The Gender Investment Gap over the Life-Cycle

Annika Bacher*

April 20, 2022

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

*European University Institute, Italy: annika.bacher@eui.eu. 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, Philipp Grübener, Lukas Nord, Jirka Slačálek, and workshop participants at the ECB, the European University Institute, the University of Wisconsin-Madison, the University of Minnesota, 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 towards 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 – to the best of my knowledge – 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, such as for example in risk aversion. 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.

towards 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 for differences in household structure (single or couple) and in gender. Individuals can get married and divorced. Single men and single women differ with regard to their income levels, their income risk, 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 preferences to be equal across single men and single women.

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 of equity shares for single men, single women and couples.

By means of counterfactual exercises, I show that heterogeneity in labor income levels and in the average number of household members (household sizes) are the most important determinants to explain the gender investment gap. Thus, the gender wage gap gets amplified in terms of wealth accumulation by making it optimal for women to invest less risky, paying on average lower returns. Importantly, not only income differences in the current period drive this result but also the fact that single men are endowed with more human capital (i.e. that they expect higher income in future periods). As shown by Cocco, Gomes, and Maenhout (2005) labor income acts as a substitute for the safe asset. Consequently, a higher human capital endowment increases households' willingness to take on financial risk.

In addition, larger female household sizes – which arise mainly through a higher likelihood of having children living in the same household – explain around half of the observed gap. Again, not only the current household size affects savings and equity shares (through different consumption needs) but larger expected household sizes act as a future consumption commitment, making single women more vulnerable to financial shocks and as a result, reduce financial risk-taking.

Lastly, I decompose the mechanisms through which heterogeneity in income levels and in household sizes translates into gender differences in equity shares 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 policy functions for equity shares conditional on the state vector, that is in households' expectations over future 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 the policy effect. Single women expect to earn less on average and to have larger household sizes than single men, making it optimal for them to invest less risky, even if they share the same (current) observable characteristics. However, as households age, single men's higher labor income and lower consumption needs translate into faster asset accumulation. Therefore, from age 55 onwards, 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 regressions that do not control for expectations have less prediction power in explaining the gender investment gap early in life when these expectations are most important, which is in line with my empirical results.

Related Literature. This paper contributes to several strands of the literature. First, it adds to a literature documenting gender differences in investment behavior and in financial choices. In general, there is large 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, Balduzzi, and Sunden (2003) show that women in the US also choose lower equity allocations in retirement saving plans. Arano, Parker, and Terry (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 stock market participation rates of women. Moreover, Goldsmith-Pinkham and Shue (2021) provide evidence that women earn lower returns on housing investments. Similarly, Andersen, Marx, Nielsen, and Vesterlund (2021) find gender differences in negotiation outcomes using transaction data of Danish residential real estate. My paper adds to this literature by being the first work – to the best of my knowledge – to analyze the gender investment gap through the lens of a structural framework.

Second, this paper relates to a mostly empirical and experimental literature documenting that women behave more risk averse than men, also with regard to financial choices (Eckel and Grossman, 2008, Croson and Gneezy, 2009, Charness and Gneezy, 2012). At first, my paper seems to contradict these findings because the model is able to replicate the gender investment gap without having to introduce gender heterogeneity in risk aversion. However, also in the current framework, single women behave observationally different to single men, conditional on state variables. The structural analysis reveals that expectations about income levels and household sizes in future periods drive this result, as opposed to underlying gender differences in risk aversion or in other structural parameters.

Third, my paper relates 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 portfolio 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 for gender differences in investment behavior over the life-cycle. Finally, Bogan and Fernandez (2017) find that having a child with mental disabilities decreases stock market participation but increases the conditional risky share.

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, more recently, Gomes (2020)). Typically, life-cycle models of portfolio choice predict the optimal equity share to be decreasing in the ratio of current financial wealth over the present value of human capital (Merton (1971), Viceira (2001)).[2] 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)

[2] The intuition behind this finding is that the correlation between labor income (human capital) and asset returns is almost zero in the data and therefore, human capital acts as a substitute for safe assets (Cocco et al., 2005).

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 a source of financial uncertainty that limits the propensity of investors to take on additional 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 decompose the gender investment along different sources of gender heterogeneity in the model. Section 7 concludes.

2 The Gender Investment Gap in the Data

The following section first explains 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 1989 until 2016 from 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-section analysis 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. To ensure the representativeness of the US population, I weigh each observation by the provided survey weight.

For income variables and demographic characteristics I work with data from the Panel Study of Income Dynamics (PSID) spanning from 1989 until 2017. The PSID is a longitudinal panel-survey of private households in the US running from 1968 until today.[3] Besides the core sample, the PSID oversamples low-income families (the 'SEO' sample) and immigrant families (the 'immigrant' sample). To make the sample comparable to that from the SCF, I drop all families belonging to those two sub-samples. All financial variables are converted into 2007 dollars using the CPI-U.[4]

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 (82,614 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 to 11,376 individuals who live in couples. The data drawn from the SCF includes information on 23,496 individuals in couples, on 4,088 single men and on 6,155 single women.

2.2 Life-Cycle Profiles of Portfolio Allocation

I define financial assets as financial wealth net of housing assets and debt (i.e. mortgages). Risky assets include direct stock holdings, corporate and foreign bonds, the fraction of mutual funds that include the former as well as the fraction of retirement accounts which is invested in stocks.[5]

Figure 1a displays the life-cycle profiles of equity shares for single men, single women and

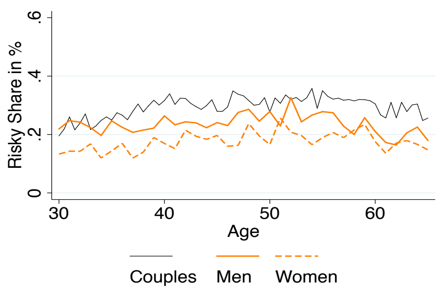
[3] Because the Survey of Consumer Finances starts in 1989, I restrict my data sample taken from the PSID to the waves from 1989 until 2017. Data were collected annually until 1997 and afterwards every two years.

[4] CPI estimates taken from the US Bureau of Labor Statistics, available under this link [Accessed May 22, 2019].

[5] In Appendix A.1, I show that my results are robust to adopting a tighter definition of risky assets that excludes risky assets held through retirement accounts.

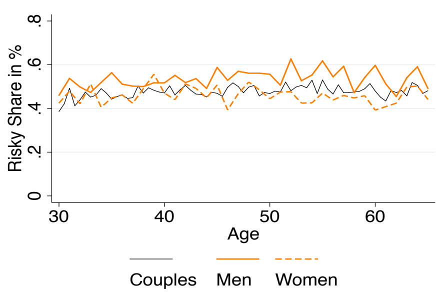
couples.[6] The equity share combines extensive margin (whether or not the households owns any risky assets) with intensive margin (conditional on holding risky asset, what portfolio share is allocated to them?). Figure 1b and Figure 1c separately plot the stock market participation rate (only the extensive margin) and the conditional risky share (only the intensive margin), respectively. All differences are statistically different from zero, see Table 1 for the corresponding regression coefficients.

Figure 1: Life Cycle Pattern of Household Finances (Data)

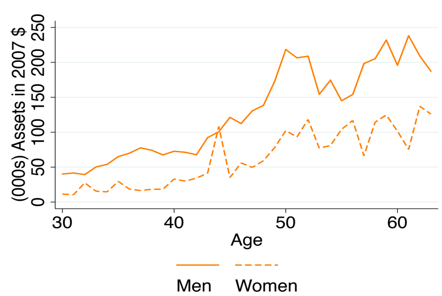


(a) Equity Share

(b) Participation Rate



(c) Conditional Risky Share



(d) Financial Assets (Singles)

Notes: Figure 1 plots the life-cycle profiles of the equity share, stock market participation rates and conditional risky shares for singles and couples as well as absolute financial assets of singles. 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 6.04%pts lower than that of men which – given an average equity share of single men of 23.79% – corresponds to 26.90% and roughly remains constant over the life-cycle. In contrast, the observed gender gap in stock market participation rates (Figure 1b) converges towards the entry to retirement.

