Department of Economics and Management "Marco Fanno"
University of Padova
XXXIV Cycle

PhD in Economics and Management
Curriculum in Economics

Risk factors and risk attitudes in household portfolio decisions

Supervisor Professor Guglielmo Weber
Co-supervisor Professor Chiara Dal Bianco
PhD coordinator Professor Enrico Rettore

PhD candidate Francesco Maura

Contents

Acknowledgments

First and foremost, I am extremely grateful to my supervisors, Professor Guglielmo Weber and Chiara Dal Bianco, for their invaluable advice, continuous support, and patience during my Ph.D. journey. They have taught me how to approach and conduct both academic research and academic life, even during the Covid-19 pandemic. Last, I want to thank them for having taught me to pursue my ideas even when I have encountered difficulties during my research.
I would also like to thank professor Viola Angelini for the warm welcome, support and supervision during my visiting period in Groningen.
A great thanks also to my Ph.D. colleagues, Isha, Debdeep and David. Thanks for the nice talks, the mutual support and the suggestions.
Thanks to my friends, Andrea, Giovanni, Luca, Giacomo, Enrico, Francesco, Lorenzo, Alice, Lia, Ilaria and Valentina, for their patience and the chilling moments during these three years and for the Zoom calls during the pandemic. Even if I'm not physically close to all of you for (almost) 5 years, nothing has changed in our relationship.
Thanks to my flatmates, Giorgia and Martina, and their boyfriends, Riccardo and Sandy, for the great time we had together, before, during and after the lockdown.
Thanks to my parents and my brothers. I really appreciate all the support and the sacrifices you do for me in these three years.
And, finally, thanks to Giulia. For all the moments and the small things that we shared in this last year.

Introduction

Economic agents are continuously exposed to situations involving imperfect knowledge and have to make decisions under incomplete information. Decisions under uncertainty are complicated tasks for households and individuals, who generally try to make the best possible choice given the available information and taking into account the risk implied by that specific choice. For example, portfolio allocation, saving for retirement or housing investment involve a high degree of uncertainty. They are important choices for households and individuals because they affect their current and future well-being.

An economic investigation of the mechanisms and factors that drive the decision-making process under uncertainty, the source of risk and the attitudes toward risk of agents are therefore relevant points in the research agenda.

This dissertation is an attempt to improve our understanding of how risk factors and risk behaviours influence household choices.

Chapter 1 focuses on households' earnings process, analyzing their heterogeneity by the education. First, I show that earnings shock distribution deviates from log-normality, violating the standard assumption of the canonical model generally used in macroeconomics to study the earnings process. I follow the nonlinear approach proposed by Arellano, R. Blundell, and Bonhomme (2017) to study the characteristics of the earnings process. This model allows for different shocks being associated with different persistence of earnings and moments of the household's life-cycle. The notion of persistence identifies the persistence of earnings history on current and future earnings. The estimates show that the nonlinear model fits better the data, following the dependence structure of earnings. The empirical estimates show that the earnings' persistence is higher when a good shock hits high-earnings households and when a bad shock hits low-earnings households. The earnings processes of high- and low-educated show similar patterns, but the former has higher persistence on average. I simulate the impacts of different shocks on low- and high-educated households' earnings to test the insurance mechanism of education. When hit by a negative shock, the low-educated show a faster recovery in the short run, while the high-educated return almost to the initial level of earnings in the long run. Thus, the simulation suggests that education acts as an insurance mechanism against negative earnings shocks. Following Guvenen et al. (2019), I quantify the risk premium implied by earnings uncertainty for high- and low-educated. The estimates show that low-educated are willing to pay a higher percentage of their consumption to avoid earnings risk. The results of Chapter 1 show that earnings are a source of uninsurable risk for households. Chang, Hong, and Karabarbounis (2018) show that earnings uncertainty prevents workers from investing in risky assets. Moreover, as labor market risk is resolved, the same workers start taking more risk in their financial portfolios. Therefore, earnings risk and financial decisions are correlated and the first is a relevant factors for the household portfolio allocation.

In Chapter 2, I study risk behaviours instead of risk sources. In particular, I investigate the effect of household decision-makers risk tolerances in the portfolio allocation decision process. The contribution of this work is to extend the collective approach to the study of financial choices and compare it with the traditional, unitary model. I develop a theoretical model of portfolio allocation where households decide first about their risk preferences and then about stock market participation and optimal share of wealth allocated in risky assets. The model measures the risk preferences of the household as the weighted average of the risk tolerance of the two household decision-makers. The weights represent the bargaining powers of the two spouses. The empirical estimates show two relevant results: first, household risk preferences do not influence stock market participation but, once the household decides to hold stocks, they affect the optimal share of wealth allocated in risky assets. Second, comparing the collective and the standard unitary model, the first fits significantly better the data. Therefore, I conclude that the risk tolerance of both household decision-makers matters in the portfolio decision process.

Last, Chapter 3 studies the relationship between the perception of individual mortality rate and investments decisions. In particular, I investigate whether subjective survival expectations play a role in the portfolio allocation process of the elderly. I compare subjective survival reports and objective mortality rates following the approach of O'Dea and Sturrock (2021), and I show that most individuals are pessimistic about their survival chances among their 50s, 60s and 70s, but they become optimistic in their early 80s. Agents with optimistic survival expectations expect to live longer and therefore have a longer investment horizon too. Those individuals may have enough time to benefit from the equity risk premium, which is always positive in the long run. The empirical results show that survival beliefs play a relevant role in the portfolio allocation process. Indeed, stock market participation is increasing in subjective survival probabilities and life expectancy, especially for younger individuals. In other words, an agent with a long time horizon is more likely to participate in the stock market. The relationship between survival optimism and stock market participation is coherent with the well-established trend that risky assets outperform treasury bills in the long run (20 years).

Chapter 1

Education, earnings dynamic and background risk

1.1 Introduction

Every year, a new generation will enter the labour market for the first time, with a multiplicity of skills and backgrounds. In their working careers, each of those individuals will face his unique job experience, which could involve both success and failure episodes as well as positive and negative periods. These events will have direct consequences on the earnings process of agents, with different impacts and persistence of their effects.

Economists are interested in understanding the evolution of labour market histories of individuals. Indeed, studying the general properties of life-cycle earnings and earnings shocks may improve the knowledge of consumption and saving behaviours of households.

One (of many) open questions concerning earnings dynamics is related to the type of shocks that could affect individual or household earnings, the effects of that shocks and the persistence of the effects through earnings and consumption over the time. Understanding the nature of shocks' persistence is of key interest not only because it affects the permanent and the transitory nature of inequality and social mobility, but also because it drives part of the variation in consumption.

The earnings dynamics literature has generally worked with the assumption of normal distribution of shocks since its birth in the late 1970s. Therefore, researchers until recently did not investigate the possible implications and effects of higher order moments beyond the variance-covariance matrix.
Thanks to the availability of large panel dataset, in recent years an increasing number of works have found evidence of non-normality and non-linearity in the earnings and earnings shocks distributions. One of the first research was Geweke and Keane (2000), which emphasized the non-Gaussian nature of earnings shocks and fitted a normal mixture model to earnings innovations. More recently, Bonhomme and Robin (2009) analyse French earnings data and model the transitory component as a mixture of Gaussian distributions using a copula model. They find the distribution of this transitory component to be left skewed and leptokurtic. Guvenen et al. (2019) use a large panel data set of earnings histories drawn from U.S. administrative records to analyse the life-cycle evolution of labour income. They reach two main conclusions using non-parametric methods: first, earnings shocks display substantial deviations from the log-normal distribution, i.e. from the standard assumption in the incomplete market literature. Second, they show how statistical properties of this distribution vary significantly over the life-cycle and with the earnings level of individuals. They also show that the persistence of earnings shocks changes depending on the magnitude of the shock and on the income level.

A second assumption that characterizes the earnings dynamic literature is that the persistence and the second and higher conditional moments of earnings shocks are independent of age and of the earnings history. However, recent papers have documented that earnings dynamics violate these assumptions, e.g. Meghir and Pistaferri (2004); Arellano, R. Blundell, and Bonhomme (2017); Guvenen et al. (2019). Earnings shocks deviate from log-normality, display strong negative skewness and high kurtosis, and their second and higher moments vary over the life cycle and across the earnings distribution.

Arellano, R. Blundell, and Bonhomme (2017) develop a new quantile-based panel data framework to study the nature of income persistence. As in the canonical model Abowd and Card (1989), log earnings are the sum of a general Markovian persistent component and a transitory innovation. However, they allow for different shocks being associated with different persistence, where the notion of persistence is the one of persistence of earnings histories. This definition of persistence implies that a particular shock may wipe out the memory of past shocks in the forthcoming earnings evolution.

I apply the estimation procedure of Arellano, R. Blundell, and Bonhomme (2017) to Italian household survey data and compare its results with the canonical earnings process. Moreover, I investigate the heterogeneity in the earnings dynamic by education level of the head of the household.

First, the estimates show that the nonlinear model of earnings fits better the data than the canonical one, reproducing almost perfectly the pattern of the earnings process. The persistence of Italian earnings history is higher when a good shock hits high-earnings households and when a bad shock hits low-earnings households, in line with the results of US and Norwegian data (Arellano, R. Blundell, and Bonhomme (2017)).

The analysis of the heterogeneity of the earnings process by households education shows that high education acts as an insurance mechanism against negative earnings shocks. In other words, the persistence of earnings history is higher for the high-educated than low-educated and the difference between the two groups is more significant among high earners. This might imply that low-educated households (i) take longer to recover from a bad shock and (ii) benefit more from a good shock. I simulate the effects of different shocks on households earnings process and this exercise provides evidence in favour of the first implication but not the second. Using the background risk approximation suggested by Guvenen et al. (2019), I provide evidence of the insurance mechanism of education, showing that the proportion (in terms of consumption) of uninsurable risk faced by highly educated households is lower.

The rest of the paper is organized as follow: Section 1.2 describes the data and the sample selection procedure, Section 1.3 presents the theoretical framework of the canonical model and its estimates using SHIW data, Section 1.4 introduces the nonlinear model and motivates its use, Section 2.4 presents the main results about earnings persistence and studies its heterogeneity by education, Section 1.6 analyses the relationship between background risk and education and Section 3.5 concludes.

1.2 Data and sample selection

I use data from the Survey of Household Income and Wealth (SHIW) conducted by the Bank of Italy. The survey is biennial on a random sample of about 8000 households, representative of the Italian population. About 50 percent of the sample is followed longitudinally. The survey collects information on demographic characteristics, occupational status and income sources of all household members. Over the years the SHIW questionnaire has undergone numerous changes, therefore, I select the waves from 1989 to 2016 in which the questionnaire structure remained rather stable. All the monetary amounts are expressed in euros, including the ones for the survey years preceding 2002. I convert nominal earnings records into real values using the deflator provided by the Historical Database, where the base year is 2010.

I pool together five balanced samples of six waves in which the first year of observation goes from 1998 to 2006, resulting in a sample covering the period 1998-2016.

This paper aims to investigate the dynamic of net household labor earnings (sum of net employment and self-employment income of all household members), then I select only those households whose head is a man aged 25–60, i.e. a man in its working age. I prefer male head of household because of the higher irregularities of females earnings and job market participation, especially in Italy (Mussida and

Picchio (2014)).

The final balanced sample is composed of 686 households. I then distinguish between low- and high-educated households considering the level of education of the household head. I define as low-educated those having at most lower secondary education (345) and as high-educated those with upper secondary education or above (341).

Table 1.1 shows the sample descriptive statistics for the entire sample and by education. The main difference between high and low-educated households is the one related to annual earnings, where low-educated earn 30% less, circa. All the other demographics (number of children, age and so on) are comparable and show little differences between the two groups.

Table 1.1: Descriptive statistics. Overall sample of households, low- and high-educated households.

	Overall sample	low-educated	high-educated
Log hh earnings	10.24	10.08	10.40
Real hh earnings (€)	32,640	27,880	37,460
Family size	3.48	3.56	3.40
Number of children	1.50	1.56	1.44
Age of hh head	47.11	47.36	46.86
Number of income recipients	1.74	1.74	1.73
North-East	19.3%	18.3%	20.3%
North-West	26.1%	26.9%	25.2%
Center	19.5%	21.2%	17.9%
South	19.5%	17.3%	21.7%
Islands	15.6%	16.3%	14.9%
Number of observations	686	345	341

As common in this literature, I adjust the earnings measure for demographic differences by regressing log earnings on a set of controls. In particular, I use cohort dummies interacted with household head's education categories (and of the partner when present), household macro-area of residence (North-West, North-East, Centre, South and Islands), a dummy for living in a large-city, family size, number of children and of income recipients. Moreover, I control for economic cycle and macro indicators such as (Italian) GDP and unemployment rate, and their interaction with cohorts[1]. Indeed, the used SHIW waves cover the period 2008–2012, in which the financial crisis hits the global economy and Italy experimented also huge government debt problems. I end up with three residualize earnings series: one for the entire sample, one for high-educated and one for low-educated. I use these last two residualized earnings separately in Section 2.4 to compare the differences implied by education on the persistence of earnings shocks.

From now on, I refer to earnings as the regression residuals.

1.3 Canonical model

In this Section, I introduce the canonical linear model largely used in macroeconomics and present the results of its estimation using the SHIW dataset.

1.3.1 Earnings process

Let i denotes individuals i=*1,...,N* and t denotes time t = *1,2,...,T* (here, time coincides with age as I assume a specific cohort is considered). As before, let $y_{i,t}$ identify the log-earnings of individual i at age t. The canonical model can be written as:

$$y_{i,t} = \eta_{i,t} + \varepsilon_{i,t} \;\; i = 1, ..., N \;\; t = 1, ..., T \tag{1.1}$$

[1]GDP and unemployment rate data are from the Italian Statistical Institute.

where the η and ε have absolutely continuous distributions. The first component $\eta_{i,t}$ is the *persistent* component of earnings and follows an autoregressive process of order one. The second component $\varepsilon_{i,t}$ is the *transitory* one, has zero mean by assumption and is independent over time such that:

$$\eta_{i,t} = \rho\eta_{i,t-1} + \zeta_{i,t} \tag{1.2}$$

$$\eta_{i,1} \overset{id}{\sim} \mathcal{N}(0, \sigma_{\eta_1}) \quad \zeta_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_{\zeta}) \quad \varepsilon_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_{\varepsilon}) \tag{1.3}$$

where ρ is the *persistence* of past earnings history on today earnings. Equation 1.2 implies that the persistence of earnings is fixed and constant over time, i.e. it does not depend on the age of the agent and current earnings shocks.

1.3.2 Parameters estimation

I estimate the parameters of the canonical linear process for earnings residual using standard minimum distance techniques (see for example R. W. Blundell et al. (2016)) [2]. I estimate the parameters of the entire sample, high- and low-educated separately, to capture the different characteristics of the groups and compare them. Table 1.2 shows the results.

Table 1.2: Estimates from the canonical earnings process, only men

Parameters	All sample	high-educated	low-educated
$\sigma^2_{\eta_1}$	0.293	0.299	0.395
σ^2_{ζ}	0.200	0.132	0.326
σ^2_{ε}	0.226	0.234	0.091
ρ	0.760	0.952	0.482

[2] Section A in the Appendix provides additional details on the moments used.

As common in the literature, the persistence ρ is close to the unit root, in particular for the entire sample and the high-educated households. The low-educated group, however, has a lower coefficient: this is in line with the expectation, because a low level of education gives access to different job market, generally characterized by a higher uncertainty. On the other hand, the estimated ρ of the high-educated is comparable to the results presented by Cappellari (2002) and Cappellari (2004) using Italian administrative data. Cappellari (2002) and Cappellari (2004) worked on Italian data, analysing the earnings dynamics with the canonical model. Their focuses were, respectively, the cross-sectional earnings differential growth, low-pay persistence and its probability and the differences in earnings growth between the private and the public sector. In both cases, authors want to study earnings shocks persistence. Their results validate the outcomes of the parameters of the canonical model that I estimate with a different sample.

1.4 Nonlinear earnings

1.4.1 Non-linearity

The canonical linear model generally used to study the household earnings process imposes three main restriction: age-independence of the persistence of the shock distribution, which implies age-independence of the second and higher moments of the conditional distributions of both the transitory and the persistent component; normality of shock distributions and linearity of the process for the persistent component.
Table 1.3 provides evidence that these assumptions are violated by the data, showing that the second and higher order moments of earnings growths by age intervals are clearly non-normal. Moreover, the moments of the earnings shocks distribution vary across age, showing that also the assumption of age independence is violated.

Table 1.3: All sample: residual earnings shocks conditional moments by household head age.

age	25-34	35-41	42-48	49-54	55-60
standard dev	0.52	0.39	0.37	0.36	0.37
skewness	-0.06	-0.08	-0.21	-0.05	-0.15
kurtosis	4.82	8.61	9.82	9.94	8.92

1.4.2 Non-linear earnings process

Because the earnings shocks distribution is non-normal, I decide to use the non-linear model proposed by Arellano, R. Blundell, and Bonhomme (2017) to study the earnings process of Italian households and its heterogeneity by education. In this framework, log-earnings are the sum of a general persistent component, which follows a Markov process and a transitory innovation. Authors interest mainly centres on the conditional distribution of the persistent component given its past, which is considered as a comprehensive measure of the earnings risk faced by households.

Let i denote household i=*1,...,N*, and t denote time t = *1,2,...,T* (time coincides with the age of the household head as I assume a specific cohort is considered). $y_{i,t}$ denote the log-earnings of household i at time t. I assume that $y_{i,t}$ is the sum of a persistent ($\eta_{i,t}$) and a transitory component ($\varepsilon_{i,t}$), such that:

$$y_{i,t} = \eta_{i,t} + \varepsilon_{i,t} \quad i = 1, ..., N \quad t = 1, ..., T \tag{1.4}$$

The persistent component $\eta_{i,t}$ follows a general first-order Markov process, with no restriction on the dependence structure. The transitory component is a zero mean variable, is independent over time and of the persistent component. The last assumption is that the probability distribution of both components is absolutely continuous.

Without making the standard parametric assumption of the canonical earnings process, following Arellano, R. Blundell, and Bonhomme (2017) I specify the dependence structure of $\eta_{i,t}$ with the function $Q_t(\eta_{i,t-1}, \tau)$. This function represents the τth conditional quantile of the persistent component $\eta_{i,t}$ given its previous realization, $\eta_{i,t-1}$. It follows that the dependence structure of $\eta_{i,t}$ can be written as:

$$\eta_{i,t} = Q_t(\eta_{i,t-1}, u_{i,t}), \qquad (u_{i,t}|\eta_{i,t-1}, \eta_{i,t-2}, ...) \sim Uniform(0,1), \qquad t = 2, \ldots, T. \tag{1.5}$$

This specification of the earnings process is flexible and has the ability to fit the nonlinear dynamics of earnings. In particular, the model captures the nonlinear persistence of η, defined as:

$$\rho_t(\eta_{i,t-1}, \tau) = \frac{\partial Q_t(\eta_{i,t-1}, \tau)}{\partial \eta}. \tag{1.6}$$

$\rho_t(\eta_{i,t-1}, \tau)$ is the function that identifies the persistence of $\eta_{i,t}$ when it is hit by a shock of rank τ. It depends both on the history of the persistent component ($\eta_{i,t-1}$) and the magnitude of the shock realization. The notion of persistence proposed by Arellano, R. Blundell, and Bonhomme (2017) is one of persistence of histories, meaning that large and rare (negative or positive) shocks are allowed to cancel out the memory and the effects of past shocks.

Similar general quantile representations are introduced for the transitory component and the initial condition. The quantile function $Q_t()$ is defined as follows:

$$Q_t(\eta_{i,t-1}, \tau) = \alpha_t(\tau) + \beta_t(\tau)' h(\eta_{i,t-1}) \tag{1.7}$$

where h is a 3^{rd}-order Hermite polynomial. With this specification, the nonlinear persistence quantities specified in Equation 1.6 become:

$$\rho_t(\eta_{i,t-1}, \tau) = \beta_t(\tau)' \frac{\partial h(\eta_{i,t-1})}{\partial \eta} \qquad \rho_t(\tau) = \beta_t(\tau)' \mathbb{E}\left[\frac{\partial h(\eta_{i,t-1})}{\partial \eta}\right] \tag{1.8}$$

thus allowing shocks to affect the persistence of $\eta_{i,t1}$ in a flexible way.

1.5 Italian household earnings dynamic

This Section presents the empirical estimates of the Italian earnings process. First, I estimate the model using the overall sample and then I split it in low- and high-educated households, separating the two estimation procedures of the persistence of earnings shocks. Then, I compare the persistence of high- and low-educated households and test the possible implication of the differences in the earnings processes simulating the effects of positive and negative shocks.

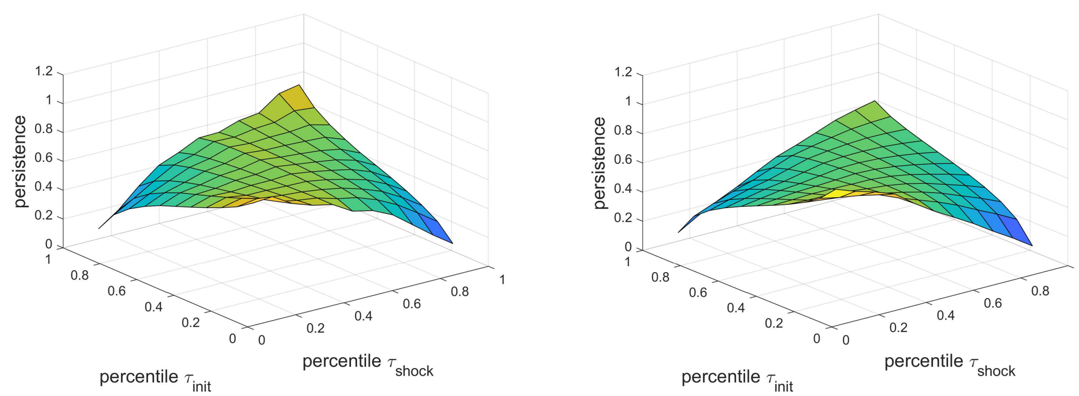
Figure 1.1a reports the average persistence of household earnings as a function of the percentiles of previous earnings $y_{i,t-1}$ and the percentiles of the shock distribution in the current period. In the plot, τ_{init} identifies the percentile of previous earnings and τ_{shock} identifies the percentile of the shock. The figure shows that the pattern of the estimates is in line with the findings of (**??**) on US and Norwegian data: when high-earnings households are hit by a good shock or a bad shock hits low-earnings households, the persistence of earnings history is higher. On the contrary, persistence decreases dramatically when high-earnings households are hit by bad shocks and low-earnings households by good shocks. Figure 1.1 reproduces the simulated nonlinear model, which is able to well reproduce the pattern of the persistence of SHIW data.

