age	-0.0020$^{\diamond}$	-0.0019$^{\diamond}$	-0.0010$^{\diamond}$	-0.0014$^{\diamond}$
	(0.0004)	(0.0005)	(0.0004)	(0.0004)
age	0.0461	0.0502$^{\diamond}$	0.0467$^{\diamond}$	0.0458$^{\diamond}$
	(0.0269)	(0.0254)	(0.0231)	(0.0189)
$age^2 * 100$	-0.0854	-0.0961	-0.0798	-0.0734
	(0.0573)	(0.0541)	(0.0491)	(0.0400)
$age^3 * 10000$	0.0524	0.0611	0.0485	0.0422
	(0.0396)	(0.0374)	(0.0339)	(0.0273)
high education		0.0027	0.0662$^{\diamond}$	0.0428$^{\diamond}$
		(0.0029)	(0.0032)	(0.0021)
no. of HH members		-0.0165$^{\diamond}$	-0.0256$^{\diamond}$	-0.0253$^{\diamond}$
		(0.0024)	(0.0018)	(0.0021)
non-asset income		0.0030$^{\diamond}$	0.0055$^{\diamond}$	0.0096$^{\diamond}$
		(0.0006)	(0.0007)	(0.0012)
safe assets			-0.0639$^{\diamond}$	-0.0687$^{\diamond}$
			(0.0009)	(0.0009)
constant	-0.2758	-0.4361	0.1380	0.0846
	(0.4085)	(0.3910)	(0.3573)	(0.2890)
Observations	4,521	4,521	4,521	3,990
Year FE	No	Yes	Yes	Yes
Industry & Occupation FE	No	No	No	Yes
ME for women at age 30	-0.0415$^{\diamond}$	-0.0340$^{\diamond}$	-0.0592$^{\diamond}$	-0.0604$^{\diamond}$
	(0.0072)	(0.0085)	(0.0065)	(0.0049)
ME for women at mean age (47)	-0.0773$^{\diamond}$	-0.0691$^{\diamond}$	-0.0782$^{\diamond}$	-0.0852$^{\diamond}$
	(0.0020)	(0.0014)	(0.0015)	(0.0022)
ME for women at age 65	-0.1101$^{\diamond}$	-0.1014$^{\diamond}$	-0.0956$^{\diamond}$	-0.1105$^{\diamond}$
	(0.0081)	(0.0078)	(0.0078)	(0.0088)

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 safe liquid assets. "ME" indicates the marginal effect of being a woman at the respective age. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

B Model Estimation – First Stage

B.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 the unexplained part of the income process (that is, the residual term ϵ_{ij} from the income equation) to follow a persistent-transitory process:

$$\tilde{y}_j = z_j + \epsilon_{\tilde{y}} \quad \text{with} \quad z_{j+1} = \rho_z z_j + \nu_z$$

A persistent-transitory process requires identification of three parameters: the persistence parameter ρ_z, the variance of the persistent shock $\sigma^2_{\epsilon_{\tilde{y}}}$ and the variance of the transitory shock $\sigma^2_{\nu_z}$ which can be identified by the following moments:

$$\frac{cov(\tilde{y}_j, \tilde{y}_{j-2})}{cov(\tilde{y}_{j-1}, \tilde{y}_{j-2})} = \frac{\rho_z^2 var(z_{j-2})}{\rho_z var(z_{j-2})} = \rho_z$$

$$var(\tilde{y}_{j-1}) - \frac{cov(\tilde{y}_j, \tilde{y}_{j-1})}{\rho_z} = var(z_{j-1}) + \sigma_{\epsilon_{\tilde{y}}} - var(z_{j-1}) = \sigma_{\epsilon_{\tilde{y}}}$$

$$var(\tilde{y}_{j-1}) - cov(\tilde{y}_j, \tilde{y}_{j-2}) - \sigma_{\epsilon_{\tilde{y}}} = \rho_z^2 var(z_{j-2}) + \sigma_{\nu_z} + \sigma_{\epsilon_{\tilde{y}}} - \rho_z^2 var(z_{j-2}) - \sigma_{\epsilon_{\tilde{y}}} = \sigma_{\nu_z}$$

I recover the parameters that minimize the distance between the covariance-variance matrices of the income process in the data and their theoretical counterparts under the assumption that $Var(z_{-1}) = 0$. In addition, the PSID collects data every two years after 1997 while the model is written in annual frequency. To account for this inconsistency, 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 in the main text displays the results.