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

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 participation 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 ones of couples in the data.

Finally, Figure 1d confirms that single women accumulate less wealth than single men also in absolute terms what is often referred to as the "gender wealth gap". Over the course of 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 linear 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 labor income, and year fixed-effects. Column (3) furthermore includes the inverse hyperbolic sine transformation of the households' safe financial assets. The corresponding marginal effects for the gender dummy along with their standard errors at various ages are reported in the last three rows of Table 1.[7]

The coefficient for being a single woman is negative (and statistically significant) across all three 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. 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' portfolio and affects stock market behavior.[8] 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.

[7] Appendix A.2 reports the corresponding specifications for the participation rate and for the conditional risky share.

[8] One 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
single woman	-0.391***	-0.230***	-0.111***
	(0.0103)	(0.00949)	(0.0110)
single woman*age	0.00573***	0.00320***	0.00122***
	(0.000207)	(0.000214)	(0.000236)
age	-0.158***	-0.0886***	-0.0630*
	(0.0269)	(0.0265)	(0.0324)
$age^2 * 100$	0.385***	0.230***	0.164**
	(0.0581)	(0.0574)	(0.0702)
$age^3 * 10000$	-0.301***	-0.185***	-0.137***
	(0.0407)	(0.0402)	(0.0493)
High education		0.380***	0.195***
		(0.00369)	(0.00419)
No. of HH members		-0.0727***	-0.0528***
		(0.00149)	(0.00211)
Income		0.0478***	0.0307***
		(0.000936)	(0.000698)
Safe assets			0.0866***
			(0.000884)
Constant	2.033***	0.208	-0.568
	(0.400)	(0.391)	(0.475)
Observations	10,243	10,239	10,239
Year FE	No	Yes	Yes
ME for women at age 30	-0.2191***	-0.1334***	-0.0746***
	(0.0048)	(0.0038)	(0.0045)
ME for women at mean age (47)	-0.1195***	-0.0778***	-0.0533***
	(0.0031)	(0.0028)	(0.0027)
ME for women at age 65	-0.0184***	-0.0213***	-0.00317***
	(0.0047)	(0.0055)	(0.0053)

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, *** p<0.01, ** p<0.05, * p<0.1

To explore whether either of these issues are present in the data, I conduct two exercises: First, if portfolio choices are a direct mapping of housing decisions, I would expect their life-cycle profiles to closely follow those in Figures 1a to 1c. Figure 2 displays the life-cycle profiles of singles for three different housing variables: the homeownership rate, housing wealth (henceforth: "HW") and the housing-wealth-to-income ratio (henceforth: "HI"). Single men hold on average slightly more housing wealth than single women. However, I do not find any gender differences in terms of homeownership rates or in the housing-wealth-to-income ratio despite significant gender differences in 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 Pattern (Singles)

(a) Homeownership Rate

(b) (Gross) Housing Wealth

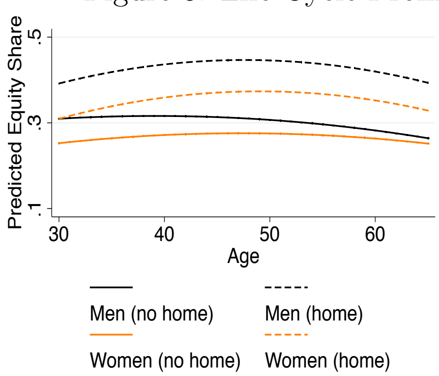
(c) Housing-Wealth-to-Income Ratio

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

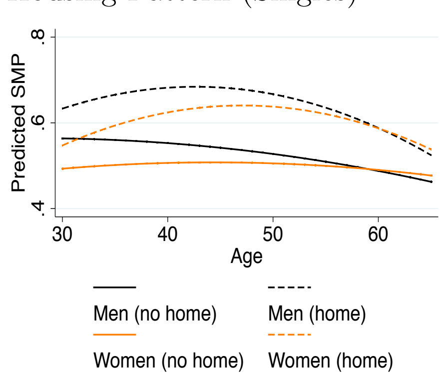
Second, I compare the predicted equity share and stock market participation rate of single homeowners to those of single non-homeowners, separately by gender. This exercise helps me to understand whether housing differently affects financial investment patterns of single men and single women. To do so, I split the sample by housing tenure and run two separate regressions on stock market participation and equity shares, respectively, controlling for observable characteristics. Figure 3 plots the predicted outcome variable from these regressions for an individual with more than 12 years of schooling, the median number of children for singles (zero), the average income and average amount of safe assets, and who is at the respective age in 2001 (which is approximately the midpoint of my sample).

In line with previous literature, the predicted equity share and participation rate differ across homeowners and renters. However, the *gender differences* in predicted participation rates and equity shares (i.e. the gap between the black and the orange line) are very similar for homeowners and non-homeowners, suggesting that housing does not explain heterogeneity in financial investment behavior across single men and single women. If anything, the predicted gender gap in equity shares is slightly larger among home-owners towards the end of the life-cycle. In contrast, the gender gap in participation rates and in equity shares among young household does not significantly differ by homeownership-status, reassuring that housing does not affect portfolio choices of single men different than those of single women.

Figure 3: Life Cycle Profiles of Housing Pattern (Singles)



(a) Equity Share



(b) Participation Rate

Notes: Figure 3 plots the predicted life-cycle profiles of the equity share and the stock market participation rates of a single individual in 2001 who has a high school degree, no children and the average level of income and safe assets, separately by gender and housing tenure. 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 ($\mathcal{M}$). For all, their life can be split in two stages: working age and retirement. Time is discrete and 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 college and non-college educated individuals 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 ($c_{\mathcal{M}}$) as well as on 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 level. 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 single or couple, 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 couples' assets to account for increased medical expenditures in the year prior to death as well as for bequests to non-spousal heirs (Jones, De Nardi, French, McGee, and Rodgers, 2020). As before, households have a portfolio choice between a safe asset and a risky asset.

3.1 Preferences

Singles. Single individuals have time-separable CRRA preferences over a consumption good c_i. Their period flow of utility is given by:

$$u(c_i) = \frac{\eta_{ij}\left(\frac{c_i}{\eta_{ij}}\right)^{1-\gamma_s}}{1-\gamma_s}$$

where γ_s is the coefficient of relative risk aversion and η is an equivalence scale that adjusts for household size and which is allowed to vary by age j and gender i.

Couples. Each couple is composed of exactly one woman and one man. As for singles, couples have time-separable CRRA preferences over the consumption good $c_{\mathcal{M}}$ which is public within the household. Their period flow of utility can therefore be expressed as:

$$u(c_{\mathcal{M}}) = \frac{\eta_{cj}\left(\frac{c_{\mathcal{M}}}{\eta_{cj}}\right)^{1-\gamma_c}}{1-\gamma_c}$$

Again, γ_c is the coefficient of relative risk aversion and η is an age-dependent equivalence term adjusting for household size.

While I do not allow for gender heterogeneity in preferences, I introduce preference heterogeneity by marital status (e.g. in the coefficient of relative risk aversion γ). This adjustment helps to accommodate the data and allows me to match empirical equity shares and wealth profile of couple households (which, in turn, affects singles' expectations, who are the focus of my study) but at the same time keeping the model tractable.