If we consider only the persistent component of the earnings process η, i.e. $y_{i,t}$ net of transitory shocks, the average persistence is higher then standard log earnings residuals, as Figure 1.2 shows. This is in line with the expectations: the persistent component of earnings is, by definition, stable over time and rarely changes. However, large negative earnings shocks (e.g.: health issue) or large positive shocks (e.g.: promotions) may change the history of past earnings of high-earners and low earners, respectively, modifying the persistent component of earnings.

Figure 1.3 shows the marginal distribution of the transitory component (Figure 1.3b) and of the persistent component (Figure 1.3a). The computation of the marginal distributions considers a households whose household head is aged 47 years, i.e. the average age in the sample. Both distributions are clearly non-normal, however the transitory component shows more pronounced deviations then the persistent

Figure 1.1: Nonlinear persistence of earnings history. Average derivative of the conditional quantile function of $y_{i,t}$ given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at τ_{shock} and at τ_{init}, computed on SHIW data (Graph (a)) and on simulated data (Graph (b)).

(a) Earnings persistence, SHIW data (b) Earnings persistence, nonlinear model



component, with a particularly high kurtosis and fat tails. Moreover, Figure 1.3b shows that the transitory component is generally close to 0, and therefore negligible in the long run of earnings. However, in few cases it might have a large impact on household finances and significantly contribute to the life-cycle variation of the earnings process.

Last, Figure 1.4 compares the quantile-based estimates of conditional dispersion, conditional skewness and conditional kurtosis of the SHIW data and the simulated earnings. The conditional moments (Figures 1.4a, 1.4b and 1.4c) document the ability of the non-linear model in replicating second and higher-order moments of earnings and capture their non-normality. Again, the non-normality of the original data that emerges from the graphs justifies the use of the non-linear model proposed by Arellano, R. Blundell, and Bonhomme (2017) and described in Section 1.4.

Figure 1.2: Estimates of the average derivative of the conditional quantile function of $\eta_{i,t}$ on $\eta_{i,t-1}$ with respect to $\eta_{i,t-1}$, based on estimates from the nonlinear earnings model.

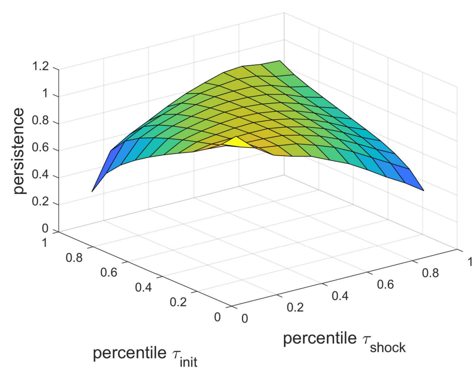
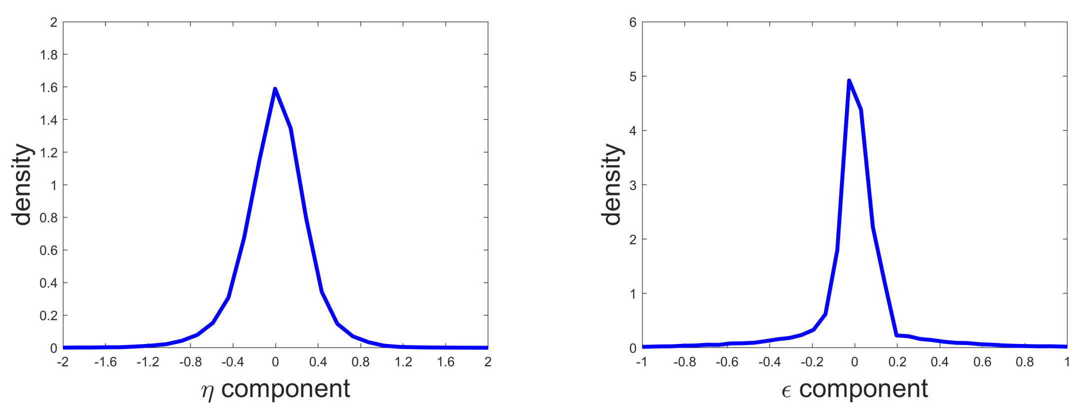


Figure 1.3: Estimated marginal densities of persistent and transitory earnings components at mean age (47 years).

(a) Persistent component ($\eta_{i,t}$) (b) Transitory component ($\epsilon_{i,t}$)



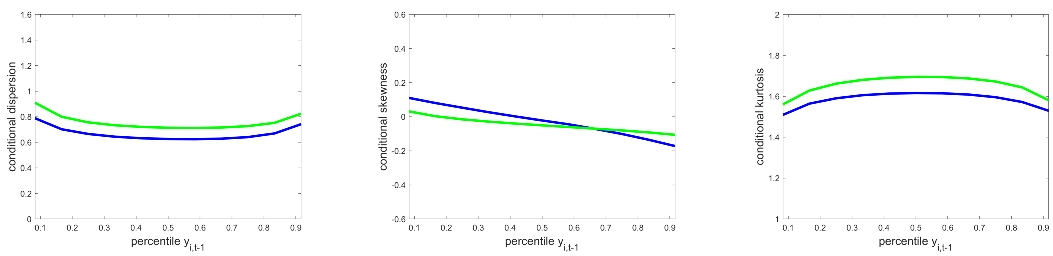
Note: non-parametric estimates of distributions based on simulated data, using a Gaussian kernel.

1.5.1 High and low-educated households earnings

In this Section, I investigate how the education level of the household head affects the nonlinear earnings dynamic. Education is a simple way to approximate the

Figure 1.4: Conditional dispersion, conditional skewness and conditional kurtosis of log earnings residuals $y_{i,t}$. Estimates from SHIW data (blue line) and simulation from the estimated nonlinear model (green line).

(a) Conditional dispersion (b) Conditional skewness (c) Conditional kurtosis

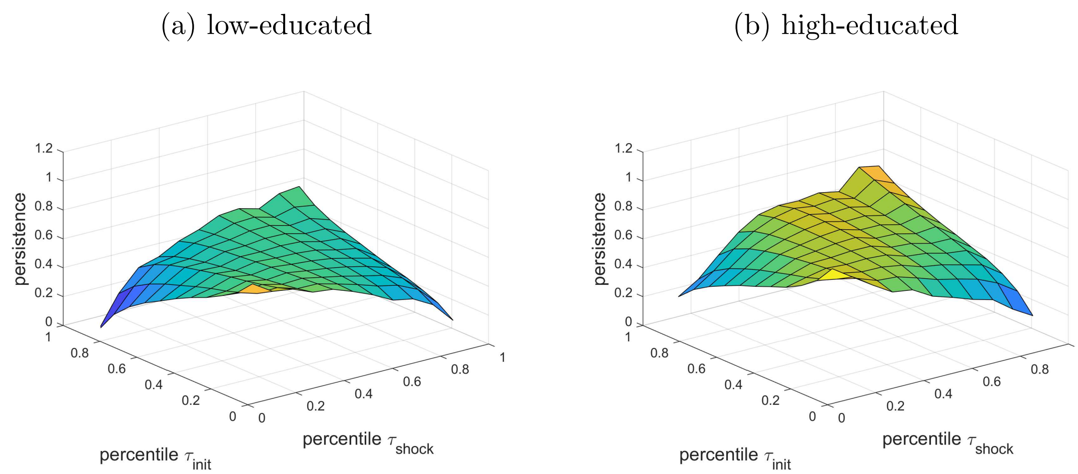


Note: Conditional dispersion is given by $\sigma(y;\tau) = Q(\tau|y_{i,t-1} = y) - Q(1-\tau|y_{i,t-1} = y)$, where $\tau = 11/12$. Conditional skewness is $skew(y,\tau) = \frac{Q(\tau)+Q(1-\tau)-2Q(1/2)}{Q(\tau)-Q(1-\tau)}$. Conditional kurtosis is $kurt(y,\tau) = \frac{Q(1-\alpha)-Q(\alpha)}{Q(\omega)-Q(1-\omega)}$ where $\tau = 11/12$, $\omega = 10/12$ and $\alpha = 1/12$.

heterogeneity in occupations of the households' head and the different labor markets they have access to. A large amount of literature shows the positive effects of a high level of education on labour, earnings and financial stability, especially during economic crisis. Figure 1.5 is in line with these results: while low- and high-educated households exhibit a similar pattern to the one of the overall sample (Figure 1.1), high-educated households have an average higher persistence of earnings (0.18–0.86) than low-educated ones (0–0.76)[3]. This means that, on average, the earnings of the high-educated are less likely to change when hit by any type of shock (positive or negative), while low-educated earnings are more volatile and subject to variation due to external conditions. Indeed, the estimated "lower-bound" of persistence is 0, i.e. there is a combination of percentiles of previous earnings and current earnings shock such that the past has no predictive power on current earnings. This result is totally in contrast with the canonical model, which implies a unique level of persistence ρ estimated to be close to 1.

[3]The appendix provides the model fit for low- and high-educated households in Figures B and C, as well as the conditional moments of high- and low-educated households earnings.

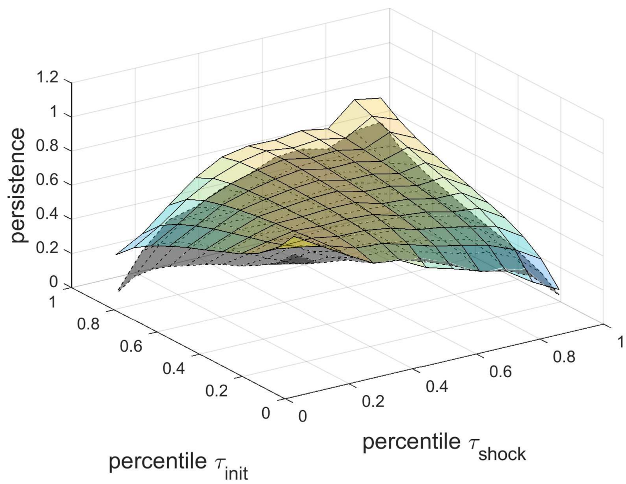
Figure 1.5: Nonlinear persistence of earnings history by education level. Average derivative of the conditional quantile function of $y_{i,t}$ given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at τ_{shock} and at τ_{init}. Low-educated (Graph (a)) and high-educated (Graph (b)) households.



Comparing the two patterns, the difference is larger among mid- and high-earnings households (Figure 1.6), therefore, if the household has low earnings, its earnings process does not depend on household head education and is the same for both groups.

Figure 1.7 plots the confidence bars, obtained with bootstrap methods, of the estimate presented in Figure 1.5. The Figure is constructed as follows: I fixed a different percentile of previous earnings τ_{init} each time and then I plot the correspondent values of the confidence bars for each percentile of current shock, τ_{shock}.
Persistence does not significantly differ for low levels of earnings in the two groups, as previously said, whereas it is significantly higher for high-educated households when τ_{init} is above the median. This seems to suggest that unusual shocks are more likely to wipe out the memory of past shocks when they hit low-educated households compared to high-educated ones, conditional on being in the upper part of the earnings distribution. Then, there are two possible implications of this estimates and this comparison: low-educated households (i) take longer to recover from a bad

Figure 1.6: Persistence of earnings history by education. Note: the grey surface shows the earnings persistence for the low-educated (Figure 1.5a), while the coloured surface the persistence for the high-educated (Figure 1.5b).



shock, and (ii) benefit more from a good shock. In the next Section I test this implication simulating the impacts of various types of shock on high- and low- educated households earnings.

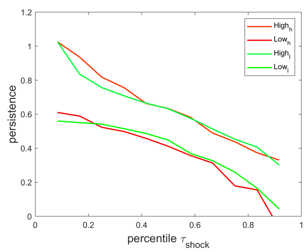
1.5.2 Simulation exercise

To verify whether low-educated households (i) take longer to recover from a bad shock, and (ii) benefit more from a good shock, I simulate the effects of a large negative ($\tau_{shock} = 0.1$) and a large positive ($\tau_{shock} = 0.9$) shock on the persistent component of the earnings process. The simulations consider a household whose household head is aged 45 years, τ_{init} (percentile of previous earnings) is set to 0.9 for everyone, and households are hit by the shock at age 47, that is the average age in the sample. The choice of $\tau_{init} = 0.9$ is justified, because Figures 1.6 and 1.7 show that the differences between high and low-educated households earnings process are among the high earners of the two distribution, i.e. those whose τ_{init} is 0.5 or more.

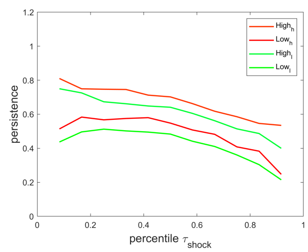
Figure 1.8 shows the difference between age-specific medians of log earnings of

Figure 1.7: Confidence bars for Figure 1.5. Every plot shows 90% confidence bars for specific values of τ_{init} for low-educated (green lines) and high-educated (red lines) households.

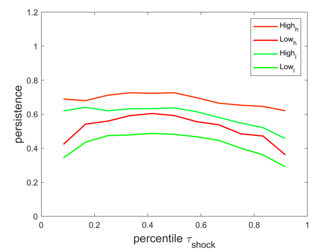
(a) Percentile $\tau_{init} = 1/11$



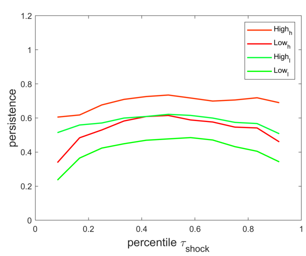
(b) Percentile $\tau_{init} = 3/11$



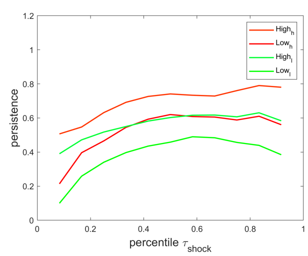
(c) Percentile $\tau_{init} = 5/11$



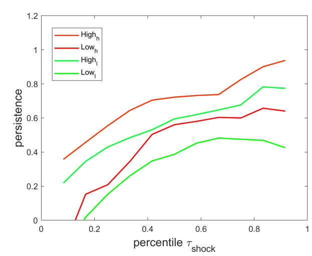
(d) Percentile $\tau_{init} = 7/11$



(e) Percentile $\tau_{init} = 9/11$



(f) Percentile $\tau_{init} = 11/11$



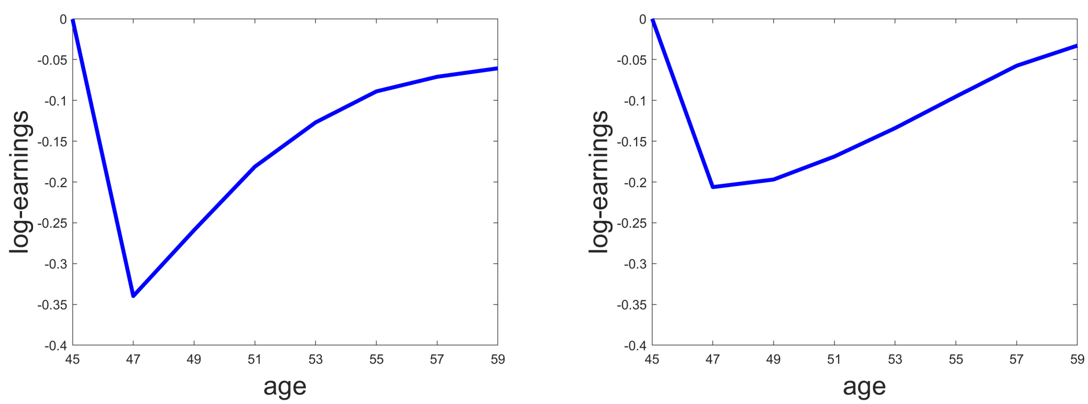
Note: Confidence bars are obtained with nonparametric bootstrap, still consistent under misspecification.

a household hit by a large negative shock ($\tau_{shock} = 0.1$) at age 47, and the same household hit by a median shock ($\tau_{shock} = 0.5$) at the same age. The large negative shock to the persistent component implies a 35% and 20% drop of log-earnings for low- and high-educated households, respectively. Analyzing the evolution of earnings in the years after the shock, the low-educated household shows a faster recovery in the short run, while the high-educated earnings return almost at their the initial level in the long period. Therefore, this simulation suggests that education acts as an insurance mechanism against negative earnings shocks.

The effects of a large positive earnings shock $\tau_{shock} = 0.9$ are shown in Figure 1.9. In both groups, this shock is much less persistent than the effect of a negative shock,

Figure 1.8: Effect of a negative shock ($\tau_{shock} = 0.1$) to the persistent component at average age in the sample (47), when the persistent component is at percentile $\tau_{init} = 0.9$ at age 45.

(a) low-educated (b) high-educated

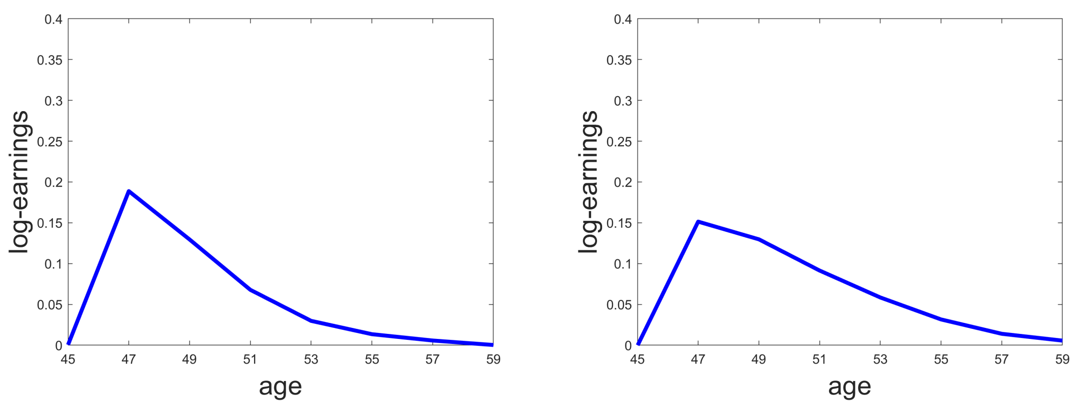


and has also a smaller initial impact on earnings. Comparing low- and high-educated groups, a positive shock is larger in magnitude for the former, 19% compared to 15% increase, but it is more persistent for the latter. Therefore, this simulation does not confirm the initial hypothesis that low-educated benefit more from a good shock: while the initial impact on earnings is slightly higher (and positive), its persistence effects runs out slightly faster.

The simulation exercise of Figures 1.8 and 1.9 compare the impact of large negative and large positive shocks on the earnings process of low- and high-educated households. However, the estimation procedure of the nonlinear model compute the persistence of earnings separately for the low- and high-educated. Therefore, the levels of τ_{init} and τ_{shock} could be very different in the two samples. In other words, the value of previous earnings for $\tau_{init} = 0.1$ may differ between the two groups because of the intrinsic characteristics of each categories (as seen in Table 1.1, low-educated households earns 30% less than high-educated).
For completeness, I run a second simulation exercise in which I impose the same

Figure 1.9: Effect of a positive shock ($\tau_{shock} = 0.9$) to the persistent component at average age in the sample (47), when the persistent component is at percentile $\tau_{init} = 0.9$ at age 45.

(a) low-educated (b) high-educated



initial conditions of earnings and shocks for the two groups. I consider the τ_{init} and τ_{shock} of the low-educated as the benchmark, and impose the same level of initial earnings and shock to the high-educated. In particular, the large positive shock $\tau_{init} = 0.9$ for the low-educated corresponds to $\tau_{init} = 0.88$ for the high-educated. The median shock ($\tau_{shock} = 0.5$) for the low-educated corresponds to $\tau_{shock} = 0.33$ for the high-educated and the large negative shock ($\tau_{shock} = 0.1$) for the low-educated to $\tau_{shock} = 0.03$ for the high-educated.

In Figure 1.10 I present the results of the simulation that used those percentile of initial earnings and current shock. Figure 1.10a is the benchmark (and reproduces Figure 1.8a): it represents the difference of the effects of a large negative shock $\tau_{shock} = 0.1$ and a median shock $\tau_{shock} = 0.5$ in a low-educated household with $\tau_{init} = 0.9$. Then, Figure 1.10b reports the effects of imposing to the high-educated the same level of initial condition and shocks of the low-educated: $\tau_{shock} = 0.03$ is the large negative shock, $\tau_{shock} = 0.33$ corresponds to the median shock.

Differently from the previous findings, high-educated households show a slower recovery than low-educated ones, even in the long period, and they are unable to reach the initial level of earnings before retirement. However, this is not in contrast with the results of the first simulation: Figure 1.10 compares shocks of the same level but quite different probability of occurrence ($\tau_{shock} = 0.1$ for the low-educated with $\tau_{shock} = 0.03$ for the high-educated). Because it is unlikely to be hit by such a large negative shock to earnings, high-educated households might consider this a negligible event and therefore they do not adequately prepare for it.

I do the same for a large positive shock using percentile $\tau_{shock} = 0.89$ for the high-educated which corresponds to $\tau_{shock} = 0.9$ for the low-educated. Results are reported in Figure 1.11.