B.2 Income Process – Deterministic Component

Table 11: Regression Coefficients for Income Estimation (Deterministic Component)

	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage
	Couples		Singles	
high educ.		0.5537$^{\diamond}$		0.4685$^{\diamond}$
		(0.0156)		(0.0236)
woman		-3.6510$^{\diamond}$		-0.7608$^{\diamond}$
		(0.0283)		(0.0223)
woman*high educ.		-0.3378$^{\diamond}$		0.1331$^{\diamond}$
		(0.0376)		(0.0292)
age	0.0884$^{\diamond}$		0.0631$^{\diamond}$	
	(0.0152)		(0.0101)	
$age^2 * 100$	-0.0965$^{\diamond}$		-0.0768$^{\diamond}$	
	(0.0161)		(0.0102)	
age*woman	0.0494$^{\diamond}$		0.0114$^{\diamond}$	
	(0.0043)		(0.0034)	
constant	7.4537$^{\diamond}$	1.5914$^{\diamond}$	9.4878$^{\diamond}$	0.1628$^{\diamond}$
	(0.3566)	(0.0120)	(0.2348)	(0.0180)
Observations	77,341	77,341	17,455	17,455
Number of unique indiv.	10,841		3,637	

Notes: Estimations are based on (fixed-effect) OLS regressions from PSID Data, waves 1989-2017. Corresponding Figure is Figure 4 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 (transfer) income is not available separately for head and spouse, I use the combined social security (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 the 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, $\diamond$ $p<0.05$.

B.3 Marriage and Divorce Probabilities

Table 12: Regression Coefficients for Marriage and Divorce Probabilities

	(1) Marriage Prob.	(2) Divorce Prob.
woman	-0.3024$^{\diamond}$	
	(0.0615)	
age	-0.0593$^{\diamond}$	-0.0252$^{\diamond}$
	(0.0033)	(0.0030)
$\mathbb{1} > 1997$	0.4675$^{\diamond}$	0.5805$^{\diamond}$
	(0.0573)	(0.0548)
high educ. (head)	0.1163	-0.5068$^{\diamond}$
	(0.0605)	(0.0617)
high educ. (spouse)		-0.1351$^{\diamond}$
		(0.0625)
constant	0.1199	-2.5911$^{\diamond}$
	(0.1377)	(0.1371)
Observations	15,287	65,419

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 the changing frequency of the PSID. Robust standard errors in parentheses, $\diamond$ $p<0.05$.

C Model Results

C.1 Solution Method & Simulation

For each set of parameters, I solve the model using backward iteration and exploit the fact that agents die with certainty in the terminal period (T) which is why I can directly solve for their optimal consumption/saving combination for each point in the state space via grid search. 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, however, 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 marital transition and labor income shocks (that, importantly, depend on each other), as well as asset return realizations, and assign each individual a certain chain of these shock processes. I simulate the model for 25,000 men and for 25,000 women who may switch marital status throughout their working life. Hence, 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 repeat the process until the minimum of the objective function is found.

I solve the model using Fortran 90 and parallelize the code with OpenMP. One solution circle for a certain set of parameters takes approximately 17 minutes when running the code on a cluster with 32 nodes.