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 be-quested assets, ω captures the luxuriousness of the bequest motive and L governs the bequest intensity. Bequest preferences are homogeneous across all types of households. 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 bequest 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 it holds that $\mu_r > r_s$. 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 $r_{tail} < \underline{\mathrm{r}}_r$ materializes. Short-selling and borrowing are not allowed.

Income Profiles. Income y_{ij} at age j for gender i can be split into a deterministic and into 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 consisting of a transitory and a persistent shock:

$$\tilde{y}_{ij} = z_{ij} + \epsilon_{\tilde{y}ij}$$

$$z_{i,j+1} = \rho_{zi} z_{ij} + \nu_{zij}$$

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}.

All parameters of the income process are allowed to vary by gender and by marital status to allow for the fact that marriage typically results in lower income for women whereas it increases earnings for men.[9] Within couples, the transitory shocks ν_{zf} and ν_{zm} are assumed to be correlated (with $\rho_{\sigma_{zf},\sigma_{zm}} = 0.3$) as spouses live in the same area and are likely to work in similar industries and are thus subject to correlated labor market shocks. In contrast, following Cocco et al. (2005), labor income shocks are uncorrelated to realizations of the stock return.

Out-of-Pocket Medical Expenditures. When being retired, agents are subject to medical expenditures m_j 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. This cost is allowed to differ between couples and singles, however, it is equal for single men and single women. As in Vissing-Jorgensen (2002), participation costs have to be paid each period irrespectively of the history of stock holdings. The main advantage to model 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 an

[9] The empirical observed "marriage penalty" for women's earnings can arise because of self-selection into marriage or because marriage itself affects income. Taking a stance on the exact mechanism is beyond the scope of this paper and therefore, I impose women's earnings drop upon marriage exogenously.

additional state variable.

3.4 Marriage and Divorce

Single individuals get married with an exogenous probability that depends on their gender, their age, their education and on their current productivity realization. Denote this marriage probability by $\mu_i(j, \theta, \tilde{y})$. 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, the education of each spouse as well as both spouses' productivity realization $\lambda(j, \theta_f, \theta_m, \tilde{y}_f, \tilde{y}_m)$. Upon divorce, assets are split equally between spouses and 25% of assets are destroyed to account for legal fees of divorce and general costs of splitting assets between spouses. There are no alimony payments.

3.5 Timing

In 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 to pay S^F in the current period t.

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:

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

subject to:

$$a'_r + a'_s + c = y(j,\theta,\tilde{y}) + (1+r_s)a_s + (1+r_r)a_r - \mathbb{1}_{a'_r > 0} S^F_s$$

$$a = (1+r_r)a_r + (1+r_s)a_s$$

and:

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

$$r_r \sim N(\mu_r,\sigma^2_r) \quad \text{with} \quad \mu_r > r_s \quad \text{and} \quad \tilde{y} \perp r_r, \quad \mathbb{E}(\tilde{y}', r'_r, \Pi | j,\theta,\tilde{y})$$

where η_j 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 realization, 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 ($\hat{y}$).

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

subject to:

$$a'_r + a'_s + c = pen(\hat{y}) + (1+r_s)a_s + (1+r_r)a_r - m_{ij} - \mathbb{1}_{a'_r > 0} S_s^F$$

$$a = (1+r_r)a_r + (1+r_s)a_s$$

$$r_r \sim N(\mu_r, \sigma_r^2) \quad \text{with} \quad \mu_r > r_s, \quad \text{and} \quad \mathbb{E}(r'_r)$$

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 value function for couples during working age is needed to compute optimal allocation for a couple that consists of a women f and a man m. The state variables of a couple 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:

$$V^C(j,\theta_f,\theta_m,a,\tilde{y}_f,\tilde{y}_m) = \max_{a'_s\geq 0, a'_r\geq 0, c\geq 0} \frac{\eta_{\mathcal{M}j}\left(\frac{c}{\eta_{\mathcal{M}j}}\right)^{1-\gamma_c}}{1-\gamma_c} +$$

$$(1-\lambda(j,\theta_f,\theta_m,\tilde{y}_f,\tilde{y}_m))\beta_c \mathbb{E}V^C(j+1,\theta_f,\theta_m,a',\tilde{y}'_f,\tilde{y}'_m)+$$

$$\lambda(j,\theta_f,\theta_m,\tilde{y}_f,\tilde{y}_m)\beta_c \sum_{i=f,m}\mathbb{E}V^S(i,j+1,\theta_i,0.75\frac{a'}{2},\tilde{y}'_i)$$

subject to:

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

$$a = (1+r_r)a_r + (1+r_s)a_s$$

and:

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

$$\begin{pmatrix} \nu_{zf} \\ \nu_{zm} \end{pmatrix} \sim \begin{pmatrix} \sigma^2_{zf} & \rho_{\sigma_{zf},\sigma_{zm}} \\ \rho_{\sigma_{zf},\sigma_{zm}} & \sigma^2_{zm} \end{pmatrix}$$

$$r_r \sim N(\mu_r,\sigma^2_r) \quad \text{with} \quad \mu_r > r_s, \quad \tilde{y} \perp r_r \quad \text{and} \quad \mathbb{E}(\tilde{y}'_f,\tilde{y}'_m,r'_r|j,\theta_f,\theta_m,\tilde{y}_f,\tilde{y}_m)$$

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 return when getting divorced.

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

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

subject to:

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

$$r_r \sim N(\mu_r,\sigma_r^2) \quad \text{with} \quad \mu_r > r_s, \quad \text{and} \quad \mathbb{E}(r'_r)$$

Thus, retired couples take the expected value over their joint asset return as well as 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.

$$\hat{V}^C(i,j,\theta_i,\theta_p,a,\tilde{y}_i,\tilde{y}_p) = \frac{\eta_{\mathcal{M}j}\left(\frac{c}{\eta_{\mathcal{M}j}}\right)^{1-\gamma_c}}{1-\gamma_c} + (1-\lambda(j,\theta_i,\theta_p,\tilde{y}_i,\tilde{y}_p))\beta_c\mathbb{E}\hat{V}^C(i,j+1,\theta_i,\theta_p,a',\tilde{y}'_i,\tilde{y}'_p) +$$
$$\lambda(j,\theta_i,\theta_p,\tilde{y}_i,\tilde{y}_p)\beta_c\mathbb{E}V^S(i,j+1,\theta_i,0.75\frac{a'}{2},\tilde{y}'_i)$$

4 Estimation & Calibration

I estimate 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 a 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_s^F, all separately for singles and couples. I collect the second stage parameters in the vector $\Theta = \{\beta_s, \beta_c, \gamma_s, \gamma_c, S_s^F, S_c^F\}$.

4.1 First Stage Estimation

Income Profiles. Figure 4 shows life-cycle profiles of average income 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. For singles, I include labor earnings, social security benefits and transfers from all members of the households to ensure that my measure of income adequately accounts for disposable income of single households in the data. For couples, I assign each spouse their own labor income, social security benefits and transfers and add half of that from other household members.[10] Lastly, I drop observations who, according to the described measure, report zero annual income (in the case of couples, if they report zero overall income).

[10] For some years, the PSID does not separately report transfer income or social security benefits of 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.