In this case the difference in probability between the adjusted and the original shock is negligible and the results do not change significantly with respect to Figure 1.9. Indeed, as in the previous exercise a positive shock is slightly more persistent over time for the high-educated than for the low-educated. Therefore, we do not find evidence in favour of low-educated households benefiting more from a good shock than the high-educated ones.

1.6 Background risk in Italy

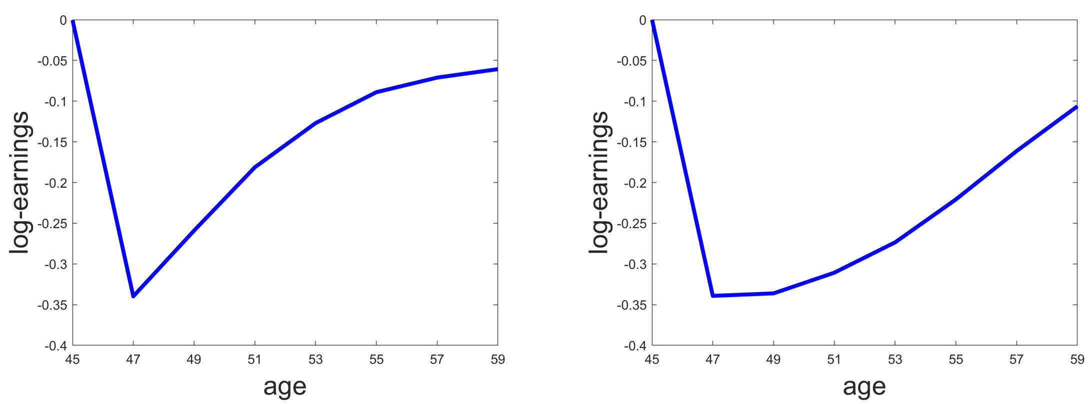
1.6.1 Quantifying background risk

In this Section, I use the collected information about life-cycle earnings dynamics to study the background risk of the Italian households. Background risk identifies those risks that households cannot insure against and integrate with the environment of the decisions. The most relevant of these risk is labour income risk .

Guvenen et al. (2019) derive an expression for individuals' background risk which accounts for the third and fourth order moments of the log-earnings distribution. The starting point is the experiment of Arrow (1965) and Pratt (1964): the decision-maker chooses between *(i)* a static gamble that changes the consumption c by a random proportion $(1 + \delta)$ (where δ is drawn from a Gaussian distribution) and *(ii)*

Figure 1.10: Effect of imposing to the high-educated the same level of initial condition, median and negative shock of the low-educated.
Note: the figure shows the difference between age-specific medians of log earnings of a household hit by a large negative shock at age 47, $\tau_{shock} = 0.1$ for the LE and 0.03 for the HE, and a household hit by a median shock for the LE that is $\tau_{shock} = 0.33$ for the HE; τ_{init} at age 45 is 0.9 for the LE and 0.88 for the HE (LE=low-educated, HE=high-educated).

(a) low-educated (b) high-educated



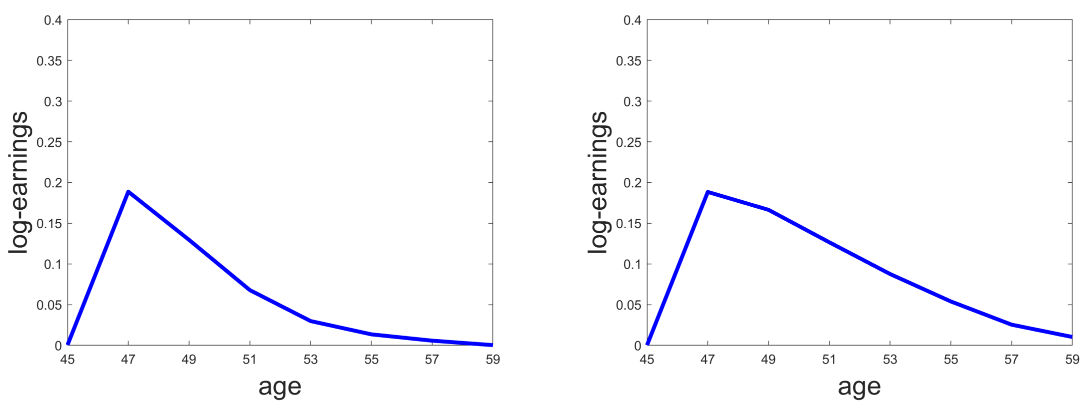
a fixed payment π to avoid this risk. Then, a utility maximizer individual solves the following problem:

$$U(c \cdot (1-\pi)) = \mathbb{E}\Big[U(c \cdot (1+\delta))\Big] \tag{1.9}$$

For simplicity, assume a constant relative risk aversion (CRRA) utility function $U(\cdot) = \frac{c^{1-\theta}}{1-\theta}$ with fixed coefficient θ, that represents risk aversion. Then, Equation 1.9 has only one unknown. Guvenen et al. (2019) take a first order Taylor expansion of the left hand side and a fourth order Taylor expansion of the right hand side. This transformation allows to include the effects of the second and higher order moments of earnings in the computation of the equivalence. Moreover, it also allows to isolate the certainty equivalent coefficient π as follows:

Figure 1.11: Effect of imposing to the high-educated the same level of initial condition, median and positive shock of the low-educated.
Note: the figure shows the difference between age-specific medians of log earnings of a household hit by a large positive shock at age 47, $\tau_{shock} = 0.9$ for the LE and 0.89 for the HE, and a household hit by a median shock for the LE that is $\tau_{shock} = 0.33$ for the HE; τ_{init} at age 45 is 0.9 for the LE and 0.88 for the HE (LE=low-educated, HE=high-educated).

(a) low-educated (b) high-educated



$$\pi \approx \frac{\theta}{2} \cdot \sigma_\delta^2 - \frac{(\theta+1)\theta}{6} \cdot \sigma_\delta^3 \cdot skew + \frac{(\theta+2)(\theta+1)\theta}{24} \cdot \sigma_\delta^4 \cdot kurt \tag{1.10}$$

In the right hand side, three different terms determine risk aversion, namely the second, third and fourth central moments of the log-earnings distribution. Indeed, $\frac{\theta}{2} \cdot \sigma^2$ is the expression generally used in the literature for "risk aversion" and in this derivation identifies a sort of *variance aversion*; $\frac{(\theta+1)\theta}{6} \cdot \sigma_\delta^3 \cdot skew$ is a *negative skewness aversion*; finally $\frac{(\theta+2)(\theta+1)\theta}{24} \cdot \sigma_\delta^4 \cdot kurt$ is the *kurtosis aversion*. These are the components of the background risk perceived by the households. Those components can significantly amplify the risk: as an example, the coefficient of kurtosis is cubic in θ and quartic in dispersion and has, therefore, a large impact on π.

Table 1.4: Estimates of the risk premium with SHIW data, $\theta = 2$

Sample	Std	Skew	Kurt	π
reference $\mathcal{N}$	0.2	0	3	4.48 %
All	0.192	-0.14	7.38	4.83 %
High education	0.187	-0.22	8.25	4.64 %
Low education	0.236	-0.71	7.87	8.95 %

1.6.2 Background risk of high and low-educated households

To estimate the risk perceived by Italian households, I evaluate Equation 1.10 substituting the central moments of log-earnings for the whole sample, high-educated and low-educated households, respectively. I compute the moments from simulated data to correct for possible measurement error.

Specifically, let the log-earnings process be $y_{i,t} = \eta_{i,t} + \varepsilon_{i,t}$; I consider the second component as comprehensive of the measurement error and the transitory shock. Therefore, ε is a measure of the noise in the data. To clear the possible noise effects, I use the log-earnings data simulated with the non-linear model procedure proposed by Arellano, R. Blundell, and Bonhomme (2017). They provide separate estimate for the persistent and the transitory component of earnings (respectively, $\hat{\eta}$ and $\hat{\varepsilon}$). Then, given $\hat{y} = \hat{\eta} + \hat{\varepsilon}$ from the expression above, I approximate the earnings process as $\hat{y} = \hat{\eta}$ and use the central moments of $\hat{\eta}$ to estimate π.

Table 1.4 shows the results of the π estimates. The first line represents a benchmark background risk where I assume that the earnings are normally distributed with a standard deviation of 0.2 (comparable to the SHIW simulated data) and skewness and kurtosis have the characteristic values of a Normal distribution, 0 and 3. Then, I use the central moments of the simulated data (considering only the persistent component of earnings) for each available sample to compute the risk premium π according to Equation 1.10.

π is the risk premium and it is expressed as percentage of consumption. Therefore,

π represents the willingness to pay of Italian households to avoid background risk in % of consumption.

As expected, the risk aversion of low-educated households is higher with respect to the one of high-educated. This means that low-educated prefer to pay a higher share of their consumption to avoid the uncertainty. These results are consistent with the findings of Section 2.4, where Figure 1.8 shows how education acts as source of insurance against earnings variation and reduces the level of background risk faced by the households. The simulation exercise of the previous Section is coherent with these findings. Indeed, low-educated households need more time to recover from a negative earnings shock. Therefore, they are more exposed to earnings risk.

1.7 Conclusion

In this paper, I analyse the Italian household earnings process and its heterogeneity by education. Then, I study the earnings as the main source of households' uninsurable risk and the differences of the risk premium by education.
First, I focus on the canonical earnings process generally used in macroeconomics, defined as the sum of a persistent and a transitory component. The model assumes the normality of log-earnings shocks and independence of the persistence of earnings from their past realization. I estimate the earnings' persistence parameter, showing that the results are in line with other similar studies on Italian data. I compare the estimates of the canonical model by education of the household, and I show that the low-educated have lower persistence of earnings history. The type of job and job market that the two groups have access to may cause these differences in the earnings process.

Second, I provide evidence that the assumptions of the canonical model are violated by the data, in line with some recent papers like (**??**). Indeed, earnings shocks deviate from log-normality and display strong negative skewness and high kurtosis compared with the value of a Gaussian distribution. Moreover, their second and

higher moments vary over the life cycle and across the earnings distribution. In particular, high kurtosis implies that most individuals experience negligible earnings shocks, few experience median shocks and a small but non-negligible group experiences extremely large shocks, both positive or negative, in a given year. These findings motivate me in studying the Italian earnings process using the non-linear approach proposed by Arellano, R. Blundell, and Bonhomme (2017). This approach allows for the dependence of persistence of earnings shocks from past earnings and current shock distribution, and therefore it reproduces better the characteristics of the earnings process.
The estimates are in line with the pattern of US and Norwegian households earnings described in Arellano, R. Blundell, and Bonhomme (2017). The persistence of earnings is higher when a positive shock hits high-income households or a negative one hits low-income households. On the other hand, a large positive shock is more likely to wipe out past earnings history when it hits low earnings households, while a negative shock has similar effects when hitting high earners.
Then, I compare the persistence of low- and high-educated households: persistence is lower for the former and shows significant differences only among the high earners of the two groups. Using simulated "impulse response" of earnings to large negative and positive shocks, I show that low-educated households have a slower recovery when hit by a large negative shock having the same occurrence probability in the two groups. Therefore, education acts as an insurance mechanism against negative shocks. However, there are no significant differences in the effect of large positive shocks on the earnings process of low- and high-educated households.

Last, I use the central moments of estimated non-linear earnings process to quantify the background risk of Italian households, following the suggestions of Guvenen et al. (2019). I decide to use the moments of the persistent component of simulated data to clear the effects of possible measurement errors in the data collection. The estimated risk premium is higher (in terms of percentage consumption) for low-educated households, whose willingness to pay to avoid risk is twice the willingness to pay of high-educated. This result confirms that education may act as a source of insurance against earnings shocks, especially the negative ones.

Appendix

Education, earnings dynamic and background risk

A GMM moments derivation

Following the suggestions of R. W. Blundell et al. (2016), I recall the canonical model for convenience:

$$y_{i,t} = \beta_0 + x_{i,t}\beta_x + \eta_{i,t} + \varepsilon_{i,t}$$

$$\eta_{i,t} = \rho * \eta_{i,t-1} + \zeta_{i,t}$$

where the indexes i and t stand for individual and time, y are earnings, x are control variables; η is the persistent earnings shock and it follows an AR(1) process with innovation ζ, and ε is the transitory earnings shock.
One possibility to identify the parameters characterizing the residual process is to use the autocovariance moments of the residual $\hat{y_{it}} = \hat{\eta_{it}} + \hat{\varepsilon_{it}}$, deriving:

$$var(y_{it}) = \rho^{2t}\sigma^2_{\eta_1} + \frac{1-\rho^{2t}}{1-\rho^2}\sigma^2_\zeta + \sigma^2_\varepsilon$$

$$cov(y_{it}, y_{i,t-1}) = \rho\left(\rho^{2(t-1)}\sigma^2_{\eta_1} + \frac{1-\rho^{2(t-1)}}{1-\rho^2}\sigma^2_\zeta\right)$$

$$cov(y_{it}, y_{i,t-l}) = \rho^l\left(\rho^{2(t-l)}\sigma^2_{\eta_1} + \frac{1-\rho^{2(t-l)}}{1-\rho^2}\sigma^2_\zeta\right)$$

Starting from the above expression, R. W. Blundell et al. (2016) suggest using 3 periods for the identification. Taking periods t = 1,2,3 (where 1 is the first observable period), they also derive:

$$\rho = \frac{E(y_{i3}y_{i1})}{E(y_{i2}y_{i1})}$$

$$\sigma^2_{\eta_1} = \frac{E(y_{i2}y_{i1})^2}{E(y_{i3}y_{i1})}$$

$$\sigma^2_{\zeta} = \frac{E(y_{i2}y_{i1})^2 - E(y_{i3}y_{i1})^2}{E(y_{i3}y_{i1})} + var(y_{i2}) - var(y_{i1})$$

$$\sigma^2_{\varepsilon} = var(y_{i1}) - \frac{E(y_{i2}y_{i1})^2}{E(y_{i3}y_{i1})}$$

B Additional material - low-educated households

Figure B.1: Low-educated households: nonlinear persistence of earnings history. Average derivative of the conditional quantile function of $y_{i,t}$ given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at τ_{shock} and at τ_{init}, computed on SHIW data (Graph (a)) and on simulated data (Graph (b)).

(a) Earnings, SHIW data

(b) Earnings, nonlinear model

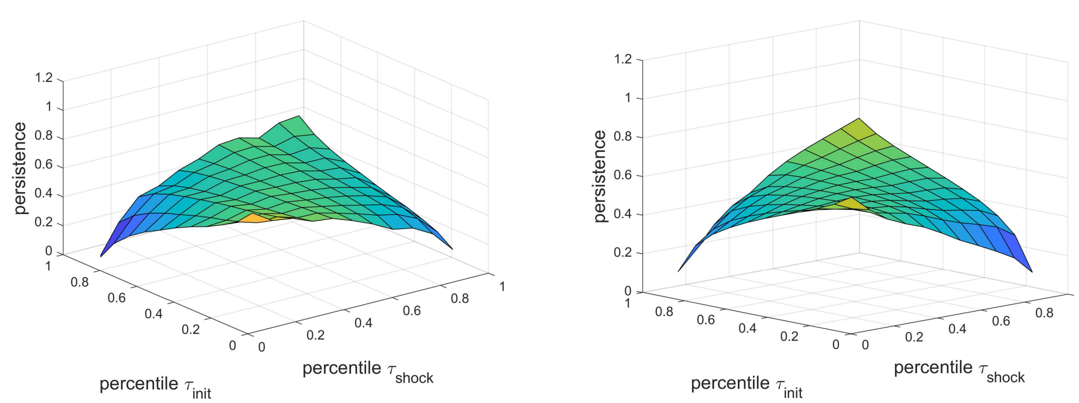


Figure B.2: Low-educated households: estimates of the average derivative of the conditional quantile function of $\eta_{i,t}$ on $\eta_{i,t-1}$ with respect to $\eta_{i,t-1}$, based on estimates from the nonlinear earnings model.

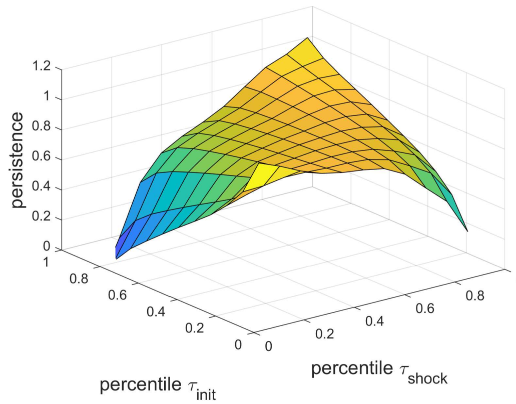
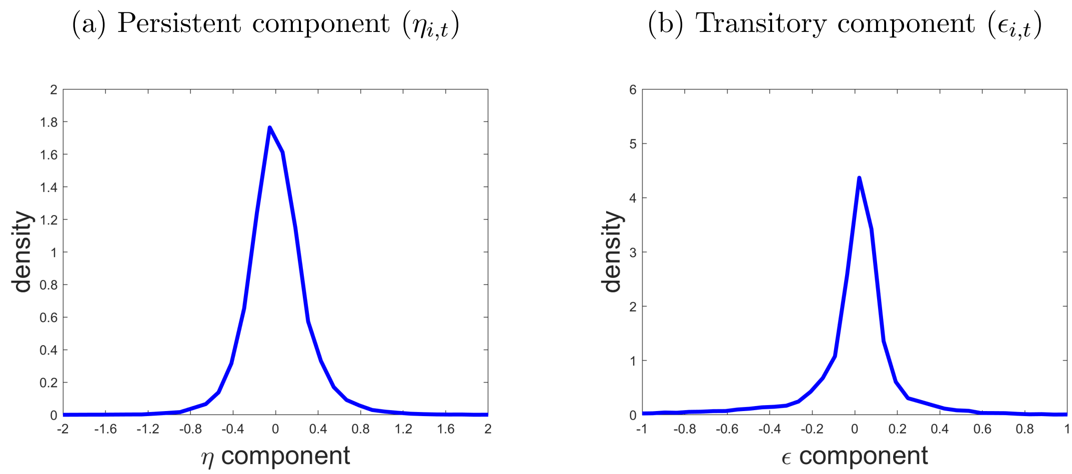


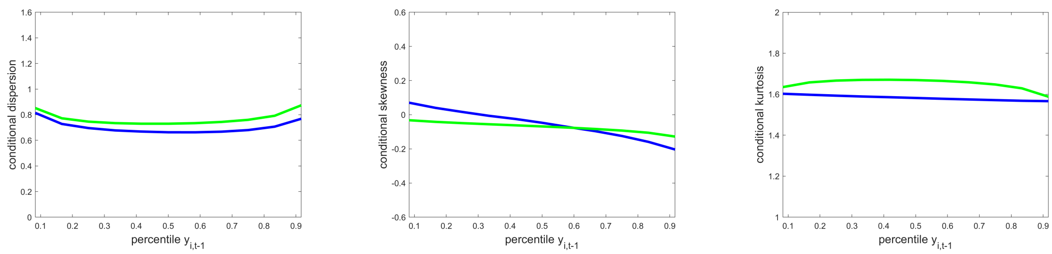
Figure B.3: Low-educated households: estimated densities of persistent and transitory earnings components at mean age (47 years).



C Additional material - high-educated households

Figure B.4: Low-educated households: conditional dispersion, conditional skewness and conditional kurtosis of log earnings residuals $y_{i,t}$. Estimates from SHIW data (blue line) and simulation from the estimated nonlinear model (green line).

(a) Conditional dispersion (b) Conditional skewness (c) Conditional kurtosis



Note: Conditional dispersion is given by $\sigma(y;\tau) = Q(\tau|y_{i,t-1} = y) - Q(1-\tau|y_{i,t-1} = y)$, where $\tau = 11/12$. Conditional skewness is $skew(y,\tau) = \frac{Q(\tau)+Q(1-\tau)-2Q(1/2)}{Q(\tau)-Q(1-\tau)}$. Conditional kurtosis is $kurt(y,\tau) = \frac{Q(1-\alpha)-Q(\alpha)}{Q(\omega)-Q(1-\omega)}$ where $\tau = 11/12$, $\omega = 10/12$ and $\alpha = 1/12$.

Figure C.1: High-educated households: nonlinear persistence of earnings history. Average derivative of the conditional quantile function of $y_{i,t}$ given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at τ_{shock} and at τ_{init}, computed on SHIW data (Graph (a)) and on simulated data (Graph (b)).

(a) Earnings, SHIW data (b) Earnings, nonlinear model

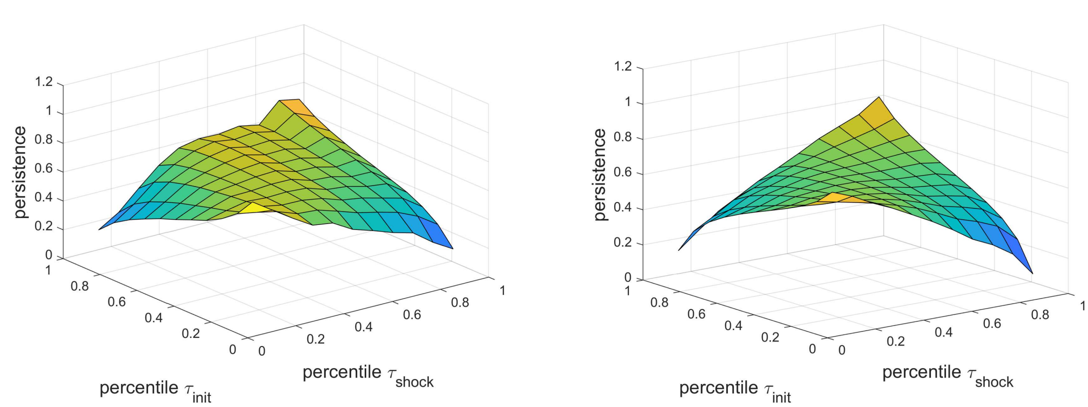


Figure C.2: High-educated households: estimates of the average derivative of the conditional quantile function of $\eta_{i,t}$ on $\eta_{i,t-1}$ with respect to $\eta_{i,t-1}$, based on estimates from the nonlinear earnings model.