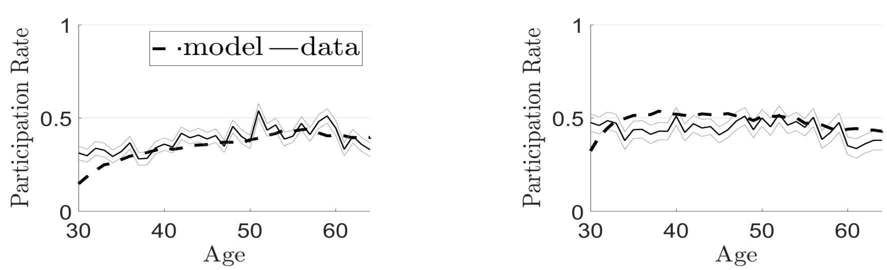
C.2 Further Results on Model Fit

This section reports further results on the model fit. Figure 13 shows that the model performs well in matching the life-cycle profiles of single households' stock market participation rates. However, in contrast to the data, it predicts the conditional risky share to be declining in age, in particular for single men (Figure 14). This difficulty of portfolio choice models to match life-cycle profiles of conditional risky shares is common: because labor income is uncorrelated to the asset return, it acts as a substitute for the safe asset. Therefore, a decreasing human-to-financial wealth ratio over the life-cycle translates into a declining optimal risky share as individuals age. Nevertheless, the model correctly matches the average levels of conditional risky shares.

Figure 15 reports the model fit for couple households, which are left entirely untargeted in the calibration exercise. While I match the equity share of couples over the life-cycle quite well (Figure 5), I slightly underpredict their participation rate and consequently overpredict their conditional risky share. Moreover, the model misses the fast asset accumulation of couples beyond age 40.

The faster empirical asset accumulation of couples 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. However, regardless of the underlying mechanism, the results of my paper do not depend on the model's ability to match the asset accumulation of couples: in a previous paper version, I improved the fit for couples by allowing their preference parameters (discount factor, coefficient of risk aversion, and stock market participation rate) to differ from those of singles, which did not alter the main results of the paper.

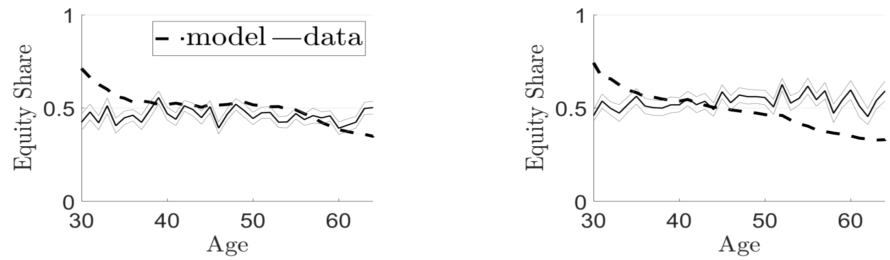
Figure 13: Model Fit of Participation Rates (Singles)



(a) Single Women (b) Single Men

Notes: Figure 13 plots the model fit of participation rates for single men and single women. The solid lines show the data (as plotted in Figure 1) whereas the dashed line display the simulated life-cycle profiles generated from the model.

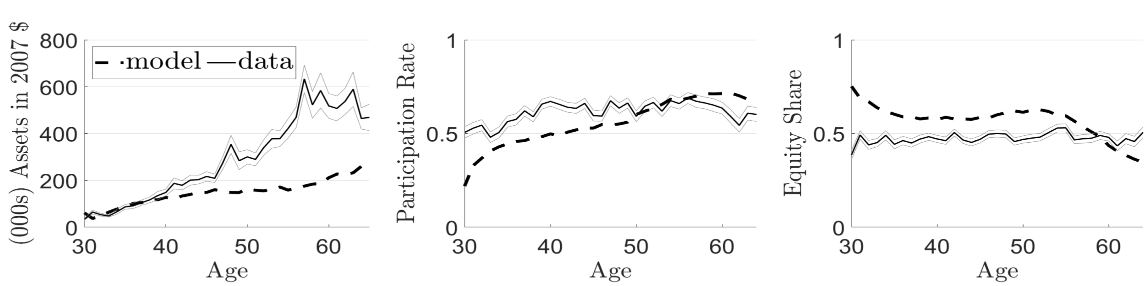
Figure 14: Model Fit of Conditional Risky Shares (Singles)



(a) Single Women (b) Single Men

Notes: Figure 14 plots the model fit of conditional risky shares for single men and single women. The solid lines show the data (including 95% confidence bands, as plotted in Figure 1) whereas the dashed line display the simulated life-cycle profiles generated from the model.

Figure 15: Model Fit of Couples



(a) Asset Accumulation (b) Participation Rates (c) Conditional Risky Share

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