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

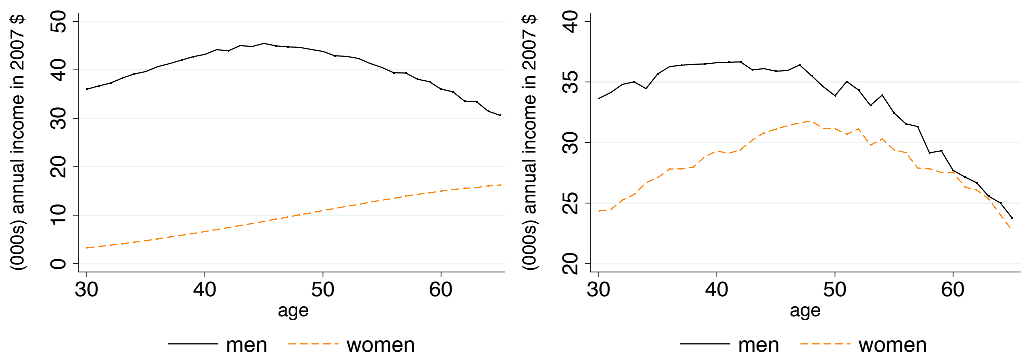
$$income_{ij} = \alpha + \beta_1 age_{ij} + \beta_2 age_{ij}^2 + \beta_3 woman_i * age_{ij} + \delta_i + u_{ij}$$

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:

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

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 Income (Deterministic Component)



(a) Couples (b) Singles

Notes: Figure 4 plots the life-cycle profiles for the deterministic part of labor income by gender and by 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 8 in Appendix B.2) inform me about the deterministic component of the income process in the model. I estimate the parameters governing the stochastic component of the income process using the minimum distance estimator as in Guvenen (2009).[11] Table 2 summarizes the results. My point

[11] Details on the estimation strategy can be found in Appendix B.1. Moreover, when estimating the stochastic

estimates imply a slightly less persistent income process for single women than for single men, 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 income process of married women exhibits a much higher variance of the transitory shock $\sigma_{\tilde{y}}^2$ than that of singles and that of 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. Figure 5 plots marriage and dviorce probabilities by age, gender and education. 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, separately for couples and singles:

$$\xi_{t+1} = \frac{exp(X_t\beta^s)}{1 + exp(X_t\beta^s)}$$

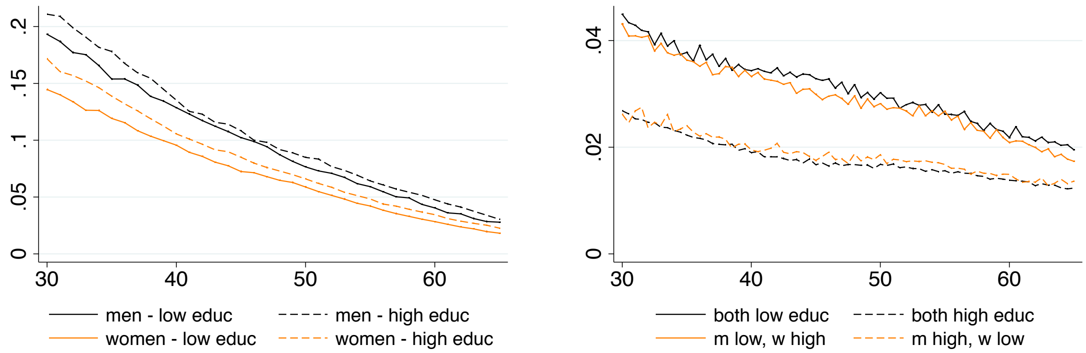
where ξ_{t+1} denotes the probability of being married (respectively divorced) next period. As explanatory variables, I include the age, age-squared, a dummy indicating whether the individual has some college education and a dummy for waves after 1997 to account for switch from annual to biannual frequency in the PSID.[12] Table 9 in Appendix B.3 reports

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.

[12] For couples, all demographic variables refer to the household head.

the corresponding regression coefficients.

Figure 5: Marital Transition Probabilities



(a) Marriage Probability

(b) Divorce Probability

Notes: Figure 5 plots the life-cycle profiles for marriage and divorce probability, respectively, separately by age and education. Data is from the waves 1989 until 2016 of the Panel Study of Income Dynamics (PSID).

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. Moreover, couples in which the husband has low education and the wife has high education face a higher chance of divorce than couples whose education is allocated in the opposite way.

Finally, I estimate the marriage market Π, i.e. the matching of spouses in terms of income, assets, and eduction, 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. Moreover, to account for the possibility of informal care arrangement among spouses, I assume that medical expenses for married individuals are 80% than that of singles.

Survival Probabilities. I take gender specific death probabilities from the Life Tables of the US Social Security Administration.[13] 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 70% of the household's asset to account for sharply increasing medical expenses in the year prior to death as well as for bequests to non-spousal heirs (Jones et al., 2020).

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 reflect 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 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 which 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 by 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 member 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,

[13] All tables available under this link [Accessed May 14, 2019].

I set the fraction of high and low educated individuals by gender to be the average share of individuals with more respectively 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.

4.2 Second Stage

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

$$\mathcal{L} = \min_{\Theta} \left(\frac{M^s(\Theta) - M^d}{M^d} \right) W \left(\frac{M^s(\Theta) - M^d}{M^d} \right)' \tag{1}$$

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 target to match life-cycle profiles in participation rates. To identify parameters referring to couples $(\beta_c, \gamma_c, S_c^F)$, I target life-cycle profiles of couples. For parameters that govern the choices of singles $(\beta_s, \gamma_s, S_s^F)$, I target the corresponding life-cycle profiles of single men. Consequently, life-cycle profiles of single women serve as untargeted moments to validate the model. In total, I work with 216 moments (36 years $\times$ 3 variables $\times$ 2 HH types).

The Weighting Matrix W. I first calibrate the 2nd stage parameters by using the identity matrix ($W = \mathcal{I}$). Consequently, every moment receives equal weight. 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 the 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. My results imply that singles discount the future less (smaller β) and have a lower degree of risk aversion (smaller γ) than couples. The stock market participation cost is larger for couples than for singles. The higher discount factor of couples is necessary to replicate their higher (per capita) asset holdings, despite them having correlated labor income shocks and hence, less need to accumulate precautionary savings against income shocks. The stronger savings pattern can arise from various sources that are outside the scope of this paper, e.g. savings for children's college. In contrast, a lower coefficient of risk aversion for singles is needed to match their lower conditional risky shares for a given level of wealth. In addition, higher wealth levels of couples result in more couple households to cross the participating threshold for risky asset holdings, requiring larger participation costs to match empirical stock market participation rates.

The parameter values across the two specifications are relatively similar except that the calibration which works with the identity matrix finds a lower discount factor and a higher coefficient of risk aversion for singles. When using the inverse variance matrix, asset moments receive the least weight. Hence, deviations from the asset profile of single men have less impact on the objective function, resulting in a larger β_s. As a consequence, a lower γ_s is required to match the empirical conditional risky share of single men.

The calibration that uses the identity matrix to weigh its moments finds an annual stock market participation cost of \$1097.5 for couples and of \$713.5 for singles. With regard to the coefficient of risk aversion, my estimates suggest $\gamma_s = 3.9527$ for singles and $\gamma_c = 4.8236$ for couples. The values for the coefficient of risk aversion are 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 quite a high degree of risk aversion with $\gamma = 11$. 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.[14] The values for β (0.8120 for couples and 0.7195 for singles) are in line with previous literature. Cooper and Zhu (2016) estimate a discount factor of 0.869, Fagereng et al. (2017) of 0.77 and Catherine (2022) of 0.88. However, given that my coefficients for the relative risk aversion are well below their estimates, my values for β are comparably low.

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 increases the precautionary savings motives for couples while at the same time generating high-asset single households (who got divorced) that are absent in models with only one generic household

[14] Catherine (2022) addresses the trade-off 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.

type. Generally, marriage and divorce introduce an additional dimension of financial risk that lowers households' willingness to invest in the risky asset. Therefore, my model is able to match empirical equity shares of households with relatively low degrees of risk aversion.