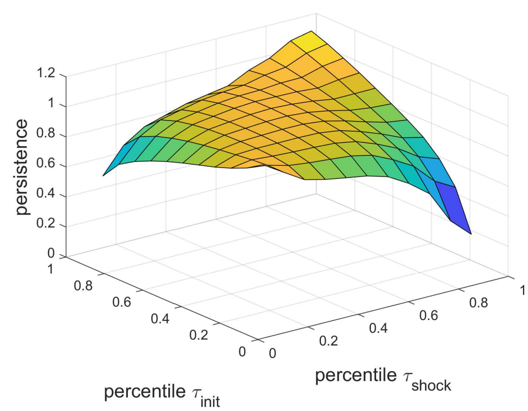
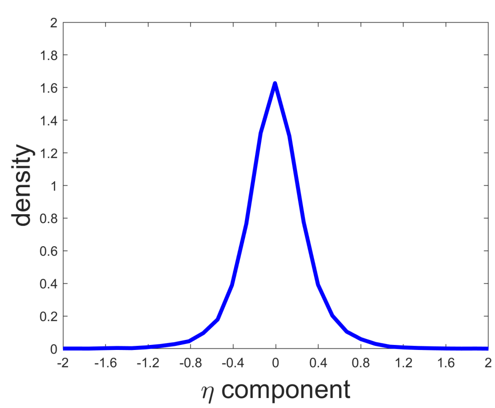


Figure C.3: High-educated households: estimated densities of persistent and transitory earnings components at mean age (47 years).

(a) Persistent component ($\eta_{i,t}$)



(b) Transitory component ($\epsilon_{i,t}$)

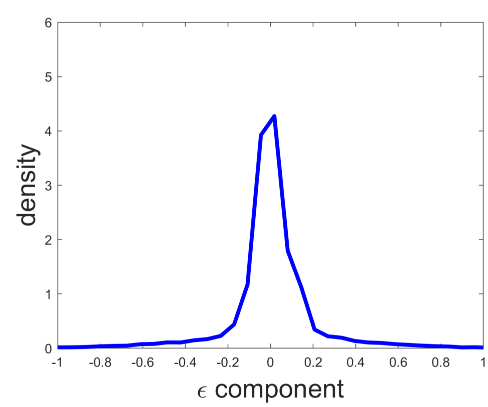
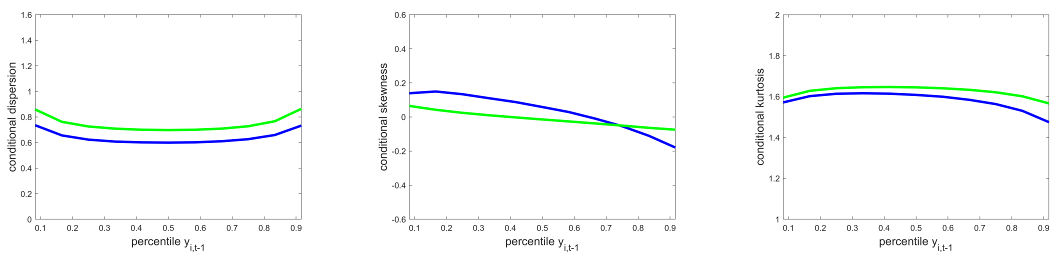


Figure C.4: High-educated households: conditional dispersion, conditional skewness and conditional kurtosis of log earnings residuals $y_{i,t}$. Estimates from SHIW data (blue line) and simulation from the estimated nonlinear model (green line).

(a) Conditional dispersion (b) Conditional skewness (c) Conditional kurtosis



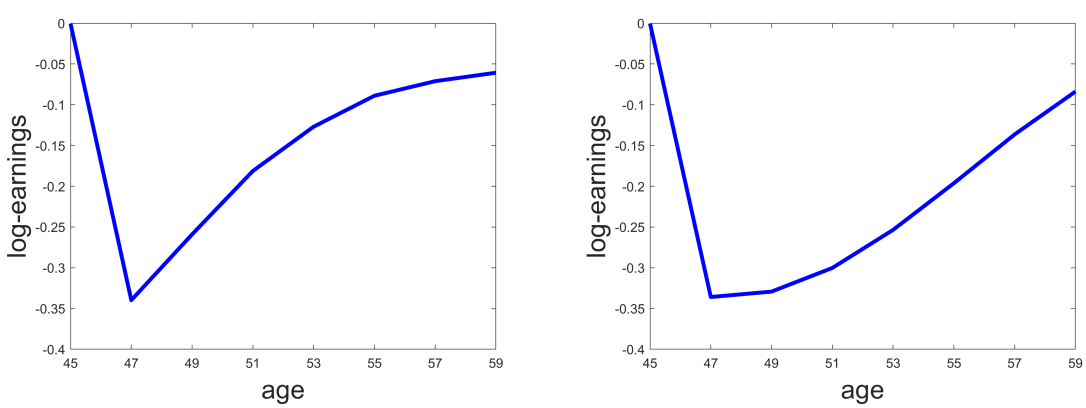
Note: Conditional dispersion is given by $\sigma(y;\tau) = Q(\tau|y_{i,t-1} = y) - Q(1-\tau|y_{i,t-1} = y)$, where $\tau = 11/12$. Conditional skewness is $skew(y,\tau) = \frac{Q(\tau)+Q(1-\tau)-2Q(1/2)}{Q(\tau)-Q(1-\tau)}$. Conditional kurtosis is $kurt(y,\tau) = \frac{Q(1-\alpha)-Q(\alpha)}{Q(\omega)-Q(1-\omega)}$ where $\tau = 11/12$, $\omega = 10/12$ and $\alpha = 1/12$.

D Simulation exercise

There are many possible ways in which this alternative simulation exercise on *levels* can be set up. I explore a different approach by fixing the initial conditions to be the same in the two groups as before ($\tau_{init} = 0.9$ and $\tau_{init} = 0.88$ respectively) but keeping the group specific median shock ($\tau_{shock} = 0.5$) and varying the positive/negative shocks to get the same percentage change in earnings. In this alternative exercise $\tau_{shock} = 0.1$ ($\tau_{shock} = 0.9$) for the low educated corresponds to $\tau_{shock} = 0.041$ ($\tau_{shock} = 0.933$) for the high educated. Results are shown in Figure D.5 and D.6. Both methodologies highlight the same patterns and bring to the same conclusions.

Figure D.5: Effect of imposing to the high educated the same level of initial condition and negative percentage change in earnings of the low educated.
Note: the figure shows the difference between age-specific medians of log earnings of a household hit by a large negative shock at age 47, $\tau_{shock} = 0.1$ for the LE and 0.041 for the HE, and a household hit by a group specific median shock; τ_{init} at age 45 is 0.9 for the LE and 0.88 for the HE (LE=low educated, HE=high educated).

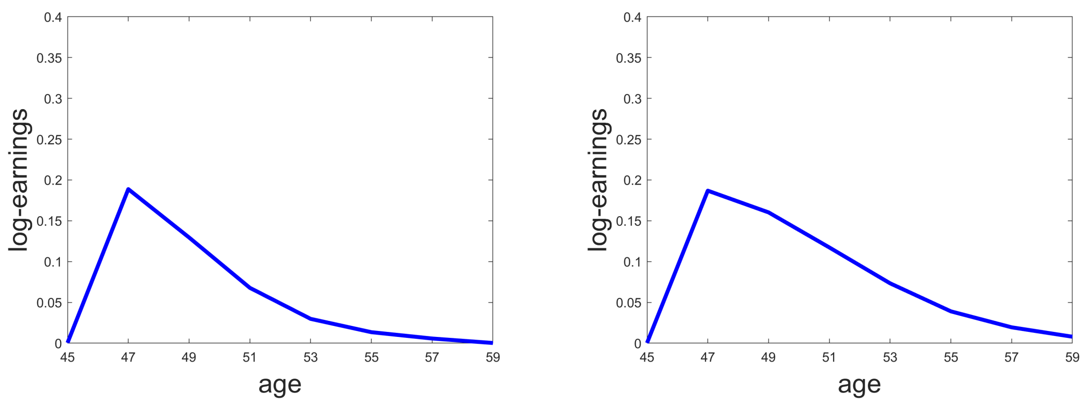
(a) Low educated (b) High educated



Our previous results are confirmed for positive shocks as $\tau_{shock} = 0.9$ for the low educated corresponds to a very similar percentile for the high educated ($\tau_{shock} = 0.89$). When considering negative shocks, we instead find that high educated households show a slower recovery than low educated ones. However, this is because we

Figure D.6: Effect of imposing to the high educated the same level of initial condition and positive percentage change in earnings of the low educated.
Note: the figure shows the difference between age-specific medians of log earnings of a household hit by a large positive shock at age 47, $\tau_{shock} = 0.9$ for the LE and 0.933 for the HE, and a household hit by a group specific median shock; τ_{init} at age 45 is 0.9 for the LE and 0.88 for the HE (LE=low educated, HE=high educated).

(a) Low educated (b) High educated



are comparing shocks of the same magnitude but quite different probability of occurrence ($\tau_{shock} = 0.1$ for the low educated corresponds to $\tau_{shock} = 0.03$ for the high educated).

Chapter 2

Household risk preferences and portfolio allocation: a collective approach

2.1 Introduction

The workflow of the household decision-making process is of core interest in economics, and understanding its mechanisms might spread new light on consumers behaviours and choices. In this contest, household finances are of particular interest because of their potential impacts on households present and future economic status. For example, the responsibility of saving for retirement falls largely into individuals' hands, as well as the decision of pension schemes or the choice of credit cards and bank accounts. Household portfolio allocation gathers large part of these financial choices and was broadly studied over past years. The standard models used to study portfolio choices predict that each household holds a fraction of its wealth in risky assets if the equity premium is positive (e.g.: Samuelson (1975) and Merton (1969)). These models rely on the so-called unitary approach, which considers the household as a unique decision unit that behaves as a single agent with well-defined preferences.

Pierre-André Chiappori (1988) shades new light on the household decision-making process, introducing the so-called collective model. In this model, households behave as a multi-dimension system of several members, which may show different preferences. An intrahousehold bargaining process is assumed to take place and to drive the final choice among the household members, combining their preferences. The collective model is largely used in studies concerning household labour supply (e.g.: Pierre-André Chiappori (1988), Pierre-André Chiappori (1992)), consumption choices (e.g.: Cherchye, De Rock, and Vermeulen (2007), Leeuwen, Alessie, and Bresser (2020)) and household production decisions (Apps and Rees (1997)), but to date only a few studies household portfolio allocation under this perspective, as Gomes, Haliassos, and Ramadorai (2020) highlight. Among the few studies that adopt a collective approach to investigate household financial decisions, Addoum, Kung, and Morales (2016) study the connection between marital decisions, consumption, and household investments. They show that changes in marital status or spouses' relative income imply a significant reallocation of the household portfolio. Olafsson and Thornquist (forthcoming) use the potential earnings of spouses, instead of actual earnings, as a proxy of the household decision-makers bargaining powers. In line with Addoum, Kung, and Morales (2016), they show that the higher is the weight of the wife, the lower is the household probability of holding equity. In other words, if the female partner has higher decision power, the household portfolio is less risky. Last, Gu, Peng, and Weilong Zhang (2021) investigate the gender gap of bargaining power in the household portfolio decision making process. Their results show that the household portfolios reflect the preferences of the male partner 44% more than the female partner characteristics, with gender norms that play a relevant role in explaining this large difference.

In this paper, I assume that a bargaining process takes place between the two partners and drives the household portfolio choice. In this decision process, household members decide about stock market participation and optimal wealth allocation simultaneously. The bargaining process mainly concerns the household risk prefer-

ences, that play a crucial role in portfolio choices as largely documented in the literature (e.g.: Weiwei Zhang (2017)). The influence of risk aversion is a natural consequence of the uncertainty and the volatility that characterizes financial markets.
This leads to the following questions: how are financial choices within a household being made? In married/co-living couples, which partner influences more the financial decision? If the household portfolio choice reflects partners' risk preferences, how these preferences combine and determine the portfolio allocation?

I derive a model of household portfolio allocation in which partners decide first about household risk preferences and then about stock market participation and optimal allocation of financial resources. Then, I study the effect of risk preferences on portfolio allocation using the English Longitudinal Study of Ageing (ELSA) panel dataset. Finally, I compare the standard unitary approach with the proposed collective approach. Results show that the collective model fits significantly better the data, and that partners risk tolerance increases the share of wealth allocated in risky assets but does not affect household stock market participation.

This paper contributes to two main fields of economic research: household portfolio choice models and risk preferences in group decision.
First, this paper contributes to the household portfolio choice models with limited stock market participation (Gomes and Michaelides (2005); Wachter and Yogo (2010)). Generally, researchers treated the household as a single decision-making unit and explain the large stock market non-participation rate with stock market participation costs. The literature refers to these models as unitary models and identifies the preferences of the decision-making unit with the male partner or the household head preferences. About costs, Vissing-Jorgensen (2002) introduces three types of participation costs: fixed or lump-sum entry costs, variable transaction costs and per period trading costs. My model departs from the standard unitary household assumption and follows the collective approach introduced by Pierre-André Chiappori (1988). The collective model was used to study households labour supply and consumption

decisions, while portfolio allocation receives little attention, as Pierre-Andre Chiappori and Mazzocco (2017) reported.
I introduce a portfolio decision model that includes the household risk preferences measured as a weighted average of partners risk preferences. The weights can be interpreted as each agent bargaining power. Therefore, the partner who holds the "purse strings" (Bertocchi, Brunetti, and Torricelli (2014)) would have a higher decision power and influence in the determination of the household risk tolerance.
Second, I investigate the effect of group risk preferences in the household portfolio allocation process, that is a particular case of group decisions. Through the last decade, many studies investigate how risk preferences differ when agents act as groups and individually. De Palma, Picard, and Ziegelmeyer (2011) studies the aggregation of preferences of spouses concluding that their decision-making process is dynamic. At the beginning of the experiments, the male partners show more decision power, while this effect gradually decreases as time passes by among the game. Abdellaoui, l'Haridon, and Paraschiv (2013) use certainty equivalent methods to derive time and risk preferences of couples, assuming prospect theory. They study the decision-making process separately for individuals and couples, revealing that the probabilistic risk attitudes of single agents and couples showed similar judgmental biases. They also show that couples risk attitudes are a combination of spouses' preferences and that the correlations between risk attitudes of couple members are weak, but significant. Charness and Sutter (2012) show that groups (composed by stranger) are more rational than individuals, and their behaviours are in line with game-theoretic prediction. In other words, groups exhibit less behavioural biases than individuals. The studies mentioned so far highlight the characteristics of the group risk preferences, while this paper focuses on how those preferences combine and affect the outcome of the portfolio decision.

The rest of the paper is organized as follow: Section 2.2 introduces the collective household model for household's portfolio allocation. Section 3.2 describes the ELSA dataset, Section 2.4 presents the empirical analysis and Section 2.5 concludes.

2.2 Theoretical model

This section introduces the theoretical model that studies the heterogeneity of household portfolio choice considering the risk preferences of both household decision-makers. The model assumes that the household members decide first about the household risk preferences, and then about the optimal portfolio allocation. I use a mean-variance utility function[1] where the household risk tolerance is measured as a weighted average of partners risk preferences.

2.2.1 Collective model: household utility and weighted risk aversion

Assume that the economy has only two assets: a risk-free asset (treasury bills) and a risky asset (representative of the stock market). The two assets have different expected returns: the risk-free asset has certain returns r while the risky asset has returns $\tilde{r} = r + \tilde{s}$, where $\tilde{s} \sim N(\mu_s, \sigma_s^2)$. Each household h has two decision makers (partners), agents a and b, and holds the initial wealth w. The household wants to maximize its utility $u(W_h)$, where W_h represents the expected household wealth after assets returns. The crucial decision is about α, that identifies the proportion of wealth w allocated in risky assets.

The household wealth W_h is:

$$W_h = w[(1 - \alpha)(1 + r) + \alpha(1 + \tilde{r})] = w[(1 + r) + \alpha\tilde{s}] \tag{2.1}$$

Thus, the household utility maximization program can be written in terms of the value function V_h as a function of α:

$$\max_{\alpha} V_h(\alpha) = \max_{\alpha}\ u(W_h) \ = \ \max_{\alpha}\ \ u(w[(1 + r) + \alpha\tilde{s}]) \tag{2.2}$$

[1]Mean-variance utility is equivalent to CARA utility, as I show in Section A.1. Section A.2 of the Appendix provides a second version of the model where household utility is a weighted sum of partners utility .

Assuming CARA (exponential) utility function of the form $u(z) = -e^{-\rho z}$, where ρ is the absolute risk aversion coefficient, household utility becomes:

$$u(W_h) = -e^{-\rho_h W_h} \tag{2.3}$$

where ρ_h identifies the household risk preference. ρ_h is a weighted sum of the partners risk preferences, where the weights represents the bargaining power of each partner in the decision making process. Therefore, household risk aversion is:

$$\rho_h = \mu_a \rho_a + \mu_b \rho_b \tag{2.4}$$

where $\rho_{a,b}$ are the risk preferences of the household decision makers, and $\mu_{a,b}$ are the bargaining powers. I normalize the weights as follow:

$$\gamma = \frac{\mu_a}{\mu_a + \mu_b} \qquad (1-\gamma) = \frac{\mu_b}{\mu_a + \mu_b} \tag{2.5}$$

such that $\gamma \in [0, 1]$ and ρ_h becomes:

$$\rho_h = \gamma \rho_a + (1-\gamma)\rho_b \tag{2.6}$$

Thus, $V_h(\alpha)$ in Equation 2.2 becomes:

$$\begin{aligned} \max_{\alpha} V_h(\alpha) &= \max_{\alpha} E[u(W_h)] = \max_{\alpha} E[-e^{-\rho_h W_h}] \\ &= \max_{\alpha} \ \rho_h w[(1+r) + \alpha\mu_s - \frac{1}{2}\alpha^2 \sigma_s^2 \rho_h w] \end{aligned} \tag{2.7}$$

Solving the first order condition for α, the optimal share of wealth allocated in risky assets is:

$$\alpha = \frac{\mu_s}{\sigma_s^2 w \rho_h} \tag{2.8}$$

The optimal α is proportional to the risk premium μ_s and decreasing in the variance

(risk) of returns and in household risk aversion.
Under the condition of Equation 2.8 and with the strong assumption that all households have the same information about stock market returns, the heterogeneity in households' portfolio depends on wealth and risk preferences[2].

2.2.2 Preference shifter

The solution proposed in Equation 2.8 implies that the heterogeneity in α depends entirely on risk preferences and household wealth, assuming that individuals have common priors (i.e., they all experience the same stock expected returns and return variance). However, it is unlikely that household demographics such as education, income or age do not affect household portfolio decision.
In what follows, I allow household characteristics to affect the household portfolio choice process through risk preferences. Risk preferences now are a function of household and partners characteristics, $\rho_h = \rho_h(z)$, where $z = \beta x + \epsilon$. x is a vector of household demographics (e.g.: ages, education, income), β is a parameter matrix and ϵ is the i.i.d. error term that represents the unobserved variation in taste shifts across households. Therefore, the maximization problem in Equation 2.7 becomes:

$$\max_{\alpha} V_h(\alpha) = \max_{\alpha} \, [\rho_h(z)]w[(1+r) + \alpha\mu_s - (\frac{1}{2}\alpha^2\sigma_s^2 w[\rho_h(z)])] \tag{2.9}$$

and the optimal α is:

$$\alpha = \frac{\mu_s}{\sigma_s^2 w[\rho_h(z)]} \tag{2.10}$$

Equation 2.8 and Equation 2.10 show that the optimal proportion of wealth al-

[2]The assumption about homogeneous expectation allows us to infer the implied degree of absolute risk aversion of the stockholders, i.e. only for those households whose share of wealth allocated in risky assets is observed.

located in risky assets depends on risk preference, stock market returns and wealth itself. However, the former implies that household or individual demographics do not affect portfolio choices, while a large amount of literature shows that they influence household finances. Education (Cooper and Zhu (2016), Poterba, Venti, and Wise (2013)), health (Poterba, Venti, and Wise (2013)), age (Ameriks and Zeldes (2004), Bertocchi, Brunetti, and Torricelli (2014)), wealth (Wachter and Yogo (2010)) and financial literacy (Jappelli and Padula (2015), Lusardi (2008)) are only some of the determinants of household portfolio decision. Thus, non including them in the model may produce severe bias in the estimation.
Equation 2.10 shows that the optimal share of wealth allocated in risky assets is always positive. In other words, if the risk premium μ_s is positive, every households invests a fraction of its financial wealth in the risky asset, as in Samuelson (1975) and Merton (1969). However, a large fraction of households does not hold stocks: this is the so-called stock holding puzzle (e.g.: Guiso, Haliassos, and Jappelli (2003) describe and discuss this issue across European countries).

2.2.3 Introducing stock market participation costs

One of the possible explanations of the stock holding puzzle are stock market participation costs. They reduce risky assets expected returns and increase the probability of losses. Therefore, depending on the amount that they have to pay, the households may decide to not participate. Vissing-Jorgensen (2002) identifies three main types of costs that affect the stock market participation choice: fixed entry costs (learning about financial markets), variable transaction costs (trading fees or bid-ask spread) and per period trading costs (broker subscription or bank fees). Working with US data, Vissing-Jorgensen (2002) estimates that a relatively low per period cost (50$ per month) explains the non participation of half of non-stockholders.

Including costs in the portfolio choice process means that the household utility maximization problem has now two steps: the first concerns the stock market participation decision (i.e.: $\alpha > 0$ or $\alpha = 0$) and the second solves the utility maximization problem, following Equation 2.10. In case of costs, the expected household wealth after assets returns W_h is described by the following framework:

$$\begin{aligned} \alpha = 0 \rightarrow & W_h = W_{hs} = w[(1+r)] \\ \alpha > 0 \rightarrow & W_h = W_{hr} = w[(1-\alpha)(1+r) + \alpha(1+r+\tilde{s})] - C \end{aligned} \tag{2.11}$$

where C are the stock market fixed entry costs payed at the end of the period. The household evaluates whether holding risky assets is convenient or not, considering that in case of stock market participation it has to pay the lump sum cost C. Then, the optimal amount allocated in risky assets α solves:

$$E[u'(w((1+r) + \alpha\tilde{s}) - C) \cdot w(\tilde{s})] = 0 \tag{2.12}$$

and $\alpha = 0$ is a solution of the maximization problem if and only if the equity premium is 0.
I define the certainty equivalent of the risk premium $\tilde{s}$ as follow:

$$E[u(w((1+r) + \alpha\tilde{s}) - C)] = u[w((1+r) + \alpha\hat{s}) - C] \tag{2.13}$$

where $\hat{s}$ represents the risk adjusted equity premium. Therefore, the household evaluates:

$$\begin{aligned} & \mathbf{E}[u(W_{hr})] > u(W_{hs}) \rightarrow \mathbf{E}[u(w((1+r) + \alpha\tilde{s}) - C)] > u(w(1+r)) \\ & \rightarrow w(1+r) + w\alpha\hat{s} - C > w(1+r) \\ & \rightarrow w > \frac{C}{\alpha\hat{s}} \end{aligned} \tag{2.14}$$

The condition derived in Equation 2.14 defines a threshold of minimum wealth for potential investors that is proportional to fixed costs C, optimal share of wealth allocated in risky assets, α, and the risk adjusted equity premium $\hat{s}$:

$$\bar{w} = \frac{C}{\alpha \hat{s}} \tag{2.15}$$

Then, when the household initial endowment of wealth is less than $\bar{w}$, non-participation is the optimal choice, otherwise the household invests α share of its financial wealth in risky asset.

2.3 ELSA data

ELSA is a longitudinal survey that collects data from a representative sample of English people aged 50+. It is a biennial survey (first wave in 2002) that aims to gather data to study all the problems and aspects of ageing, like social care, retirement, pension policies and social participation. The original sample of ELSA (first wave) was selected from the Health Survey for England (HSE[3]) respondents in the period 1998-2001. After the first survey in 2002, younger age groups are refreshed to balance the panel over time.

This paper works with Wave 8 of ELSA, which collects data about 8445 individuals, interviewed between May 2016 and June 2017. Researchers introduced a series of new and innovative measures that have broadened the scope of the study. Among the new questions, Wave 8 includes three self-assessed measures of risk preferences: one related to the general propensity to take risks, one to financial risk taking and one to the health domain. The purpose of this paper is to study the household portfolio decision process and these questions are of particular interest because of the crucial role of risk tolerance in financial choices. Several papers point out that these qualitative questions predict behaviour across various domains (Caliendo, Fossen, and Kritikos (2009), Fouarge, Kriechel, and Dohmen (2014)), including risk preferences, when experimental data (like lottery choices) are not available.

In ELSA, the participants answer to the following general risk tolerance question:

[3] More information about HSE at `http://healthsurvey.hscic.gov.uk`.

Are you generally a person who is fully prepared to take risk, or do you try to avoid taking risks?
The respondent chooses an integer between 0 (Avoid taking risks) and 10 (Fully prepared to take risks)[4]. The predefined structure of the answers implies that they return a self-assessed measure of risk tolerance, rather than risk aversion. These measures should be positively correlated with α according to Equation 2.10, where the risk measure, ρ_h, represents relative risk aversion.

The survey provides a second question, related to respondents patience:
Are you generally an impatient person, or someone who always shows great patience?
The respondent chooses an integer between 0 (Very impatient) and 10 (Very patient)[5].

The correlation between partners' risk preferences (both general and financial) is always positive and significant, but relatively weak (coefficient vary between 7.7% and 17.7%, depending on the specific item considered). The highest correlation is the one between the financial risk tolerance of husband and wife, that may be a signal of assortative matching of partners.

2.3.1 Sample selection and description

The survey distinguishes between three financial unit categories: singles, couples with separate finances and couples with joint finances. I use the data about individuals who are in a *couple with joint finances* aged less than 90 years. I select the male-female couples that have positive income (labour and pension income, including state benefit transfers), non-negative net financial wealth (the sum of savings and investments, subtracting financial debts from credit cards, overdrafts and other private debts but not mortgages) and share of net financial wealth allocated in risky assets between 0 and 1 (computed as the ratio between the amount of risky assets and the

[4]Financial risk tolerance question is: *Thinking specifically about your finances, spending and savings, are you a person who is fully prepared to take risk, or do you try to avoid taking risks?*

[5]Financial patience question is: *Thinking specifically about your finances, spending and savings, are you generally an impatient person, or someone who always shows great patience?*

household net financial wealth). The final sample is composed of 1441 male-female couples, i.e. 2882 individuals[6].

Table 2.1 presents the households' basic demographics, while Tables 2.2 and 2.3 show the sample statistics of partners. Table 2.2 distinguishes partners by gender, while Table 2.3 distinguishes by financial respondent (and non-financial respondent) partners. The financial respondent is the partner that answers to the Income & Asset section of the ELSA survey. Note that the selected households have *joint finances*: in these cases the ELSA interviewer asks financial information only to one of the two spouses (financial respondent) and her/his answers are copied in the survey of the partner (non-financial respondent).

Table 2.1: ELSA Wave 8 summary statistics: *couples with joint finances*

	All	Non-stockholders	Stockholders
hh obs	1441	371	1070
financial respondent: male	60.8%	54.9%	62.8%
hh income: mean (weekly £)	649.5	533.3	689.7
hh income: median (weekly £)	572.9	459.8	615.5
hh income: std (weekly £)	408.5	328.4	425.5
net financial wealth: mean (thousand £)	152.8	44.6	190.2
net financial wealth: median (thousand £)	66.9	13.0	101.0
net financial wealth: std (thousand £)	267.5	179.8	282.4
gross financial wealth: mean (thousand £)	153.7	45.3	191.1
gross financial wealth: median (thousand £)	67.1	15.0	103.0
gross financial wealth: std (thousand £)	267.4	179.8	282.2
stock share of financial wealth	32.6%	-	43.9%
share of hh income : male (γ)	66.6%	66.2%	66.7%
share of hh income : respondent (γ)	56.3%	54.9%	56.8%

Table 2.1 shows that around 75% of the households hold risky assets, where risky assets are defined as shares, bonds, stocks and shares ISAs or life insurance ISAs[7].

[6] Table 2.1 in the Appendix shows the sample selection procedure and the correspondent number of observations.

[7] ISA (Individual Saving Account) is a class of retail investment arrangement available to residents of the United Kingdom, with favorable tax condition. They offer four types of account: cash ISA, stocks & shares ISA, innovative finance ISA (IFISA) and lifetime ISA.

Labour and pension income are higher among stockholders, as well as net and gross household financial wealth. This is in line with the fixed entry costs assumption of Section 2.2.3: the higher is the household wealth, the lower is the impact of fixed entry costs on the portfolio returns, the higher is the probability of participation. Let's assume that there are two households with the same risk preferences and demographics, i.e., the households have the same optimal α, but different financial wealth. The one with a higher financial wealth invests a higher amount of money in risky assets and obtains higher returns (in absolute terms), with a lower impact of fixed costs on its finances. The last two rows of the table show the share of household labour and pension income of the male partner and the financial respondent partner. On average, males contribute more to household income, with no difference between stockholders and non-participant. The financial respondent also generally holds a larger share of the household income, but the difference between the two partners is now ten percentage points lower. Note that around 40% of the financial respondent are females, which have lower salary/pension on average.

Table 2.2 shows partner characteristics by gender. There is a significant difference in the education level of partners comparing participants and non-participants. Highly educated partners are almost one-third among stockholders, while only one over ten non-stockholders completed college/university. The demographics show a second relevant difference, in line with the literature concerning risk preferences: males have higher risk tolerance than females. Wives show lower risk tolerance both when comparing stockholders and non-stockholders partners. These differences are persistent across the two types of risk tolerance, general and financial. On the other hand, there are no significant differences in patience between males and females and between stockholders and non-stockholders. Table 2.3 shows that the differences in risk tolerance between financial respondent and non-financial respondent are lower than that between male and female partners, especially when financial risk preferences are considered. As mentioned above, this attenuation derives from the fact that 40% of the financial respondents are females, which show a lower risk tolerance. However, risk tolerance is still higher among the stockholders. As in Table 2.2, there are no differences in financial and general patience scores between financial respondent and non-financial respondent and between stockholders and non-stockholders. Last, there are no changes in the proportion of highly educated and low educated individuals between stockholders and non-stockholders, confirming the pattern of Table 2.2.

Table 2.2: ELSA Wave 8 summary statistics: male and female partners

	All	Non-stockholders	Stockholders
age: male	68.5	68.4	68.6
age: female	66.2	66.3	66.1
low edu: male	32.5%	50.7%	26.2%
mid edu: male	41.9%	37.7%	43.3%
high edu: male	25.6%	11.6%	30.4%
low edu: female	29.9%	43.9%	25.0%
mid edu: female	48.5%	46.4%	49.2%
high edu: female	21.6%	9.7%	25.8%
general risk: male	5.0	4.8	5.1
general risk: female	4.2	4.1	4.3
general patience: male	6.2	6.4	6.2
general patience: female	6.8	6.8	6.7
financial risk: male	3.6	3.2	3.8
financial risk: female	2.9	2.8	2.9
financial patience: male	6.9	6.9	6.9
financial patience: female	7.0	6.8	7.1

Table 2.4 gives an overview of the selected households portfolio allocation. Total household wealth is increasing in household income and is higher among stockholders. Comparing the first three quartiles of the income distribution, the wealth of stockholders is almost twice the wealth of non-stockholders. Last, the households with higher income are generally younger. On average, housing accounts for more than 70% and less than 65% of wealth for non- and stockholders, respectively, with five percentage points of difference between the two categories. Risky assets account for 10-15% of the household wealth, and their share is independent of income and

Table 2.3: ELSA Wave 8 summary statistics: financial respondent and non respondent partners

	All	Non-stockholders	Stockholders
financial respondent: female	39.2%	45.0%	37.2%
age: respondent	67.5	67.7	67.4
age: non respondent	67.2	67.0	67.2
low edu: respondent	29.2%	47.2%	23.0%
mid edu: respondent	45.1%	41.0%	46.6%
high edu: respondent	25.6%	11.9%	30.4%
low edu: non respondent	33.2%	47.4%	28.3%
mid edu: non respondent	45.2%	43.1%	46.0%
high edu: non respondent	21.6%	9.4%	25.8%
general risk: respondent	4.8	4.7	4.8
general risk: non respondent	4.5	4.3	4.6
general patience: respondent	6.4	6.6	6.4
general patience: non respondent	6.5	6.6	6.5
financial risk: respondent	3.3	3.0	3.5
financial risk: non respondent	3.1	3.0	3.1
financial patience: respondent	7.0	6.9	7.1
financial patience: non respondent	6.9	6.7	7.0

rather stable among stockholders portfolio. This pattern is in line with the entry costs hypothesis presented in Section 2.2.3: comparing two households with the same optimal α and different financial wealth, the household with lower financial wealth invests a lower amount, therefore the potential financial returns may not cover the fixed entry costs and the decison-makers decide to not hold risky assets.

These statistics are in line with the findings of Banks and Smith (2000). They study the evolution of English households portfolio composition between 1980 and 2000, working with the Family Expenditure Survey (FES) and the Financial Research Survey (FRS) data. They show that housing and pensions funds account for the largest share of the household portfolio, with a progressive shift from housing towards financial assets over time. This shift was the consequence of tax-favoured products (TESSAs, replaced by ISAs in 1999) created by the government to try to encourage pension savings during the 90s.

Table 2.4: Composition of total gross household wealth by stock market participation and household labour and pension income quartile[a].

stock market participation	income quartile	age male	obs	wealth (thousands £)	housing (in wealth %)	safe assets (in wealth %)	risky assets (in wealth %)	physical wealth (in wealth %)	debt (£)	mortgage (thousands £)
No	1^{st}	71.0	140	255	68.7	27.6	-	4.51	462	1.9
	2^{nd}	68.8	108	223	70.5	32.3	-	2.47	626	5.0
	3^{rd}	66.0	67	323	76.2	21.2	-	3.28	1002	8.8
	4^{th}	63.9	50	680	73.8	17.1	-	9.46	1196	45.4
Yes	1^{st}	69.8	217	514	63.4	15.1	13.4	8.34	485	6.4
	2^{nd}	69.5	254	461	66.6	15.8	11.7	6.13	493	4.0
	3^{rd}	68.9	290	608	65.4	14.9	14.1	6.04	1070	4.2
	4^{th}	66.7	313	903	60.1	14.1	15.5	10.7	1445	17.4

[a] *wealth* is net total household wealth, *housing* is gross housing wealth (the value of owner occupied primary housing before mortgage debt is subtracted), *safe assets* are money invested in "safe" assets such as bank accounts, savings accounts and cash ISAs, *risky assets* are money invested in "risky" assets such as shares, bonds, stocks and shares ISAs or life insurance ISAs, *physical wealth* represents alternative investments (second homes, farm or business property, works of art etc), *debt* is credit cards, overdrafts, other private debt but not mortgages.

2.3.2 Weights

The model in Section 2.2 defines ρ_h as a weighted household risk tolerance, where weights are the normalized bargaining powers of partners, γ and $(1-\gamma)$ respectively. Measuring the bargaining powers of household members is one of the main issues of collective models. These powers may change relative to the kind of decision that the household is taking (e.g., the decision-making processes of financial choices may differ from other consumption decision processes like grocery or clothes).

The first study that investigates the within household allocation sharing rule and its determinants is Browning et al. (1994). They show that the allocation of resources is proportional to the relative share of household income of partners. In other words, the income pooling hypothesis, i.e. that it is the household total income that matter for the decision outcomes, is not consistent with the data. The conclusion shows that it is the share of household income of each partner that affects the intrahousehold allocation of resources. Recently, Attanasio and Lechene (2014) study the within-household shift in bargaining power using data about the cash transfer program *Progresa* in rural Mexico. This state program generates a large variation in the wife's relative household income (about 20% of household total expenditures), explicitly changing the control of resources within the treated households. They show that this shift changes the balance of power within the couple, concluding that one of the determinants of the within-household bargaining power is the share of the current income of each partner.

I follow these results to construct a measure of partners bargaining powers. The couples of the sample manage their finances jointly and have a unique, shared household wealth. I assume that the bargaining power of each partner is proportional to their own share of household income. This assumption implies that the income insurance that the high-income partner receives is lower than the one that he/she provides. To compensate this risk gap, the high-income partner has a higher control over household finances. Then, he/she can adjust the household risk, and consequently the household portfolio, to cover the partial income insurance that the partner can not provide. Concluding, I use the partners share of household income (labour, pension and state benefit income) as a proxy of the bargaining powers. The idea is that the higher is the wife/husband share of income, the higher is her/his control on the portfolio allocation decision.

Table 2.5 shows the income percentiles of ELSA partners by gender and by financial respondent[8]. Male and financial respondent partners earn more. However, males median income is twice the median income of females, while this difference reduces among financial and non-financial respondents, even if it remains large and significant (the gap is about 35% of non-respondent median income).

[8]Note that there is one household where the female reports negative earnings: in this case, I consider the male partner as the only income supplier with a share of household income equal to 1. Therefore, there are 74 one-income couples where males are the only income source, while only 73 females with no income.

Figures 2.3.1, 2.3.2 and 2.3.3 compare the box plots, the histograms and the kernel densities of wives and husbands weekly income distributions, measured in pounds. Both distributions are clearly non-normal and show positive skewness, with long right tails, and a concentration of the mass of the distribution on the left. Males income is higher (the husbands median income is two times the wives median income) and has a higher variability ($sd_m = 341.4$ £ against $sd_f = 202.4$ £ of wives) and longer right tails. Wage differences may depend on individual circumstances (e.g.: number of dependent children, company size and type of occupation) and reduce since the introduction of the Equal Pay Act in 1975 in the UK, but still affect the English society.

Table 2.5: Partners weekly income percentiles. Income is the sum of employment and self-employment income, private and state pension income, and state benefit transfers.

		Income percentiles				
	individual income = 0	10%	25%	50%	75%	90%
male	15	177 £	257 £	366 £	541 £	726 £
female	73	51 £	96 £	172 £	277 £	406 £
financial respondent	28	91 £	178 £	309 £	488 £	690 £
financial non-respondent	60	69 £	130 £	230 £	351 £	531 £

Table 2.6: Partners share of household income - percentiles. Income is the sum of employment and self-employment income, private and state pension income, and state benefit transfers.

		Share of hh income percentiles					
	share of hh income = 0	10%	25%	50%	75%	90%	share of hh income = 1
male	15	41.8%	55.4%	68.2%	80.3%	90.1%	74
financial respondent	28	20.9%	38.2%	58.4%	75.7%	87.8%	61

Table 2.6 shows the percentile of share of household income (γ) of males and financial respondent partners. On average, the male partner earns the 66.5% (sd:

Figure 2.3.1: Female and Male income - boxplot.

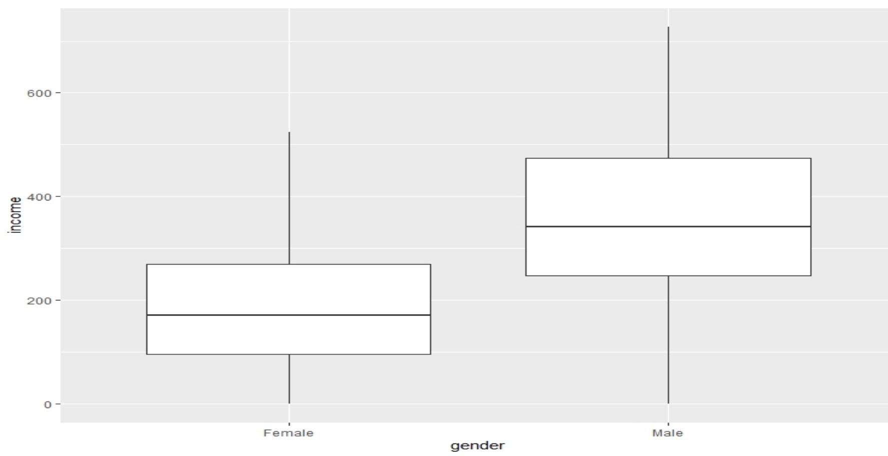
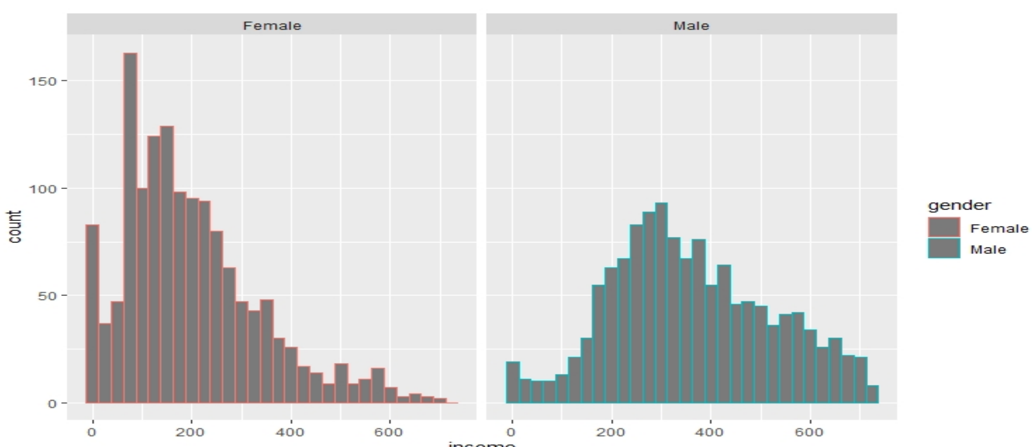
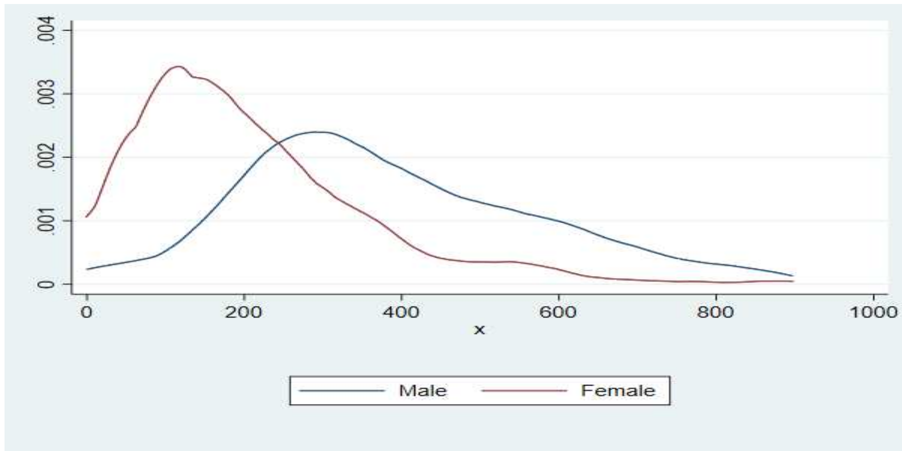


Figure 2.3.2: Female and Male income - histogram.



19.9%) of household labour and pension income, while the financial responding partner share is 56.3% (sd: 25.2%). The fact that the financial respondent has 10% less power depends mainly on the share of females financial respondents (40% circa), which earn a lower income than their partners, as Figure 2.3.1, 2.3.2 and 2.3.3 and Table 2.5 show.

Figure 2.3.3: Female and Male partners income - kernel densities.



2.4 Results

This Section analyses empirically the implication of the theoretical model described in Section 2.2.