Table 3: 2nd Stage Parameters

β_c	β_s	γ_c	γ_s	S_c^F	S_s^F	$\mathcal{L}$	W
0.8120	0.7195	4.8236	3.9527	\$1097.5	\$713.5	26.7546	$\mathcal{I}$
0.7956	0.7563	4.8807	3.5242	\$1035.5	\$825.5	56.3326	$\frac{1}{\mathcal{V}}$

5.2 Model Fit

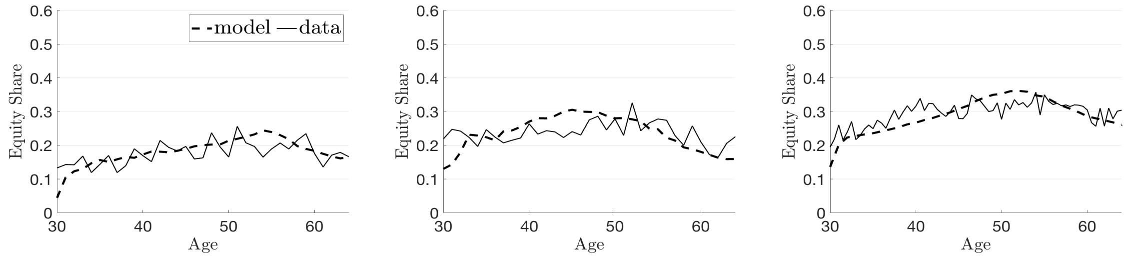
Figures 6 and 7 contrast the life-cycle profiles of equity shares and asset accumulation from the data with those generated by the model.

The model matches very well the life-cycle profiles of equity shares for single men, single women, and couples. Hence, 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 7) for couples , single men, and single women. For all household types, the asset accumulation profiles is 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 for single men and single women. Appendix C.2 furhter reports the model fit for the participation rate and the conditional risky share by household type.

5.3 Simulated Regressions

To compare reduced form regressions from Section 2 with those generated by the model, Table 4 replicates the same regressions on generated data from model simulations. Columns

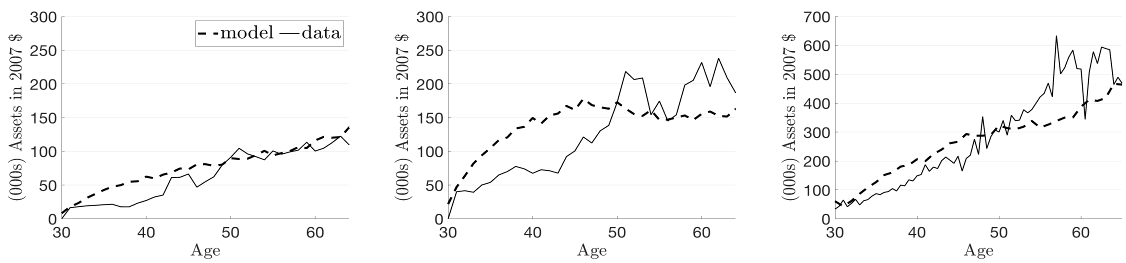
Figure 6: Model Fit of Equity Shares



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

Notes: Figure 6 plots the model fit of equity shares for single women, single men and couples. 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 7: Model Fit of Asset Accumulation



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

Notes: Figure 7 plots the model fit of asset accumulation for single women, single men and couples. 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.

(1) - (3) show the regression estimates from model simulations whereas Columns (4) - (6) re-report the results with data from the Survey of Consumer Finances. All of these coefficients are untargeted in the calibration exercise.

The model slightly over-predicts the effect of gender on portfolio choices, meaning that the absolute values for the coefficient "single woman" and for the interaction term of single woman and age are larger than in the data. However, the simulated data captures well the increasing marginal effect ("ME") of being a woman on the equity share over the life-cycle. Both in Column (1) and in Column (4), the marginal effect is smallest (most negative) at young ages, and increases as households age. Hence, reduced form regressions that control for observable household characteristics fail to fully explain the gender investment gap, in particular among young households, even if the underlying data generating process assumes preferences homogeneity across men and women.

Thus, it appears that either factors which cannot as easily be controlled for (such as expectations) or non-linearities explain the residual part of the gap. To uncover these factors to quantify their relative importance over the life-cycle, Section 6 performs several counterfactual exercises.

6 Counterfactual Simulations

6.1 Decomposing the Gender Investment Gap

In this section, I decompose the gender gap in equity shares and in wealth levels along the dimensions of gender heterogeneity within the model, that is along income levels, income risk, 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 in 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 differences in income levels, income risk and in household size explain the largest fraction of the wealth gap between single men and single women. Lower income levels of women translate into fewer asset holdings, explaining 23.20% of the "gender wealth gap". At the same time, the income process of single women is less volatile than that of single men. Therefore, assigning single women the stochastic part of the male income process increases

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

	(1) Equity Share	(2) SMP	(3) Cond. Share	(4) Equity Share	(5) SMP	(6) Cond. Share
	Model Simulations			Data (SCF)		
single woman	-0.503***	-0.277***	0.192***	-0.230***	-0.178***	0.0239
	(0.0545)	(0.0361)	(0.0251)	(0.00949)	(0.00955)	(0.0222)
single woman*age	0.00761***	0.00376***	-0.00254***	0.00320***	0.00310***	-0.00193***
	(0.00108)	(0.000724)	(0.000489)	(0.000214)	(0.000196)	(0.000460)
age	0.232***	0.127***	-0.0706***	-0.0886***	-0.0926***	0.0502**
	(0.0443)	(0.0300)	(0.0204)	(0.0265)	(0.0141)	(0.0254)
$age^2 * 100$	-0.404***	-0.224***	0.169***	0.230***	0.223***	-0.0961*
	(0.0945)	(0.0646)	(0.0431)	(0.0574)	(0.0308)	(0.0541)
$age^3 * 10000$	0.224***	0.131***	-0.137***	-0.185***	-0.171***	0.0485
	(0.0658)	(0.0453)	(0.0298)	(0.0402)	(0.0218)	(0.0339)
High education	-0.0489***	-0.0570***	0.0388***	0.380***	0.275***	0.00267
	(0.0108)	(0.00770)	(0.00475)	(0.00369)	(0.00201)	(0.00288)
No. of HH member				-0.0727***	-0.0417***	-0.0165***
				(0.00149)	(0.000731)	(0.00243)
Income	0.412***	0.298***	0.0188***	0.0478***	0.0325***	0.00296***
	(0.00651)	(0.00360)	(0.00284)	(0.000936)	(0.000453)	(0.000628)
Constant	-8.565***	-5.004***	1.315***	0.208	1.151***	-0.436
	(0.683)	(0.455)	(0.317)	(0.391)	(0.208)	(0.391)
Observations	10,239	10,239	4,011	10,239	10,239	4,521
Year FE	No	No	No	Yes	Yes	Yes
ME at age 30	-0.2745***	-0.1638***	0.1160***	-0.1334***	-0.0846***	-0.0340***
	(0.0236)	(0.0155)	(0.0110)	(0.0038)	(0.0040)	(0.0085)
ME at average age	-0.1452***	-0.0999***	0.0703***	-0.0778***	-0.0308***	-0.0691***
	(0.0106)	(0.0074)	(0.0047)	(0.0028)	(0.0021)	(0.0013)
ME at age 65	-0.0082	-0.0322**	0.0272**	-0.0213***	0.0239***	-0.1014***
	(0.0194)	(0.0138)	(0.0084)	(0.0055)	(0.0041)	(0.0078)

Notes: Estimations are based on linear regressions on the sample of individuals who live in households with no spouse present. Columns (1)-(3) are model simulations, columns (4)-(6) refer to data from the SCF waves 1989 until 2016. Equity Share = Unconditional risky share. SMP = Stock market participation rate. Cond. 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, *** p<0.01, ** p<0.05, * p<0.1

their precautionary savings and reduces the gender wealth gap. This channel in isolation explains 24.91% of the gap. Gender differences in household size further explain 20.11% 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 high educated individuals increases the gender wealth gap. Longer female life expectancies and higher medical expenses increase their need to save for retirement. However, as most gender differences in survival risk and in medical expenditures materialize at the very end of the life-cycle, this effect is quantitatively small when averaging over working life. In addition, the fraction of highly educated individuals among single women is higher than among single men. Hence, assigning women the male education distribution lowers the share of highly educated single women and consequently widens the gender wealth gap.