First, I compare two econometric models that estimate household portfolio allocation using the general and the financial risk tolerance, respectively. This first step drives the choice of the baseline model, which uses the financial measure. Then, I discuss the results of the estimates, focusing on the determinants of household stock market participation and household share of wealth allocated in risky assets. Finally, I provide evidence that the collective model fits significantly better the data compared to the standard unitary model generally used in the literature.

This Section presents a reduce form analysis, that studies the effect of the collective risk tolerance on household portfolio allocation conditioning on a set of household demographics. The estimation of a structural model would allow studying how the demographics determine the household risk preferences and the role of the unobserved heterogeneity component of risk tolerance (see Section 2.2.2), however, this goes beyond the scope of this paper and is left for future work.

The empirical analysis does not consider distribution factors that are relevant elements of the collective models. Distribution factors are defined as exogenous conditions that may affect the bargaining power of the spouses without altering their preferences. Some examples are divorce law or other policies that explicitly change the distribution of resources within the household (e.g., see Attanasio and Lechene (2014)). Due to data limitations, it is not possible to identify proper distribution factors and study their effects in this analysis.

2.4.1 Heckman correction method

The analysis of the optimal portfolio allocation presents a possible problem of incidental truncation: the explanatory variables are always observed, while the dependent variable (the share of household wealth invested in risky assets, α) is available only for a subset of the population. In other words, I observe α only for those households that decide to invest in risky assets. Therefore, the rule determining whether α is observed or not does not depend directly on the outcome of α itself. Concluding, the truncation of the dependent variable is incidental because it depends on household decision to participate in the stock market.

I use the approach proposed by Heckman (1979) to estimate the effects of risk preferences on household portfolio allocation. This method allows the correction of bias from non-randomly selected samples or incidentally truncated dependent variables. First, it estimates the probability of observing the dependent variable using a probit model (selection equation) and then includes these results in the linear OLS estimation of the dependent variable (outcome equation). In other words, the second stage corrects for non-random selection by incorporating a transformation of the predicted probabilities of observing the dependent variable as an additional explanatory variable. This is the so-called Heckman two-step procedure.

The first stage is a probit model of the probability of observing the dependent variable:

$$Prob(s = 1|Z) = \Phi(Z\delta) \tag{2.16}$$

where s is a dummy variable that indicates whether the dependent variable (α, in this case) is observed or not, Z is the vector of the explanatory variables, δ is the vector of unknown parameters, and Φ is the cumulative distribution function of the standard normal distribution. This model is used to predict the probability of observing the dependent variable for each observation.
The second stage equation incorporates a transformation of these predicted probabilities as an additional explanatory variable to correct for self-selection. The outcome equation (second step) becomes:

$$\begin{aligned} y &= X\beta + u \\ E[y|X, s=1] &= X\beta + E[u|X, s=1] \\ E[y|X, s=1] &= X\beta + \rho\sigma_u\lambda(Z\delta) \end{aligned} \tag{2.17}$$

where the second row of Equation 2.17 highlights that the conditional expectation of y depends on the probability of observing it and the third row assumes that the error terms of Equations 2.16 and 2.17 are jointly normal. ρ represents the correlation between the two error terms, ϵ and u respectively, σ_u is the standard deviation of u and λ is the inverse Mills ratio estimated at the first step. Therefore, the Heckman selection model implies that incidental truncation is a form of omitted-variables bias.

In this analysis, the two dependent variables of the Heckman two-step procedure are household stock market participation and the share of wealth invested in risky assets, respectively. In other words, the first stage estimates the probability $Prob(\alpha) > 0$, while the second stage estimates α.

2.4.2 Exclusion restriction

An exclusion restriction is required for non-parametric identification. There must be at least one variable that appears with a non-zero coefficient in the selection equation but does not appear in the outcome equation: this variable is essentially an instrument.

In this case, the Z in Equation 2.16 includes age and age squared of the male partner, a dummy for large age difference in the couple (more than 10 years), household income and net total wealth quartiles, job market participation, a sickness index[9], education, risk tolerance, patience scores and numeracy score of both partners. The X of Equation 2.17 (second step) do not include partners' numeracy scores, which serves as exclusion restriction.

Numeracy is based on a set of questions about simple math exercises, like computing percentages, fractions, additions and subtractions. I use the 5 questions that ask to compute a sequence of subtractions: respondents have to subtract 7 from 100, and then 7 from the previous result and so on, five times. I compute individual numeracy score as the sum of correct answers of the respondent. I aggregate the scores from 2 to 4, obtaining a total of 4 possible categories for each partner: no numeracy skills (0 correct answers), low numeracy (1 correct answer), medium numeracy (2 to 4 correct answers) and high numeracy (5 correct answers) [10]. On average, the numeracy score is higher among stockholders. Moreover, males and financial respondents, i.e., those with higher bargaining power, show higher numeracy skills than females and non-financial respondents, respectively.

Numeracy approximates the individual cognitive skills. I assume that agents with low numeracy need more time to improve their financial literacy, increasing the stock market fixed entry costs. These higher costs decrease the potential returns of their investments. Then, those households with low numeracy may decide to not hold stocks. This assumption implies that all the households below a minimum threshold of numeracy do not participate in the stock market, while all the stockholders are above that threshold. Thus, numeracy affects household stock market participation, but the heterogeneity in α of those who participate does not depend on cognitive skills.

[9] It is constructed with principal component, using the numerous questions of ELSA related to participants health conditions. For more information see Dal Bianco (2020).

[10] Section B.1 in the Appendix shows the distribution of numeracy and numeracy categories.

2.4.3 Empirical estimates

This Section presents the empirical results of the paper. First, I focus on the selection of the collective model that serves as a benchmark through the analysis, then, I study the determinants of the household portfolio allocation and, last, I compare the collective and the unitary approach.

Baseline model selection

The ELSA survey provides two self-assessed measures of individual risk preferences. The former asks about risk preferences in general, while the latter asks about risk preferences in financial contexts (spending, savings).

Table 2.7 compares the outcome equation of the Heckman estimates of two specification that differ because of the type of risk tolerance used: Column (1) uses the financial risk tolerance, and Column (2) uses the general risk tolerance. Note that both risk measures are at the household level. Then, they are the weighted sum of partners risk preferences, where the weights are the share of household income of each spouse.

I use maximum likelihood estimation, which provides two evaluation metrics of the goodness of fit of the models, the information criteria Akaike (AIC) and Schwarz (BIC). Results are qualitatively the same and, based on AIC and BIC of Table 2.7, Column (1), i.e. the model that uses financial risk preferences, has to be preferred. Therefore, in what follows I focus on financial risk preferences and use Column (1) Table 2.7 as the baseline model.

Household portfolio allocation

This Section discusses the results of the empirical estimates that study the determinants of the household portfolio allocation decision.

I estimate two different collective specifications that capture the effects of risk preferences of both partners on household portfolio allocation. Table 2.8 presents the

Table 2.7: Heckman outcome equation: household share of net financial wealth allocated in risky assets. General vs financial risk tolerance measure. Partners characteristics by gender (selection equation is Table 2.6 in the Appendix).

share of risky assets	Financial risk measure (1)	General risk measure (2)
demographics[a]	*	*
financial risk: hh	0.0314***	
	(0.0053)	
financial patience: hh	0.0015	
	(0.0061)	
general risk: hh		0.0109**
		(0.0052)
general patience: hh		-0.0107*
		(0.0056)
Constant	0.1944	0.3804
	(0.6672)	(0.6740)
AIC	1,928	1,950
BIC	2,165	2,187
Number of observation	1,441	1,441

[a]Demographics include male age and age squared, dummy for large age difference between partners (more than 10 years), dummies of income and wealth quartiles, job market participation of partners, education of partners and the health index of partners.

results of the Heckman first stage: Column (1) uses the measure of husband and wife (financial) risk tolerance separately, while Column (2) uses the household weighted risk measures.

Income and wealth strongly affect the probability of stock market participation: income effects are stable across the income distribution, while wealth effects are increasing in magnitude. These findings are in line with the entry costs assumption in Section 2.2.3: the higher is the household wealth, the lower is the impact of lump-sum costs on household finances and the probability of being/becoming a stockholder increases. The numeracy of both partners is significant and have a positive effect on the stock market participation[11]. Male numeracy effects are increasing, while female numeracy does not show a similar pattern. This difference might be a consequence of the higher bargaining power of males: when the husband has high cognitive skills, he increase the probability of household stock market participation because of the higher (average) influence on portfolio choices. Male job market participation decreases the household probability of being a stockholder. Husbands may hold the "last say" on portfolio decision because of their higher (average) bargaining power. When the husband is working, he has less time to gather and study the necessary information about financial markets and then he decides to not invest in risky assets. Moreover, workers contributing to a Defined Contribution pension plan receive a lump-sum payment which amounts to about 25% of their pension pot when transiting to retirement. Therefore, it is more likely that they have higher liquidity to invest in the stock market at retirement. Male education has positive and significant effects, as largely documented in the literature. Last, there are no significant effects of financial risk tolerance on stock market participation, in line with the theoretical model. Indeed, risk tolerance must affect the share of wealth allocated in risky assets, but not participation, which depends on wealth, stock market entry costs and numeracy.

The outcome equation, in Table 2.9, estimates a linear model where the dependent variable is α (share of household net financial wealth allocated in risky assets) and the regressors are the same of the probit estimates of the first stage. However, the outcome equation excludes partners numeracy scores and includes the inverse Mills ratio of the first stage. The structure of Table 2.9 is the same of Table 2.8.
α increases in net financial wealth, with wealthier households that invest a larger share of their finances in risky assets. The female partner high education increases the portfolio share allocated in risky investments, while the male education has no effects. Notice that the opposite is true in the first stage. This effect might be due to

[11] I test the joint significance of the numeracy scores of the male and female partners in each model. The Wald tests reject the null hypothesis of joint non-significance at the 1% significance level.

partners sorting into marriage based on education. Therefore, couples participating in the stock market tend to have higher education on average. Once the stockholders are considered, what seems to matter is the high education level of the wife, because a woman with a college degree is rarer than men with the same title.

Risk tolerance positively affects the share of wealth allocated in risky assets in both cases, when the partners' or the household risk preferences is considered. The difference is in the marginal effects: male and female partner risk show similar coefficients, while the weighted household risk tolerance impact is twice the risk tolerance of the two partners.

The results are in line with the collective portfolio decision model, which describes the household financial decision process as a two-step procedure. In the first step, partners choose the shared degree of risk tolerance and then decide about stock market participation. In the second step, the choice concerns the optimal share of wealth allocated in risky assets, if any. Therefore, the household takes decisions as a system of individuals that combine their preferences, and not as a single unit. The specifications in Table 2.8 and Table 2.9 follow this idea: each of them includes the preferences of all the household decision-makers, combined in a unique (weighted) measure or not, and these preferences affect the household portfolio allocation. Last, the inverse Mills ratio presented in Table 2.9 is non-significant. However, as I will discuss in the robustness section (Section 2.4.4), controlling for selection remains important in this context.

Collective vs unitary approach

This Section aims to assess whether the collective approach proposed in this paper fits the data significantly better than the standard unitary approach generally used in the literature to study the household portfolio allocation.

As stated in the Introduction, the unitary model describes the household as a single decision-making unit that solves the utility maximization problem with well-defined preferences. Empirically, it is common practice to proxy the household behaviour using the husband or the head of the household. In this paper, I represent the unitary household with the risk preferences of the husband. In the Appendix, I present the result obtained using the household head as a proxy of the unitary household, i.e. with the preferences of the partner who hold the last say on the decision (e.g.: Bertocchi, Brunetti, and Torricelli (2014)). In this context, I proxy the household head using the financial respondent of the interview, i.e. the partner that answers to the Income & Asset section of the survey.

I compare three specifications, two collective models and one unitary model, using the information criteria and the likelihood ratio test. The selection and the outcome equation of each specification are jointly estimated using maximum likelihood, which provides the values of the AIC and BIC to compare the goodness of fit of the models. Last, the likelihood ratio test is constructed by considering the unitary model as a special case of the collective models. Then, the unitary model is the nested (or reduced) model of the test.

In line with the selected baseline specification of Section 2.4.3, the three estimated models use the financial risk preferences of households and partners. The three specifications share the same first-stage selection equation, which uses the numeracy of both partners as exclusion restriction, and includes the household demographics and the risk tolerances of each partner, separately. On the other hand, the specifications have three different outcome equations, whose estimation results are shown in Table 2.10. In particular, the first collective model (Column (2) Table 2.10) uses the risk tolerance of each partner separately (wife and husband), the second collective model (Column (3) Table 2.10) includes the household weighted risk tolerance and, additionally, the risk preference of the husband, while the unitary model (Column (4) Table 2.10) uses only the husband risk preferences (as a proxy of the household preferences). Therefore, the risk tolerance of the husband appears in each outcome equation, such that the unitary model (Column (4)) becomes a nested model of the two collective specifications (Columns (2) and (3)).
I compare the three specifications using the measures of goodness of fit, testing if

the risk preferences of wives add information to the collective models and matter in the household portfolio allocation process.

Table 2.10 shows the likelihood ratio test results between Column (2) and Column (4) and between Column (3) and Column (4), i.e. between the two collective models and the unitary model. The test rejects the null hypothesis in both cases. Thus, the additional variable used in Column (2) and (3), i.e. the risk tolerance of the wives, has an important role in explaining household portfolio allocation.

Last, Column (1) Table 2.10 reports the baseline model selected in Section 2.4.3[12]. Comparing the AIC and BIC information criteria of Column (1) and (4), I conclude that the collective model proposed in this paper fits better the data than the standard unitary approach.

2.4.4 Robustness check

The Heckman selection model has been shown to be sensitive to the choice of the exclusion restrictions. This Section provides the estimation of alternative models as a robustness check, changing the variables used as exclusion restrictions. In particular, Tables 2.11 and 2.12 show the estimates of two specifications whose exclusion restrictions rely on the numeracy of the male partner only. Overall, the coefficients are stable and consistent with the findings of Tables 2.8 and 2.9 show small changes. Therefore, I conclude that the results of Section 2.4.3 do not depend on the assumption made on the exclusion restrictions. Moreover, the robustness analysis presented in Table 2.12 shows that the inverse Mills ratio is significant, documenting the importance of controlling for truncated selection using the Heckman model.

Last, I perform the analysis using financial and non-financial respondent partners instead of male and female partners characteristics. In other words, I check the consistency of the results distinguishing partners and partners' weights and demographics by the household main financial decision-maker. Results are in Tables 2.10,

[12]It differs from Table 2.7 because it shares the selection equation with the specifications in Columns (2), (3) and (4) of Table 2.10, that use the risk preferences of each partner separately and not weighted households risk tolerance.

2.11 and 2.14 of the Appendix; they are consistent with the main findings of Section 2.4.3, but slightly attenuated.

Table 2.8: Heckman 1^{st} stage: household probability of holding risky assets. Partners characteristics by gender.

participation	Financial risk	
	Individual	Baseline
	(1)	(2)
low numeracy: male	0.4519**	0.4569**
	(0.2249)	(0.2247)
mid numeracy: male	0.6240***	0.6343***
	(0.2353)	(0.2351)
high numeracy: male	0.6539***	0.6541***
	(0.2118)	(0.2117)
low numeracy: female	0.4096**	0.4015**
	(0.1803)	(0.1802)
mid numeracy: female	0.3795*	0.3807**
	(0.1942)	(0.1942)
high numeracy: female	0.4082**	0.4073**
	(0.1712)	(0.1712)
age: male	0.1278**	0.1271**
	(0.0632)	(0.0634)
age^2: male	-0.0009*	-0.0009*
	(0.0005)	(0.0005)
age difference > 10	0.1215	0.1062
	(0.1659)	(0.1660)
2^{nd} hh income quartile	0.2439**	0.2440**
	(0.1066)	(0.1063)
3^{rd} hh income quartile	0.3470***	0.3385***
	(0.1176)	(0.1173)
4^{th} hh income quartile	0.3493***	0.3564***
	(0.1343)	(0.1341)
2^{nd} hh wealth quartile	0.7377***	0.7376***
	(0.1069)	(0.1068)
3^{rd} hh wealth quartile	0.8949***	0.8965***
	(0.1141)	(0.1138)
4^{th} hh wealth quartile	1.4268***	1.4228***
	(0.1448)	(0.1437)
in work: male	-0.2224**	-0.2301**
	(0.1123)	(0.1122)
in work: female	-0.0105	-0.0178
	(0.1084)	(0.1079)
mid education: male	0.3017***	0.3001***
	(0.0974)	(0.0974)
high education: male	0.3392**	0.3263**
	(0.1372)	(0.1368)
mid education: female	0.0582	0.0552
	(0.0952)	(0.0951)
high education: female	0.1452	0.1488
	(0.1455)	(0.1452)
health index: male	-0.0200	-0.0191
	(0.0190)	(0.0190)
health index: female	-0.0127	-0.0156
	(0.0182)	(0.0181)
financial risk: male	0.0177	
	(0.0172)	
financial risk: female	-0.0052	
	(0.0176)	
financial patience: male	-0.0281	
	(0.0196)	
financial patience: female	0.0259	
	(0.0183)	
financial risk: hh		0.0173
		(0.0205)
financial patience: hh		-0.0188
		(0.0236)
Constant	-6.1592***	-5.9883***
	(2.2244)	(2.2331)
Numeracy - joint significance test		
p-value	0.003	0.003
Number of observations	1,441	1,441

Table 2.9: Heckman 2^{nd} stage: household share of net financial wealth allocated in risky assets. Partners characteristics by gender (1^{st} stage is Table 2.8).

share of risky assets	Financial risk	
	Individual	Baseline
	(1)	(2)
age: male	0.0142	0.0140
	(0.0196)	(0.0197)
age^2: male	-0.0001	-0.0001
	(0.0001)	(0.0001)
age difference > 10	0.0238	0.0149
	(0.0410)	(0.0412)
2^{nd} hh income quartile	0.0082	0.0076
	(0.0316)	(0.0323)
3^{rd} hh income quartile	0.0114	0.0123
	(0.0338)	(0.0347)
4^{th} hh income quartile	-0.0008	-0.0010
	(0.0350)	(0.0362)
2^{nd} hh wealth quartile	0.0694	0.0756
	(0.0636)	(0.0658)
3^{rd} hh wealth quartile	0.1735**	0.1794**
	(0.0712)	(0.0739)
4^{th} hh wealth quartile	0.2413***	0.2467***
	(0.0875)	(0.0911)
in work: male	-0.0310	-0.0385
	(0.0276)	(0.0284)
in work: female	-0.0085	-0.0009
	(0.0254)	(0.0256)
mid education: male	0.0263	0.0282
	(0.0297)	(0.0301)
high education: male	0.0239	0.0245
	(0.0354)	(0.0356)
mid education: female	0.0339	0.0308
	(0.0248)	(0.0250)
high education: female	0.1045***	0.1067***
	(0.0322)	(0.0325)
health index: male	-0.0001	0.0003
	(0.0056)	(0.0056)
health index: female	-0.0029	-0.0025
	(0.0049)	(0.0049)
financial risk: male	0.0161***	
	(0.0043)	
financial risk: female	0.0186***	
	(0.0044)	
financial patience: male	-0.0017	
	(0.0048)	
financial patience: female	0.0017	
	(0.0048)	
financial risk: hh		0.0308***
		(0.0051)
financial patience: hh		-0.0004
		(0.0060)
Constant	-0.4977	-0.4875
	(0.7527)	(0.7597)
Inverse Mills ratio		
lambda	0.0720	0.0849
	(0.1203)	(0.1291)
Number of observations	1,441	1,441
Selection	1,070	1,070

Table 2.10: Heckman outcome equation: household share of net financial wealth allocated in risky assets. Collective vs unitary approach. Partners characteristics by gender (selection equation is Table 2.7 in the Appendix).

share of risky assets	Baseline model (1)	Collective models: Partner risk (2)	Collective models: Household risk (3)	Unitary model (4)
demographics[a]	*	*	*	*
financial risk: hh	0.0317***		0.0439***	
	(0.0052)		(0.0102)	
financial patience: hh	0.0015		0.0019	
	(0.0061)		(0.0109)	
financial risk: male		0.0158***	-0.0118	0.0192***
		(0.0044)	(0.0084)	(0.0044)
financial risk: female		0.0200***		
		(0.0046)		
financial patience: male		0.0010	0.0000	0.0006
		(0.0049)	(0.0088)	(0.0050)
financial patience: female		-0.0009		
		(0.0048)		
Constant	0.1802	0.1695	0.1721	0.2531
	(0.6666)	(0.6673)	(0.6665)	(0.6707)
likelihood ratio test		Col (2) and (4)	Col (3) and (4)	
p-value		0.0001	0.0001	
AIC	1,929			1,946
BIC	2,177			2,194
Number of observations	1,441	1,441	1,441	1,441

[a]Demographics include male age and age squared, dummy for large age difference between partners (> 10 years), dummies of income quartile, job market participation of partners, education of partners and the health index of partners.

Table 2.11: Heckman 1^{st} stage: household probability of holding risky assets. Partners characteristics by gender. Male numeracy is the only exclusion restriction.

participation	Financial risk	
	Individual	Baseline
	(1)	(2)
low numeracy: male	0.4956**	0.5015**
	(0.2228)	(0.2227)
mid numeracy: male	0.6659***	0.6777***
	(0.2331)	(0.2329)
high numeracy: male	0.7038***	0.7053***
	(0.2093)	(0.2092)
age: male	0.1286**	0.1276**
	(0.0632)	(0.0633)
age^2: male	-0.0009*	-0.0009*
	(0.0005)	(0.0005)
age difference > 10	0.1227	0.1071
	(0.1648)	(0.1649)
2^{nd} hh income quartile	0.2418**	0.2425**
	(0.1061)	(0.1059)
3^{rd} hh income quartile	0.3450***	0.3370***
	(0.1172)	(0.1170)
4^{th} hh income quartile	0.3478***	0.3553***
	(0.1340)	(0.1338)
2^{nd} hh wealth quartile	0.7530***	0.7530***
	(0.1065)	(0.1065)
3^{rd} hh wealth quartile	0.9071***	0.9094***
	(0.1137)	(0.1134)
4^{th} hh wealth quartile	1.4484***	1.4461***
	(0.1442)	(0.1431)
in work: male	-0.2073*	-0.2144*
	(0.1118)	(0.1118)
in work: female	-0.0002	-0.0090
	(0.1082)	(0.1076)
mid education: male	0.2867***	0.2850***
	(0.0968)	(0.0967)
high education: male	0.3269**	0.3120**
	(0.1363)	(0.1359)
mid education: female	0.0694	0.0671
	(0.0946)	(0.0945)
high education: female	0.1563	0.1602
	(0.1442)	(0.1439)
health index: male	-0.0187	-0.0178
	(0.0190)	(0.0189)
health index: female	-0.0183	-0.0213
	(0.0178)	(0.0177)
financial risk: male	0.0181	
	(0.0172)	
financial risk: female	-0.0079	
	(0.0175)	
financial patience: male	-0.0281	
	(0.0196)	
financial patience: female	0.0245	
	(0.0183)	
financial risk: hh		0.0162
		(0.0204)
financial patience: hh		-0.0203
		(0.0235)
Constant	-5.8757***	-5.7029**
	(2.2203)	(2.2273)
Number of observations	1,441	1,441

Table 2.12: Heckman 2^{nd} stage: household share of net financial wealth allocated in risky assets. Partners characteristics by gender. Male numeracy is the only exclusion restriction (1^{st} stage is Table 2.11).

share of risky assets	Financial risk	
	Individual	Baseline
	(1)	(2)
age: male	0.0239	0.0230
	(0.0205)	(0.0206)
age^2: male	-0.0002	-0.0001
	(0.0001)	(0.0001)
age difference > 10	0.0297	0.0205
	(0.0435)	(0.0432)
2^{nd} hh income quartile	0.0291	0.0252
	(0.0343)	(0.0343)
3^{rd} hh income quartile	0.0387	0.0351
	(0.0372)	(0.0371)
4^{th} hh income quartile	0.0254	0.0217
	(0.0385)	(0.0388)
2^{nd} hh wealth quartile	0.1446**	0.1418**
	(0.0697)	(0.0705)
3^{rd} hh wealth quartile	0.2594***	0.2560***
	(0.0782)	(0.0793)
4^{th} hh wealth quartile	0.3528***	0.3451***
	(0.0974)	(0.0984)
in work: male	-0.0440	-0.0494*
	(0.0301)	(0.0300)
in work: male	-0.0095	-0.0011
	(0.0272)	(0.0269)
mid education: male	0.0489	0.0477
	(0.0319)	(0.0319)
high education: female	0.0484	0.0447
	(0.0382)	(0.0378)
mid education: female	0.0378	0.0355
	(0.0263)	(0.0263)
high education: female	0.1110***	0.1143***
	(0.0346)	(0.0344)
health index: male	-0.0021	-0.0015
	(0.0059)	(0.0058)
health index: female	-0.0044	-0.0040
	(0.0052)	(0.0052)
financial risk: male	0.0172***	
	(0.0046)	
financial risk: female	0.0180***	
	(0.0047)	
financial patience: male	-0.0033	
	(0.0052)	
financial patience: female	0.0032	
	(0.0051)	
financial risk: hh		0.0318***
		(0.0054)
financial patience: hh		-0.0012
		(0.0063)
Constant	-1.0400	-0.9707
	(0.7990)	(0.7982)
Inverse Mills Ratio		
lambda	0.2470*	0.2358*
	(0.1381)	(0.1399)
Number of observations	1,441	1,441

2.5 Conclusion

This paper studies the role of partners risk preferences in the household portfolio choice process. I develop a theoretical model that describes the household portfolio allocation decision following the collective approach proposed by Pierre-André Chiappori (1988). This approach considers the households as groups of agents which combine their preferences through a bargaining process. Formally, the model assumes that the decision-making process is a weighted combination of individual behaviours, where the weights are the bargaining power of each household's member. Assuming exponential utility, the model shows how the optimal portfolio allocation depends on a weighted average of the household decision-makers risk preferences, fixed entry costs and stock market returns.

I use the ELSA survey to study the effects of the partners risk preferences on household portfolio allocation, using the share of household income as a proxy of the bargaining powers. The empirical estimate relies on the Heckman selection model, which corrects for the bias produced by non-randomly selected samples. Indeed, the household decision to participate or not in the stock market creates a problem of incidental truncation. In other words, the variable of interest is observed or not depending on the stock market participation decision. The exclusion restriction that guarantees non-parametric identification relies on partners' numeracy scores, which I consider a proxy of cognitive ability. I assume that low cognitive skills increase the stock market fixed entry costs because the household needs more time to learn about the stock market and its mechanism. Then, the partners decide to not hold risky assets because of lower potential returns. However, once the household has the minimum knowledge about the stock market (i.e., it has sufficiently high cognitive abilities) and holds stocks, numeracy does not affect the heterogeneity of the portfolio allocation.

The estimates show that stock market participation increases in household income, household wealth, partners numeracy and partners education, while risk tolerance has no effects. On the other hand, a higher risk tolerance increases the share

of wealth allocated in risky assets, once the household is a stockholder.
Finally, I compare the collective and the unitary approach estimating three different specifications with maximum likelihood. I use Akaike's and Schwarz's Bayesian information criteria and the likelihood ratio test to compare their goodness of fit. Results show that the collective approach performs significantly better than the unitary one, and therefore the preferences of both spouses matter in the household portfolio allocation process, when the partners manage their finances jointly.

Appendix
Household risk preferences and portfolio allocation: a collective approach

A Theoretical model

A.1 CARA and mean-variance utility

Household wealth w can be invested in two types of assets: a risk-free asset with constant return (r) and a risky asset with return $\tilde{r} = r + \tilde{s}$ where $\tilde{s} \sim N(\mu_s, \sigma_s^2)$. The household (or agent) i maximizes its own utility u_i, choosing the optimal share of wealth allocated in risky asset, α.

Wealth after the investments is:

$$W_i = (w - \alpha)(1 + r) + w\alpha(1 + \tilde{r}) = w(1 + r + \alpha\tilde{s})$$

Assuming CARA utility with risk aversion parameter ρ_i, we have:

$$u_i(W_h) = -e^{-\rho_i W_h}$$

Because risky asset returns are normally distributed, the value function $V_i(\alpha)$ can be written as:

$$\begin{aligned}
\max_\alpha V_i(\alpha) &= \max_\alpha E[u_i(W_h)] = \max_\alpha E[-e^{-\rho_h W_h}] \\
&= \max_\alpha -E[e^{-\rho_i(w[(1+r)+\alpha\tilde{s}])}] = \max_\alpha -ln\ (E[e^{-\rho_i w[(1+r)+\alpha\tilde{s}]}]) \\
&= \max_\alpha -(E[-\rho_i w[(1+r)+\alpha\tilde{s}]] + \frac{1}{2}Var[\rho_h w[(1+r)+\alpha\tilde{s}]]) \qquad (2.1) \\
&= \max_\alpha\ \rho_i w[(1+r)+\alpha\mu_s] - \frac{1}{2}[\alpha^2\sigma_s^2\rho_i^2 w^2] \\
&= \max_\alpha\ \rho_i w[(1+r)+\alpha\mu_s - \frac{1}{2}\alpha^2\sigma_s^2\rho_i w]
\end{aligned}$$

A.2 Household utility as a weighted sum of decision makers utilities

I propose a second model of household portfolio allocation. With respect to Section 2.2.1, the household maximizes a weighted sum of the two decision-makers utilities. I assume a CARA utility function $u_i(W_h)$ for each decision-maker (a and b), and egoistic preferences of agents (i.e.: their utility does not depends on partners utility). Wealth can be invested in two types of assets: a risk-free asset with constant return (r) and a risky asset with return $(1+\tilde{r}) = (1+r+\tilde{s})$ where $\tilde{s} \sim N(\mu_s, \sigma_s^2)$. Therefore, the problem of agent i is to maximize his/her own utility u_i, choosing the optimal share of wealth allocated in risky asset, α.
Wealth is defined as

$$W_h = w(1 + r + \alpha\tilde{s})$$

and u_i becomes:

$$u_i(W_h) = -e^{-\rho_i W_h}$$

Individual i maximizes:

$$\begin{aligned} \max_{\alpha} V_i(\alpha) &= \max_{\alpha} E[u_i(W_h)] = \max_{\alpha} E[-e^{-\rho_h W_h}] \\ &= \max_{\alpha} \ \rho_i w[(1+r) + \alpha\mu_s - \frac{1}{2}\alpha^2\sigma_s^2\rho_i w] \end{aligned} \tag{2.2}$$

while household utility $u_h(W_h)$ is:

$$u_h(W_h) = \mu_a u_a(W_h) + \mu_b u_b(W_h) \tag{2.3}$$

where the weights (μ_a,μ_b) are the bargaining powers of the two individuals. I normalize the weights as follow:

$$\gamma = \frac{\mu_a}{\mu_a + \mu_b} \qquad (1-\gamma) = \frac{\mu_b}{\mu_a + \mu_b} \tag{2.4}$$

such that $\gamma \in [0, 1]$ represents agent's a bargaining power. Thus, the household maximization problem in Equation 2.3 becomes:

$$u_h(W_h) = \gamma u_a(W_h) + (1 - \gamma)u_b(W_h) \tag{2.5}$$

In terms of the value function $V_h(\alpha)$, the program becomes:

$$\begin{aligned} \max_{\alpha} \; V_h(\alpha) &= \max_{\alpha} E[u_h(W_h)] \\ &= \max_{\alpha} E[\gamma u_a(W_h) + (1 - \gamma)u_b(W_h)] \\ &= \max_{\alpha} \; \gamma E[u_a(W_h)] + (1 - \gamma)E[u_b(W_h)] \\ &= \max_{\alpha} \; \gamma \rho_a w[(1 + r) + \alpha\mu_s - \frac{1}{2}\alpha^2\sigma_s^2 w\rho_a] + \\ &+ (1 - \gamma)\rho_b w[(1 + r) + \alpha\mu_s - \frac{1}{2}\alpha^2\sigma_s^2 w\rho_b] \end{aligned} \tag{2.6}$$

Then, solving the first order conditions for α:

$$\alpha = \frac{\mu_s}{\sigma_s^2 w} \cdot \frac{\rho_h}{\phi} \tag{2.7}$$

where $\rho_h = \gamma\rho_a + (1 - \gamma)\rho_b$ and $\phi = \gamma\rho_a^2 + (1 - \gamma)\rho_b^2$.

B ELSA dataset

B.1 Numeracy

Table 2.1: Sample selection

Selection	Decrease in observation
Initial sample (individuals)	8445
Individuals in a couple with joint finances	4965
Couple with joint finances (both partners)	2325
Couple with both partners younger than 90s	2304
Couple with both partners self-reported risk valid answers	1582
Couple with positive income	1577
Couple with non-negative net wealth	1465
Couple with risky share of wealth $\in [0, 1]$	1445
Health index missing values	1441

Table 2.2: Numeracy: partners by gender

Male numeracy

numeracy	classification	Freq.	Percent	Cum.
0	0	45	3.1	3.1
1	low	220	15.3	18.4
2	mid	84	5.8	24.2
3	mid	33	2.3	26.5
4	mid	54	3.7	30.3
5	high	1,005	69.7	100.00
Total		1,441	100.00	-

Female numeracy

numeracy	classification	Freq.	Percent	Cum.
0	0	76	5.3	5.3
1	low	331	23.0	28.3
2	mid	90	6.2	34.5
3	mid	45	3.1	37.6
4	mid	61	4.2	41.8
5	high	838	58.1	100.00
Total		1,441	100.00	-

Table 2.3: Numeracy: partners by financial respondent

Respondent numeracy

numeracy	classification	Freq.	Percent	Cum.
0	0	35	2.4	2.4
1	low	254	17.6	20.0
2	mid	72	5.0	25.0
3	mid	31	2.2	27.2
4	mid	66	4.6	31.8
5	high	983	68.2	100.00
Total		1,441	100.00	-

Non-respondent numeracy

numeracy	classification	Freq.	Percent	Cum.
0	0	86	6.0	6.0
1	low	297	20.6	26.6
2	mid	102	7.1	33.7
3	mid	47	3.3	37.0
4	mid	49	3.4	40.4
5	high	860	59.6	100.00
Total		1,441	100.00	-

Table 2.4: Numeracy and stock market participation: partners by gender

Male numeracy

numeracy category	non-stockholders obs	%	stockholders obs	%
0	26	7,0%	19	1,8%
low	79	21,3%	141	13,2%
mid	44	11,9%	127	11,9%
high	222	59,8%	783	73,2%

Female numeracy

numeracy category	non-stockholders obs	%	stockholders obs	%
0	38	10,2%	38	3,6%
low	91	24,5%	240	22,4%
mid	53	14,3%	143	13,4%
high	189	50,9%	649	60,7%

Table 2.5: Numeracy and stock market participation: partners by financial respondent

Respondent numeracy

numeracy category	non-stockholders obs	%	stockholders obs	%
0	21	5,7%	14	1,3%
low	79	21,3%	175	16,4%
mid	45	12,1%	124	11,6%
high	226	60,9%	757	70,7%

Non-respondent numeracy

numeracy category	non-stockholders obs	%	stockholders obs	%
0	43	11,6%	43	4,0%
low	91	24,5%	206	19,3%
mid	52	14,0%	146	13,6%
high	185	49,9%	675	63,1%

C Results - robustness check

C.1 Baseline model - financial vs general risk measure - selection equation

Table 2.6: Heckman selection equation: household probability of holding risky assets. Baseline model selection: financial and general risk tolerance. Partners characteristics by gender.

participation	Financial risk measure (1)	General risk measure (2)
low numeracy: male	0.3349	0.3344
	(0.2126)	(0.2129)
mid numeracy: male	0.4983**	0.5036**
	(0.2228)	(0.2232)
high numeracy: male	0.4909**	0.4865**
	(0.2020)	(0.2027)
low numeracy: female	0.4710***	0.4671***
	(0.1721)	(0.1724)
mid numeracy: female	0.4513**	0.4378**
	(0.1855)	(0.1852)
high numeracy: female	0.4554***	0.4540***
	(0.1637)	(0.1638)
age: male	0.1175*	0.1174*
	(0.0623)	(0.0618)
age^2: male	-0.0008*	-0.0008*
	(0.0004)	(0.0004)
age difference > 10	0.0732	0.0729
	(0.1641)	(0.1641)
2^{nd} hh income quartile	0.2332**	0.2543**
	(0.1057)	(0.1059)
3^{rd} hh income quartile	0.3167***	0.3262***
	(0.1163)	(0.1165)
4^{th} hh income quartile	0.3337**	0.3225**
	(0.1331)	(0.1327)
2^{nd} hh wealthquartile	0.6967***	0.7190***
	(0.1034)	(0.1031)
3^{rd} hh wealthquartile	0.9503***	0.9848***
	(0.1090)	(0.1085)
4^{th} hh wealth quartile	1.5015***	1.5366***
	(0.1380)	(0.1362)
in work: male	-0.2238**	-0.2083*
	(0.1107)	(0.1104)
in work: female	-0.0074	0.0218
	(0.1071)	(0.1069)
mid education: male	0.3034***	0.3097***
	(0.0963)	(0.0963)
high education: male	0.3293**	0.3254**
	(0.1349)	(0.1354)
mid education: female	0.0441	0.0431
	(0.0943)	(0.0944)
high education: female	0.1309	0.1224
	(0.1428)	(0.1430)
health index: male	-0.0164	-0.0192
	(0.0189)	(0.0189)
health index: female	-0.0144	-0.0133
	(0.0179)	(0.0178)
financial risk: hh	0.0099	
	(0.0203)	
financial patience: hh	-0.0172	
	(0.0232)	
general risk: hh		-0.0050
		(0.0193)
general patience: hh		-0.0533**
		(0.0211)
Constant	-5.4767**	-5.2259**
	(2.1921)	(2.1709)
Number of observations	1,441	1,441

C.2 Collective vs unitary approach - selection equation

Table 2.7: Heckman selection equation: household probability of holding risky assets. Collective vs unitary approach. Partners characteristics by gender.

share of risky assets	Baseline model (1)	Collective Individual (2)	Collective Weighted (3)	Unitary (4)
low numeracy: male	0.3293	0.3338	0.3277	0.3178
	(0.2135)	(0.2134)	(0.2129)	(0.2131)
mid numeracy: male	0.4910**	0.4944**	0.4900**	0.4779**
	(0.2238)	(0.2237)	(0.2231)	(0.2233)
high numeracy: male	0.4917**	0.4958**	0.4897**	0.4770**
	(0.2029)	(0.2027)	(0.2022)	(0.2026)
low numeracy: female	0.4829***	0.4872***	0.4861***	0.4826***
	(0.1728)	(0.1729)	(0.1725)	(0.1720)
mid numeracy: female	0.4583**	0.4617**	0.4594**	0.4625**
	(0.1860)	(0.1861)	(0.1857)	(0.1853)
high numeracy: female	0.4606***	0.4634***	0.4626***	0.4606***
	(0.1642)	(0.1642)	(0.1639)	(0.1634)
age: male	0.1175*	0.1170*	0.1176*	0.1150*
	(0.0622)	(0.0621)	(0.0622)	(0.0621)
age^2: male	-0.0008*	-0.0008*	-0.0008*	-0.0008*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
age difference > 10	0.0828	0.0895	0.0835	0.0784
	(0.1641)	(0.1642)	(0.1641)	(0.1639)
2^{nd} hh income quartile	0.2344**	0.2341**	0.2335**	0.2353**
	(0.1060)	(0.1060)	(0.1059)	(0.1059)
3^{rd} hh income quartile	0.3242***	0.3243***	0.3244***	0.3241***
	(0.1167)	(0.1167)	(0.1167)	(0.1166)
4^{th} hh income quartile	0.3289**	0.3269**	0.3289**	0.3264**
	(0.1334)	(0.1334)	(0.1333)	(0.1332)
2^{nd} hh wealth quartile	0.6982***	0.6958***	0.6971***	0.6931***
	(0.1035)	(0.1035)	(0.1033)	(0.1033)
3^{rd} hh wealth quartile	0.9521***	0.9494***	0.9528***	0.9496***
	(0.1094)	(0.1094)	(0.1092)	(0.1091)
4^{th} hh wealth quartile	1.5097***	1.5060***	1.5086***	1.5077***
	(0.1393)	(0.1392)	(0.1391)	(0.1390)
in work: male	-0.2171**	-0.2210**	-0.2183**	-0.2174**
	(0.1107)	(0.1107)	(0.1106)	(0.1105)
in work: female	-0.0042	0.0003	-0.0018	-0.0112
	(0.1075)	(0.1075)	(0.1075)	(0.1074)
mid education: male	0.3043***	0.3065***	0.3033***	0.3040***
	(0.0963)	(0.0963)	(0.0963)	(0.0962)
high education: male	0.3374**	0.3390**	0.3393**	0.3288**
	(0.1352)	(0.1353)	(0.1352)	(0.1349)
mid education: female	0.0470	0.0486	0.0458	0.0495
	(0.0943)	(0.0943)	(0.0943)	(0.0942)
high education: female	0.1270	0.1300	0.1257	0.1266
	(0.1431)	(0.1431)	(0.1430)	(0.1429)
health index: male	-0.0175	-0.0179	-0.0175	-0.0177
	(0.0189)	(0.0189)	(0.0189)	(0.0189)
health index: female	-0.0122	-0.0116	-0.0120	-0.0120
	(0.0179)	(0.0179)	(0.0179)	(0.0179)
financial risk: male	0.0056	0.0097	0.0092	0.0066
	(0.0170)	(0.0171)	(0.0171)	(0.0171)
financial risk: female	0.0030	-0.0057	-0.0010	0.0142
	(0.0169)	(0.0174)	(0.0171)	(0.0169)
financial patience: male	-0.0269	-0.0270	-0.0272	-0.0267
	(0.0189)	(0.0192)	(0.0192)	(0.0192)
financial patience: female	0.0213	0.0223	0.0210	0.0207
	(0.0176)	(0.0181)	(0.0179)	(0.0175)
Constant	-5.5942**	-5.5781**	-5.5943**	-5.5328**
	(2.1842)	(2.1823)	(2.1835)	(2.1821)
Number of observations	1,441	1,441	1,441	1,441

C.3 Household portfolio allocation - general risk measure

Table 2.8: Heckman 1^{st} stage: household probability of holding risky assets. Partners characteristics by gender. General risk tolerance.

participation	General risk	
	Individual	Weighted
	(1)	(2)
low numeracy: male	0.4725**	0.4724**
	(0.2269)	(0.2270)
mid numeracy: male	0.6623***	0.6634***
	(0.2374)	(0.2376)
high numeracy: male	0.6753***	0.6756***
	(0.2140)	(0.2142)
low numeracy: female	0.3985**	0.3855**
	(0.1819)	(0.1816)
mid numeracy: female	0.3704*	0.3559*
	(0.1952)	(0.1953)
high numeracy: female	0.4125**	0.3967**
	(0.1727)	(0.1725)
age: male	0.1268**	0.1244**
	(0.0630)	(0.0630)
age^2: male	-0.0008*	-0.0008*
	(0.0005)	(0.0005)
age difference > 10	0.1082	0.1091
	(0.1662)	(0.1663)
2^{nd} hh income quartile	0.2713**	0.2670**
	(0.1069)	(0.1067)
3^{rd} hh income quartile	0.3530***	0.3541***
	(0.1177)	(0.1177)
4^{th} hh income quartile	0.3543***	0.3569***
	(0.1342)	(0.1341)
2^{nd} hh wealth quartile	0.7633***	0.7628***
	(0.1076)	(0.1075)
3^{rd} hh wealth quartile	0.9259***	0.9208***
	(0.1146)	(0.1143)
4^{th} hh wealth quartile	1.4403***	1.4396***
	(0.1437)	(0.1434)
in work: male	-0.2068*	-0.2120*
	(0.1127)	(0.1125)
in work: female	0.0010	0.0051
	(0.1087)	(0.1081)
mid education: male	0.3028***	0.3013***
	(0.0976)	(0.0976)
high education: male	0.3268**	0.3263**
	(0.1378)	(0.1376)
mid education: female	0.0495	0.0508
	(0.0954)	(0.0953)
high education: female	0.1382	0.1443
	(0.1456)	(0.1456)
health index: male	-0.0224	-0.0221
	(0.0191)	(0.0191)
health index: female	-0.0159	-0.0156
	(0.0182)	(0.0181)
general risk: male	0.0021	
	(0.0162)	
general risk: female	0.0066	
	(0.0160)	
general patience: male	-0.0435**	
	(0.0171)	
general patience: female	-0.0257	
	(0.0170)	
general risk: hh		0.0020
		(0.0195)
general patience: hh		-0.0608***
		(0.0213)
Constant	-5.7200***	-5.6305**
	(2.2179)	(2.2169)
Number of observations	1,441	1,441