Finally, when 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 early in life, increasing consumption of young single women.

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

When single women receive the male income level, 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 wealth level (and other state variables).[15] Because of the bond-like nature of labor income, a higher human capital endowment (i.e. more expected income in future periods) increases the willingness of single women to invest in the risky asset for any given level of wealth and income. Hence, 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, current portfolio choices are also affected by *expectations* over future income realizations. It is rational for a single woman who has the same current income, wealth and education level as a single man to invest less risky because she expects a lower income in the future.

The same mechanism applies when lowering female household sizes to that of single men. First, the sample composition shifts towards richer women because smaller household sizes decrease per-period consumption. Second, the policy function for the optimal equity share becomes more "risky". Assigning single women the male household size not only decreases their consumption needs today but also that 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 gap in equity shares by 59.07%, whereas eliminating heterogeneity in household sizes narrows the gap by 63.17%.

In contrast, when single women face the same income risk as single men, the gender gap in equity share widens by 13.34%, despite single women being on aggregate wealthier. Because the income process of single men is more volatile than the income process of single women (see Table 2), giving single women the male income process increases their exposure to labor

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

market risk and lowers their willingness to to take additional risk in the financial market.

Lastly, lowering the fraction of highly educated women to the same levels as for single men widens the gender investment gap by 9.25%. Less educated women have a lower income, reducing aggregate female investment in the risky asset.

Table 5: Decomposition Results

Gap in Asset Holdings in (000s) 2007 $	Model	% explained	Data
Baseline	62.34		61.28
Male income level	47.88	23.20%	
Male income risk	46.81	24.91%	
Male HH size	49.80	20.11%	
Male marriage probability	60.47	3.00%	
Male marriage market	58.72	5.81%	
Male education distribution	64.09	-2.81%	
Male medical expenses	62.52	-0.29%	
Male survival probability	62.63	-0.47%	
Male initial wealth	60.80	2.57%	
Gap in Equity Share in %	**Model**	**% explained**	**Data**
Baseline	5.62%		5.93%
Male income level	2.30%	59.07%	
Male income risk	6.37%	-13.34%	
Male HH size	2.07%	63.17%	
Male marriage probability	5.65%	-0.53%	
Male marriage market	5.55%	1.24%	
Male education distribution	6.14%	-9.25%	
Male medical expenses	5.56%	1.06%	
Male survival probability	5.68%	-1.07%	
Male initial wealth	5.35%	4.80%	

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

Aggregate portfolio allocations in the model are determined by the policy function for the optimal risky share $\alpha = \phi(X)$, conditional on state variables X, and the distribution of individuals across the state space. Thus, differences in aggregate investment patterns between the baseline and a counterfactual scenario can arise because the distribution of individuals across the state space changes ("composition effect") or because individual policy functions at any given point in the state space differ ("policy effect").

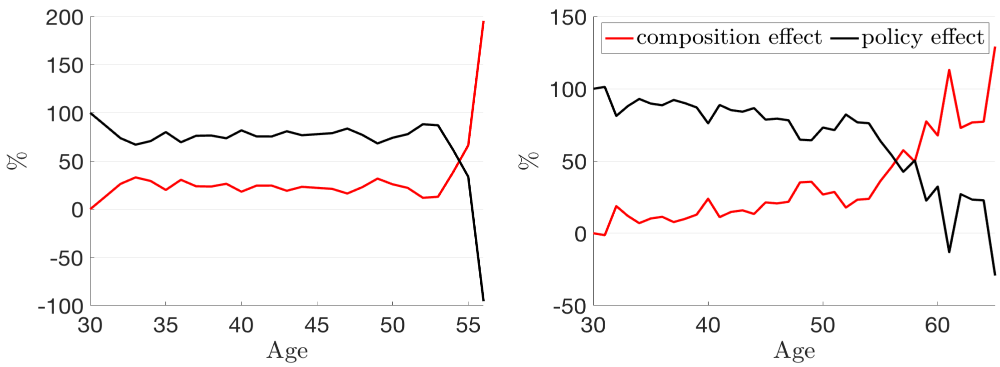
The objective of this section is to quantify the relative importance of the composition effect vs. the policy effect on the gender investment gap along the life-cycle. To do so, I re-simulate the income and the household size counterfactual but restrict policy functions for the equity share to be the same as in the baseline model. Consequently, any difference in life-cycle profiles between this simulation and the counterfactual with unrestricted policy functions can be attributed to the policy effect, that is, because single women choose a more risky portfolio allocation conditional on their state vector (through different expectations). Figure 8 reports the results of this exercise.[16]

In the household size counterfactual (Figure 8b) the policy effect explains the majority of differences in aggregate equity shares between the baseline and the counterfactual early in the life-cycle. That is, higher equity shares for relatively young single women arise from differences in the policy functions through expectations rather than from differences in the sample composition. Young single women are more likely than single men to have children living with them in the future what increases their expected consumption needs and reduces their willingness to take on financial risk already today. As individuals age, the composition effect becomes more important. The older one gets, the fewer periods are remaining in which single women expect to have larger household sizes. In addition, children move out, reducing gender differences in the number of household members. However, larger household sizes in previous years have translated into less savings, affecting the sample composition. In line with this finding, reduced form regressions (see Table 4) that do not control for expectations fail to explain the gender investment gap, in particular among young households.

[16] In Figure 8, I drop some years prior to retirement during which the overall difference between the baseline model and the unrestricted counterfactual is small. Additionally, in these years, the composition effect over-explains the gap between the baseline and the unrestricted counterfactual, which results in extremely large values when reporting the relative importance of each channel, despite the overall effect being small. Figure 13 in Appendix C.3 further illustrates that point.

A similar pattern emerges for the income counterfactual (Figure 8a). When young, single women are endowed with less human capital (i.e. they expect to have a lower labor income) than single men what reduces their willingness to take financial risk. In contrast, the older they get, lower income levels in the past have translated into less savings, shifting the sample composition towards asset-poorer single women households.

Figure 8: Composition vs. Policy Effect (Single Women)



(a) Income Counterfactual (b) HH Size Counterfactual

Notes: Figure 8 decomposes the difference in the aggregate female equity shares between the baseline model and the income resp. household size counterfactual into a composition and into a policy effect. The composition effect (red line) shows which percentage of the overall gap can be explained through on aggregate richer individuals in the simulated sample whereas the policy effect (black line) shows which percentage of the gap can be explained by differences in policy functions. Both lines mechanically add up to 100 at every age.

7 Conclusion

This paper studies the determinants of the gender investment 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 into 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 men and women is able to replicate the 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," *Working Paper, Columbia University*.

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.

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

Bogan, V. L. and J. M. Fernandez (2017): "How children with mental disabilities affect household investment decisions," *American Economic Review*, 107, 536–40.

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.

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

Catherine, S. (2022): "Countercyclical Labor Income Risk and Portfolio Choices over the Life Cycle," *The Review of Financial Studies, forthcoming.*

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.

Cocco, J. F., F. J. Gomes, and P. J. Maenhout (2005): "Consumption and portfolio choice over the life cycle," *The Review of Financial Studies*, 18, 491–533.

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.

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.

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.

Goldsmith-Pinkham, P. and K. Shue (2021): "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. (1971): "Optimum consumption and portfolio rules in a continuous-time model," *Journal of Economic Theory*, 3, 373–413.

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*.

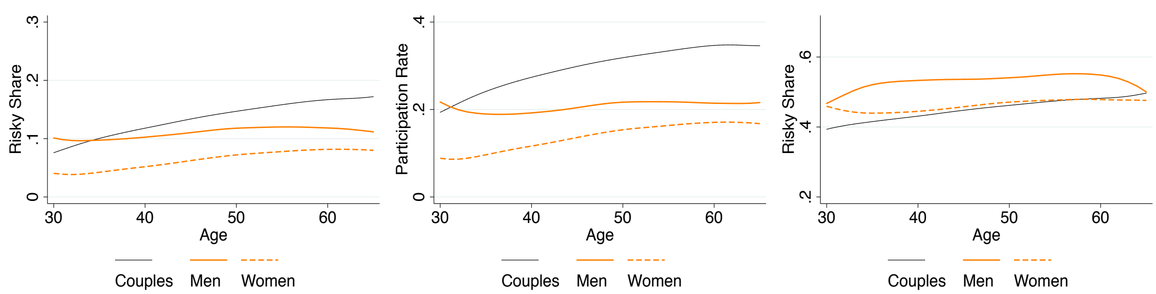
Yao, R. and H. H. Zhang (2005): "Optimal consumption and portfolio choices with risky housing and borrowing constraints," *The Review of Financial Studies*, 18, 197–239.

A Data Appendix

A.1 Supplementary Figures

Excluding Retirement Accounts. If single men are more likely to hold retirement accounts (e.g. because of the type of job they have) than single women, and if individuals, regardless of gender, tend to invest retirement savings more risky than other types of wealth, the gender investment gap would reflect gender heterogeneity in the labor market rather than in investment choices. Figure 9 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. It shows that, the gender gap in equity shares (Figure 9a) increases compared to the baseline, alleviating concerns that empirical investment differences across gender are mainly driven through savings that are linked to certain types of jobs.

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

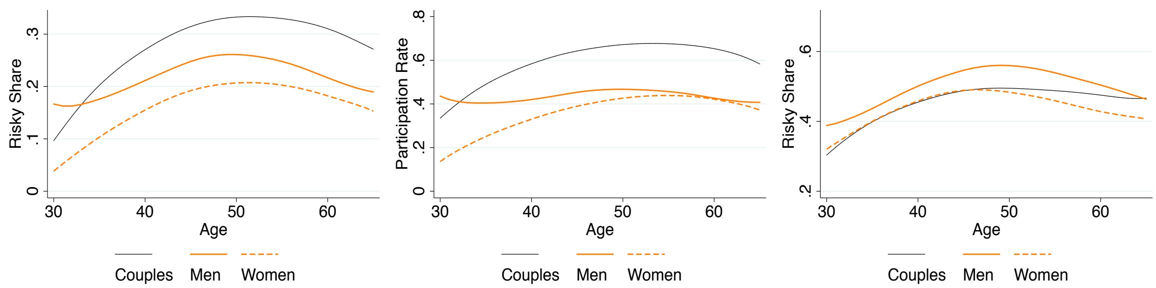


(a) Equity Share (b) Participation Rate (c) Conditional Risky Share

Notes: Figure 9 plots the life-cycle profiles of the equity share, stock market participation rates and conditional risky shares for singles and couples. 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 as well as 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 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 for cohort effects. Therefore, 10 plots the empirical patters 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.

Figure 10: Life Cycle Pattern of Household Finances – One Cohort



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

Notes: Figure 10 plots the life-cycle profiles of the equity share, stock market participation rates and conditional risky shares for singles and couples. Data is from the waves 1989 until 2016 of the Survey of Consumer Finances (SCF) for individulas 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.2 Regression Coefficients and Marginal Effects

Table 6: Regression Coefficients & Marginal Effects – Participation Rates of Singles

	(1) SMP	(2) SMP	(3) SMP
single women	-0.287***	-0.178***	-0.0853***
	(0.0103)	(0.00955)	(0.0107)
single woman * age	0.00485***	0.00310***	0.00158***
	(0.000205)	(0.000196)	(0.000226)
age	-0.136***	-0.0926***	-0.0770***
	(0.0140)	(0.0141)	(0.0167)
$age^2 * 100$	0.320***	0.223***	0.182***
	(0.0305)	(0.0308)	(0.0365)
$age^3 * 10000$	-0.244***	-0.171***	-0.141***
	(0.0216)	(0.0218)	(0.0259)
High education		0.275***	0.124***
		(0.00201)	(0.00246)
No. of HH members		-0.0417***	-0.0259***
		(0.000731)	(0.000742)
Income		0.0325***	0.0199***
		(0.000453)	(0.000256)
Safe assets			0.0550***
			(0.000593)
Constant	2.326***	1.151***	0.749***
	(0.206)	(0.208)	(0.246)
Observations	10,243	10,239	10,239
R^2	0.010	0.142	0.301
Year FE	No	Yes	Yes
ME for women at age 30	-0.1414***	-0.0846***	-0.0379***
	(0.0045)	(0.0040)	(0.0042)
ME for women at mean age (47)	-0.0572***	-0.0308***	-0.0105***
	(0.0024)	(0.0021)	(0.0019)
ME for women at age 65	0.0282***	0.0239***	0.0173***
	(0.0041)	(0.0041)	(0.0046)

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, *** p<0.01, ** p<0.05, * p<0.1

Table 7: Regression Coefficients & Marginal Effects – Conditional Risky Share of Singles

	(1) Cond. Share	(2) Cond. Share	(3) Cond. Share
single women	0.0174	0.0239	-0.0279
	(0.0198)	(0.0222)	(0.0185)
single woman * age	-0.00196***	-0.00193***	-0.00104***
	(0.000424)	(0.000460)	(0.000403)
age	0.0461*	0.0502**	0.0467**
	(0.0269)	(0.0254)	(0.0231)
$age^2 * 100$	-0.0854	-0.0961*	-0.0798
	(0.0573)	(0.0541)	(0.0491)
$age^3 * 10000$	0.0524	0.0611	0.0485
	(0.0396)	(0.0374)	(0.0339)
High education		0.00267	0.0662***
		(0.00288)	(0.00324)
No. of HH members		-0.0165***	-0.0256***
		(0.00243)	(0.00182)
Income		0.00296***	0.00553***
		(0.000628)	(0.000671)
Safe assets			-0.0639***
			(0.000893)
Constant	-0.276	-0.436	0.138
	(0.408)	(0.391)	(0.357)
Observations	4,521	4,521	4,521
R^2	0.019	0.049	0.242
Year FE	No	Yes	Yes
ME for women at age 30	-0.0415***	-0.0340***	-0.0591***
	(0.0072)	(0.0085)	(0.0065)
ME for women at mean age (48)	-0.0773***	-0.0691***	-0.0782***
	(0.0020)	(0.0013)	(0.0014)
ME for women at age 65	-0.1101***	-0.1014***	-0.0956***
	(0.0081)	(0.0078)	(0.0078)

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, *** p<0.01, ** p<0.05, * p<0.1

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). In particular, I assume the unexplainable part of the income process (that is, the residual term ϵ_{it} from the income equation) to follow a persistent-transitory process:

$$\tilde{y}_j = z_j + \epsilon_{\tilde{y}}$$

$$z_{j+1} = \rho_z z_j + \nu_z$$

A persistent-transitory process requires identification of three parameters: The persistence parameter ρ_z, the variance of the persistent shock $\sigma_{\epsilon^2_{\tilde{y}}}$ 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 $Var(z_{-1}) = 0$. However, the PSID only 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 8: Regression Coefficients for Income Estimation (Deterministic Component)

	(1) First Stage	(2) Second Stage	(3) First Stage	(4) Second Stage
	Couples		Singles	
high educ.		0.625***		0.558***
		(0.0172)		(0.0256)
Woman		-3.759***		-0.804***
		(0.0338)		(0.0260)
Woman*high educ.		-0.381***		0.202***
		(0.0433)		(0.0332)
age	0.0939***		0.0595***	
	(0.0167)		(0.0121)	
$age^2 * 100$	-0.102***		-0.0743***	
	(0.0177)		(0.0123)	
age*woman	0.0510***		0.0117***	
	(0.00493)		(0.00415)	
Constant	7.347***	1.526***	9.643***	0.0588***
	(0.392)	(0.0137)	(0.287)	(0.0199)
Observations	64,000	64,000	12,750	12,750
Number of unique indiv.	9,154		2,717	
R^2	0.015	0.374	0.010	0.221

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: Log of annual income (labor income, social security income and transfers) of the household head. In years where social security (transfer) income is not available separately by head and spouse, I use 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 again 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 woman; Robust standard errors in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$

B.3 Marriage and Divorce Probabilities

Table 9: Regression Coefficients for Marriage and Divorce Probabilities

	(1) Marriage Prob.	(2) Divorce Prob.
Woman	-0.309***	
	(0.0616)	
Age	-0.0118	-0.0360
	(0.0320)	(0.0281)
$age^2 * 100$	-0.0537	0.0124
	(0.0368)	(0.0306)
$\mathbb{1} > 1997$	0.466***	0.578***
	(0.0572)	(0.0536)
High Educ. (Head)	0.118*	-0.517***
	(0.0606)	(0.0602)
High Educ. (Spouse)		-0.137**
		(0.0610)
Constant	-0.879	-2.352***
	(0.675)	(0.620)
Observations	15,292	67,889

Notes: Estimations are based on Logit regressions from PSID Data, waves 1989-2017. Corresponding Figure is Figure 5 in the main text. Dependent variable: Likelihood of getting married (resp. divorced) within the next year, conditional on not being married (resp. being married) today. The age of a couple is the average age of both spouses. For education within couple, head refers to the husband and spouse refers to the wife. In contrast, singles are always labeled as head. *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 woman; $\mathbb{1} > 1997$ indicates observations that were interviewd after 1997 to account for the changing frequency of the PSID. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

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 in the terminal period (T) die with certainty 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 choice in period T (for each point in the state space), I iterate one period backwards 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 additionally 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).

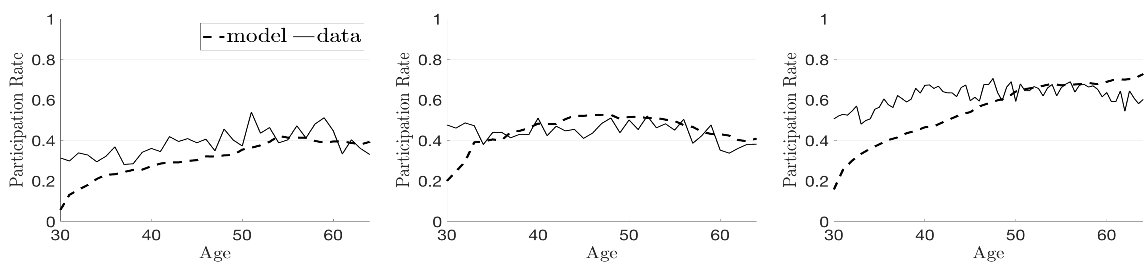
After having solved for the policy functions, I simulate the model for a large number of individuals over their life-cycle. At age 30, I assign each individual an initial level of wealth, education, and marital status as in the data. Next, I simulate a chain of marital transition and labor income shocks (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, its partner is assigned from "outside" of 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 21 minutes when running the code on a cluster with 36 nodes.

C.2 Further Results on Model Fit

This section reports the model fit for the life-cycle profiles when splitting the equity share along the extensive margin (participation rate) and the intensive margin (conditional risky share). Whereas the life-cycle profiles for couples and single men were targeted in the calibration exercise, those of single women serve as undtargeted moments. The model does a good job at matching the life-cycle profiles of participation rates, especially for single households (Figure 11). However, in contrast to the data, it predicts the conditional risky share to be declining in age (Figure 12) and overshoots the data in particular for single women. This difficulty of portfolio choice models to match the 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, in particular for single men.

Figure 11: Model Fit of Participation Rates



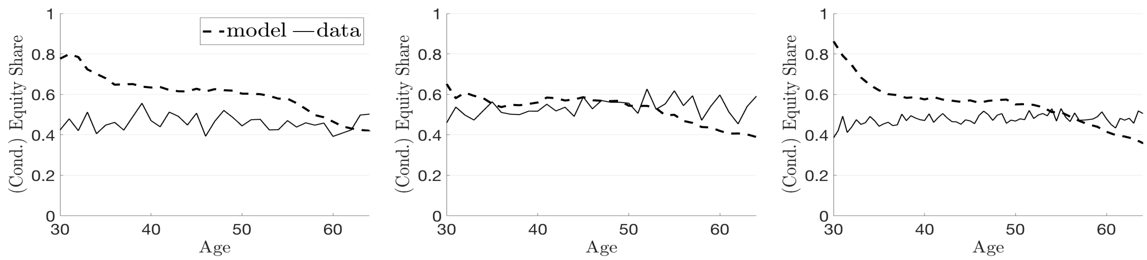
(a) Single Women (b) Single Men (c) Couples

Notes: Figure 11 plots the model fit of participation rates for single women, single men and couples. 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.

C.3 Further Results on Composition vs. Policy Effect

Figure 13 reports the life-cycle profiles for the equity share of single women in the baseline model (black solid line), the unrestricted counterfactual (black dashed line) and the counterfactual that restricts policy functions for the risky share to be the same as in the baseline

Figure 12: Model Fit of Conditional Risky Shares

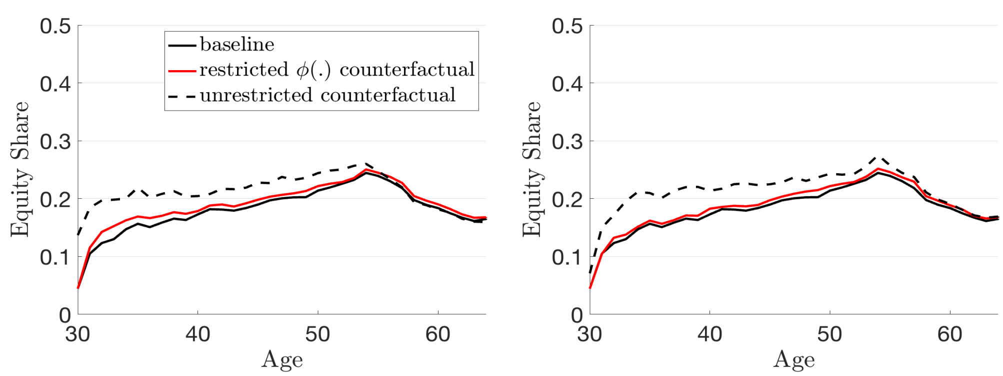


(a) Single Women (b) Single Men (c) Couples

Notes: Figure 12 plots the model fit of conditional risky shares for single women, single men and couples. 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.

(red line). First, it demonstrates that the policy effect (i.e. the gap between the red line and the black dashed line) is important in explaining differences in investment patterns of single women between the baseline model and both the income as well as household size counterfactual. Moreover, beyond age 55, the absolute difference across all three specification is small which is why I abstract from reporting the relative importance of each channel for these ages in Figure 8.

Figure 13: Composition vs. Policy Effect (Single Women)



(a) Income Counterfactual (b) HH Size Counterfactual

Notes: Figure 13 contrasts the life-cycle profiles of single women's equity share from the baseline model (black solid line) to the unrestricted income counterfactual (black dashed line) and to the counterfactual that restricts the policy function for the risky share to be the same as in the baseline model (red line).