Table 2.9: Heckman 2^{nd} stage: household probability of holding risky assets. Partners characteristics by gender (1^{st} stage is Table 2.8).

share of risky assets	General risk	
	Individual	Weighted
	(1)	(2)
age: male	0.0106	0.0121
	(0.0198)	(0.0198)
age^2: male	-0.0001	-0.0001
	(0.0001)	(0.0001)
age difference > 10	0.0269	0.0216
	(0.0415)	(0.0416)
2^{nd} hh income quartile	0.0168	0.0184
	(0.0329)	(0.0329)
3^{rd} hh income quartile	0.0093	0.0105
	(0.0351)	(0.0353)
4^{th} hh income quartile	-0.0081	-0.0054
	(0.0363)	(0.0365)
2^{nd} hh wealth quartile	0.0750	0.0826
	(0.0668)	(0.0668)
3^{rd} hh wealth quartile	0.1819**	0.1928**
	(0.0753)	(0.0750)
4^{th} hh wealth quartile	0.2525***	0.2658***
	(0.0920)	(0.0918)
in work: male	-0.0201	-0.0274
	(0.0283)	(0.0285)
in work: female	0.0025	0.0043
	(0.0258)	(0.0258)
mid education: male	0.0305	0.0317
	(0.0306)	(0.0305)
high education: male	0.0269	0.0290
	(0.0359)	(0.0359)
mid education: female	0.0331	0.0334
	(0.0252)	(0.0253)
high education: female	0.1042***	0.1047***
	(0.0326)	(0.0328)
health index: male	-0.0012	-0.0007
	(0.0057)	(0.0057)
health index: female	-0.0037	-0.0041
	(0.0050)	(0.0050)
general risk: male	0.0054	
	(0.0040)	
general risk: female	0.0046	
	(0.0040)	
general patience: male	-0.0142***	
	(0.0046)	
general patience: female	-0.0014	
	(0.0044)	
general risk: hh		0.0106**
		(0.0049)
general patience: hh		-0.0150***
		(0.0058)
Constant	-0.2054	-0.2739
	(0.7528)	(0.7535)
Inverse Mills Ratio		
lambda	0.0540	0.0763
	(0.1302)	(0.1301)
Number of observations	1,441	1,441

C.4 Heckman estimates - partners characteristics by financial respondent

Table 2.10: Heckman 1^{st} stage: household probability of holding risky assets. Partners characteristics by financial respondent.

participation	General risk		Financial risk	
	Individual	Weighted	Individual	Weighted
low numeracy: respondent	0.5288**	0.5318**	0.5383**	0.5466**
	(0.2538)	(0.2539)	(0.2534)	(0.2529)
mid numeracy: respondent	0.5511**	0.5507**	0.5560**	0.5617**
	(0.2647)	(0.2647)	(0.2646)	(0.2640)
high numeracy: respondent	0.5809**	0.5814**	0.5859**	0.5926**
	(0.2441)	(0.2442)	(0.2437)	(0.2432)
low numeracy: non-respondent	0.3904**	0.3668**	0.3736**	0.3673**
	(0.1729)	(0.1727)	(0.1716)	(0.1711)
mid numeracy: non-respondent	0.4459**	0.4288**	0.4320**	0.4325**
	(0.1841)	(0.1840)	(0.1830)	(0.1826)
high numeracy: non-respondent	0.5105***	0.4908***	0.4832***	0.4835***
	(0.1613)	(0.1611)	(0.1600)	(0.1594)
age: respondent	0.0456	0.0461	0.0472	0.0463
	(0.0612)	(0.0610)	(0.0610)	(0.0608)
age^2: respondent	-0.0003	-0.0003	-0.0003	-0.0003
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
age difference > 10	0.0614	0.0588	0.0552	0.0552
	(0.1634)	(0.1636)	(0.1634)	(0.1634)
2^{nd} hh income quartile	0.2853***	0.2847***	0.2616**	0.2604**
	(0.1070)	(0.1066)	(0.1065)	(0.1062)
3^{rd} hh income quartile	0.3966***	0.4002***	0.3815***	0.3818***
	(0.1169)	(0.1170)	(0.1166)	(0.1165)
4^{th} hh income quartile	0.3926***	0.3930***	0.3900***	0.3919***
	(0.1337)	(0.1335)	(0.1335)	(0.1335)
2^{nd} hh wealth quartile	0.7485***	0.7527***	0.7187***	0.7273***
	(0.1073)	(0.1074)	(0.1067)	(0.1067)
3^{rd} hh wealth quartile	0.9373***	0.9368***	0.9041***	0.9104***
	(0.1157)	(0.1154)	(0.1150)	(0.1149)
4^{th} hh wealth quartile	1.4419***	1.4480***	1.4221***	1.4309***
	(0.1449)	(0.1448)	(0.1452)	(0.1451)
in work: respondent	-0.1460	-0.1458	-0.1651	-0.1675
	(0.1096)	(0.1096)	(0.1094)	(0.1095)
in work: non-respondent	-0.1999*	-0.1931*	-0.2027*	-0.2071*
	(0.1080)	(0.1078)	(0.1079)	(0.1077)
mid edu: respondent	0.2848***	0.2844***	0.2709***	0.2684***
	(0.0967)	(0.0967)	(0.0965)	(0.0965)
high edu: respondent	0.3403**	0.3405**	0.3471**	0.3423**
	(0.1379)	(0.1378)	(0.1376)	(0.1375)
mid edu: non-respondent	0.0316	0.0341	0.0529	0.0516
	(0.0969)	(0.0967)	(0.0961)	(0.0960)
high edu: non-respondent	0.1167	0.1247	0.1178	0.1228
	(0.1466)	(0.1464)	(0.1457)	(0.1456)
health index: respondent	-0.0405**	-0.0409**	-0.0389*	-0.0387*
	(0.0206)	(0.0206)	(0.0205)	(0.0204)
health index: non-respondent	-0.0038	-0.0027	-0.0009	-0.0016
	(0.0173)	(0.0173)	(0.0173)	(0.0173)
female respondent	0.0765	0.0914	0.0911	0.0815
	(0.0900)	(0.0870)	(0.0880)	(0.0867)
general risk: respondent	0.0000			
	(0.0162)			
general risk: non-respondent	0.0151			
	(0.0160)			
general patience: respondent	-0.0332*			
	(0.0170)			
general patience: non-respondent	-0.0337**			
	(0.0171)			
general risk: hh		0.0075		
		(0.0195)		
general patience: hh		-0.0599***		
		(0.0213)		
financial risk: respondent			0.0169	
			(0.0174)	
financial risk: non-respondent			0.0018	
			(0.0175)	
financial patience: respondent			-0.0073	
			(0.0192)	
financial patience: non-respondent			0.0098	
			(0.0183)	
financial risk: hh				0.0199
				(0.0205)
financial patience: hh				-0.0142
				(0.0235)
Constant	-2.6718	-2.6897	-3.1306	-2.9798
	(2.1347)	(2.1276)	(2.1310)	(2.1254)
Number of observations	1,441	1,441	1,441	1,441

Table 2.11: Heckman 2^{nd} stage: household share of net financial wealth allocated in risky assets. Partners characteristics by financial respondent (1^{st} stage is Table 2.10).

share of risky assets	General risk		Financial risk	
	Individual	Weighted	Individual	Weighted
age: respondent	-0.0054	-0.0054	-0.0063	-0.0062
	(0.0165)	(0.0164)	(0.0163)	(0.0163)
age^2: respondent	0.0001	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
age difference > 10	0.0310	0.0301	0.0271	0.0223
	(0.0414)	(0.0413)	(0.0408)	(0.0408)
2^{nd} hh income quartile	0.0047	0.0058	-0.0041	-0.0053
	(0.0336)	(0.0335)	(0.0329)	(0.0328)
3^{rd} hh income quartile	-0.0106	-0.0049	-0.0066	-0.0048
	(0.0366)	(0.0368)	(0.0362)	(0.0361)
4^{th} hh income quartile	-0.0270	-0.0215	-0.0192	-0.0190
	(0.0376)	(0.0377)	(0.0373)	(0.0374)
2^{nd} hh wealth quartile	0.0143	0.0410	0.0143	0.0308
	(0.0686)	(0.0693)	(0.0682)	(0.0687)
3^{rd} hh wealth quartile	0.1174	0.1464*	0.1094	0.1292*
	(0.0786)	(0.0792)	(0.0781)	(0.0785)
4^{th} hh wealth quartile	0.1705*	0.2063**	0.1600*	0.1817*
	(0.0960)	(0.0970)	(0.0964)	(0.0969)
in work: respondent	0.0069	0.0048	-0.0055	-0.0060
	(0.0269)	(0.0268)	(0.0269)	(0.0269)
in work: non-respondent	-0.0293	-0.0327	-0.0374	-0.0367
	(0.0276)	(0.0273)	(0.0270)	(0.0271)
mid edu: respondent	0.0098	0.0154	0.0050	0.0077
	(0.0313)	(0.0314)	(0.0309)	(0.0308)
high edu: respondent	0.0391	0.0450	0.0399	0.0448
	(0.0366)	(0.0365)	(0.0364)	(0.0363)
mid edu: non-respondent	0.0297	0.0321	0.0295	0.0309
	(0.0249)	(0.0248)	(0.0245)	(0.0245)
high edu: non-respondent	0.0604*	0.0643**	0.0545*	0.0581*
	(0.0326)	(0.0324)	(0.0320)	(0.0320)
health index: respondent	-0.0039	-0.0045	-0.0029	-0.0033
	(0.0063)	(0.0063)	(0.0062)	(0.0062)
health index: non-respondent	0.0005	0.0004	0.0014	0.0018
	(0.0047)	(0.0047)	(0.0046)	(0.0046)
female respondent	0.0049	-0.0033	-0.0075	-0.0086
	(0.0222)	(0.0214)	(0.0218)	(0.0210)
general risk: respondent	0.0100**			
	(0.0040)			
general risk: non-respondent	-0.0006			
	(0.0041)			
general patience: respondent	-0.0080*			
	(0.0044)			
general patience: non-respondent	-0.0041			
	(0.0046)			
general risk: hh		0.0101**		
		(0.0049)		
general patience: hh		-0.0128**		
		(0.0059)		
financial risk: respondent			0.0198***	
			(0.0043)	
financial risk: non-respondent			0.0138***	
			(0.0044)	
financial patience: respondent			-0.0002	
			(0.0047)	
financial patience: non-respondent			-0.0010	
			(0.0047)	
financial risk: hh				0.0296***
				(0.0052)
financial patience: hh				-0.0008
				(0.0059)
Constant	0.4913	0.4400	0.3710	0.3530
	(0.5985)	(0.5974)	(0.6038)	(0.6005)
Inverse Mills Ratio				
lambda	-0.0698	-0.0157	-0.0445	-0.0162
	(0.1344)	(0.1359)	(0.1358)	(0.1361)
Number of observations	1,441	1,441	1,441	1,441

Table 2.12: Heckman 1^{st} stage: household probability of holding risky assets. Partners characteristics by financial respondent. Restriction : numeracy of the financial respondent partner.

participation	General risk		Financial risk	
	Individual	Weighted	Individual	Weighted
low numeracy: respondent	0.5962**	0.5949**	0.5995**	0.6094**
	(0.2516)	(0.2518)	(0.2515)	(0.2508)
mid numeracy: respondent	0.6279**	0.6231**	0.6263**	0.6345**
	(0.2624)	(0.2624)	(0.2625)	(0.2618)
high numeracy: respondent	0.6682***	0.6636***	0.6661***	0.6747***
	(0.2412)	(0.2414)	(0.2412)	(0.2404)
age: respondent	0.0458	0.0460	0.0480	0.0461
	(0.0609)	(0.0608)	(0.0608)	(0.0606)
age^2: respondent	-0.0003	-0.0003	-0.0004	-0.0003
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
age difference > 10	0.0524	0.0520	0.0488	0.0491
	(0.1630)	(0.1631)	(0.1630)	(0.1629)
2^{nd} hh income quartile	0.3000***	0.2994***	0.2782***	0.2758***
	(0.1065)	(0.1061)	(0.1061)	(0.1057)
3^{rd} hh income quartile	0.4057***	0.4095***	0.3914***	0.3911***
	(0.1162)	(0.1162)	(0.1159)	(0.1158)
4^{th} hh income quartile	0.4095***	0.4097***	0.4065***	0.4085***
	(0.1330)	(0.1329)	(0.1328)	(0.1329)
2^{nd} hh wealth quartile	0.7519***	0.7571***	0.7223***	0.7323***
	(0.1068)	(0.1069)	(0.1063)	(0.1063)
3^{rd} hh wealth quartile	0.9421***	0.9416***	0.9093***	0.9167***
	(0.1149)	(0.1147)	(0.1144)	(0.1142)
4^{th} hh wealth quartile	1.4650***	1.4691***	1.4418***	1.4518***
	(0.1443)	(0.1441)	(0.1446)	(0.1445)
in work: respondent	-0.1386	-0.1372	-0.1563	-0.1586
	(0.1092)	(0.1092)	(0.1090)	(0.1091)
in work: non-respondent	-0.1994*	-0.1936*	-0.2015*	-0.2077*
	(0.1074)	(0.1073)	(0.1074)	(0.1072)
mid edu: respondent	0.2726***	0.2730***	0.2587***	0.2564***
	(0.0962)	(0.0962)	(0.0960)	(0.0960)
high edu: respondent	0.3187**	0.3191**	0.3253**	0.3199**
	(0.1372)	(0.1371)	(0.1369)	(0.1368)
mid edu: non-respondent	0.0455	0.0466	0.0649	0.0636
	(0.0964)	(0.0963)	(0.0957)	(0.0956)
high edu: non-respondent	0.1567	0.1632	0.1567	0.1615
	(0.1453)	(0.1452)	(0.1445)	(0.1443)
health index: respondent	-0.0390*	-0.0397*	-0.0377*	-0.0374*
	(0.0205)	(0.0205)	(0.0204)	(0.0203)
health index: non-respondent	-0.0114	-0.0102	-0.0083	-0.0092
	(0.0170)	(0.0170)	(0.0170)	(0.0170)
female respondent	0.0981	0.1105	0.1125	0.1000
	(0.0892)	(0.0862)	(0.0872)	(0.0859)
general risk: respondent	0.0012			
	(0.0162)			
general risk: non-respondent	0.0124			
	(0.0159)			
general patience: respondent	-0.0312*			
	(0.0169)			
general patience: non-respondent	-0.0326*			
	(0.0170)			
general risk: hh		0.0072		
		(0.0194)		
general patience: hh		-0.0592***		
		(0.0212)		
financial risk: respondent			0.0192	
			(0.0173)	
financial risk: non-respondent			-0.0023	
			(0.0173)	
financial patience: respondent			-0.0059	
			(0.0191)	
financial patience: non-respondent			0.0092	
			(0.0183)	
financial risk: hh				0.0195
				(0.0205)
financial patience: hh				-0.0145
				(0.0234)
Constant	-2.3654	-2.3770	-2.8462	-2.6639
	(2.1216)	(2.1149)	(2.1194)	(2.1131)
Number of observations	1,441	1,441	1,441	1,441

Table 2.13: Heckman 2^{nd} stage: household share of net financial wealth allocated in risky assets. Partners characteristics by financial respondent. Financial respondent partner numeracy is the only exclusion restriction (1^{st} stage is Table 2.12).

share of risky assets	General risk		Financial risk	
	Individual	Weighted	Individual	Weighted
age: respondent	-0.0056	-0.0050	-0.0058	-0.0055
	(0.0167)	(0.0166)	(0.0165)	(0.0165)
age^2: respondent	0.0001	0.0000	0.0001	0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
age difference > 10	0.0307	0.0306	0.0277	0.0231
	(0.0416)	(0.0414)	(0.0409)	(0.0410)
2^{nd} hh income quartile	0.0036	0.0083	-0.0014	-0.0016
	(0.0367)	(0.0364)	(0.0359)	(0.0358)
3^{rd} hh income quartile	-0.0115	-0.0014	-0.0026	0.0004
	(0.0408)	(0.0406)	(0.0406)	(0.0405)
4^{th} hh income quartile	-0.0280	-0.0182	-0.0153	-0.0140
	(0.0415)	(0.0413)	(0.0415)	(0.0417)
2^{nd} hh wealth quartile	0.0114	0.0502	0.0250	0.0447
	(0.0848)	(0.0841)	(0.0849)	(0.0854)
3^{rd} hh wealth quartile	0.1138	0.1573	0.1220	0.1458
	(0.0986)	(0.0976)	(0.0988)	(0.0992)
4^{th} hh wealth quartile	0.1661	0.2201*	0.1761	0.2028
	(0.1217)	(0.1206)	(0.1234)	(0.1239)
in work: respondent	0.0070	0.0037	-0.0070	-0.0078
	(0.0274)	(0.0272)	(0.0275)	(0.0275)
in work: non-respondent	-0.0290	-0.0341	-0.0390	-0.0388
	(0.0285)	(0.0281)	(0.0280)	(0.0282)
mid edu: respondent	0.0095	0.0181	0.0083	0.0116
	(0.0338)	(0.0336)	(0.0333)	(0.0332)
high edu: respondent	0.0387	0.0477	0.0433	0.0489
	(0.0389)	(0.0386)	(0.0391)	(0.0388)
mid edu: non-respondent	0.0295	0.0324	0.0300	0.0316
	(0.0250)	(0.0248)	(0.0246)	(0.0246)
high edu: non-respondent	0.0599*	0.0654**	0.0556*	0.0597*
	(0.0331)	(0.0329)	(0.0325)	(0.0326)
health index: respondent	-0.0038	-0.0049	-0.0033	-0.0038
	(0.0065)	(0.0065)	(0.0064)	(0.0064)
health index: non-respondent	0.0006	0.0003	0.0013	0.0017
	(0.0047)	(0.0047)	(0.0046)	(0.0046)
female respondent	0.0046	-0.0028	-0.0069	-0.0079
	(0.0223)	(0.0216)	(0.0220)	(0.0212)
general risk: respondent	0.0099**			
	(0.0040)			
general risk: non-respondent	-0.0006			
	(0.0041)			
general patience: respondent	-0.0078*			
	(0.0047)			
general patience: non-respondent	-0.0040			
	(0.0048)			
general risk: hh		0.0102**		
		(0.0050)		
general patience: hh		-0.0132**		
		(0.0065)		
financial risk: respondent			0.0199***	
			(0.0045)	
financial risk: non-respondent			0.0138***	
			(0.0044)	
financial patience: respondent			-0.0002	
			(0.0047)	
financial patience: non-respondent			-0.0008	
			(0.0047)	
financial risk: hh				0.0298***
				(0.0053)
financial patience: hh				-0.0009
				(0.0059)
Constant	0.5041	0.4078	0.3291	0.2984
	(0.6260)	(0.6224)	(0.6398)	(0.6353)
Inverse Mills ratio				
lambda	-0.0756	0.0049	-0.0199	0.0156
	(0.1736)	(0.1719)	(0.1774)	(0.1776)
Number of observations	1,441	1,441	1,441	1,441

C.5 Collective vs unitary - partners characteristics by financial respondent

Table 2.14: Heckman outcome equation: household share of net financial wealth allocated in risky assets. Collective vs unitary approach. Partners characteristics by financial respondent.

share of risky assets	Baseline model (1)	Collective models Partner risk (2)	Household risk (3)	Unitary model (4)
demographics[a]	*	*	*	*
financial risk: hh	0.0305***		0.0229***	
	(0.0053)		(0.0078)	
financial patience: hh	0.0004		-0.0027	
	(0.0062)		(0.0086)	
financial risk: respondent		0.0198***	0.0084	0.0225***
		(0.0043)	(0.0065)	(0.0044)
financial risk: non-respondent		0.0138***		
		(0.0044)		
financial patience: respondent		-0.0002	0.0021	-0.0003
		(0.0047)	(0.0070)	(0.0050)
financial patience: non-respondent		-0.0010		
		(0.0047)		
Constant	0.4753	0.3853	0.4510	0.4126
	(0.5831)	(0.5874)	(0.5834)	(0.5863)
likelihood ratio test		Col (2) and (4)	Col (3) and (4)	
p-value		0.0001	0.0121	
AIC	1,939			1,949
BIC	2,187			2,208
Number of observations	1,441	1,441	1,441	1,441

[a]Demographics include financial respondent age and age squared, dummy for large age difference between partners (> 10 years), dummies of income quartile, job market participation of partners, education of partners and the health index of partners.

Chapter 3

Subjective survival expectations and individual portfolio choices

3.1 Introduction

The standard portfolio allocation theory implies that all agents should hold risky assets and have a well-diversified portfolio, as Merton (1969) and Samuelson (1975) explain. However, data show that households do not behave according to the theory: only a few participate in the stock market (Haliassos and Bertaut (1995)) and their portfolios are generally not well-diversified (Kelly (1995)). These differences between theory and data, in particular the low stock market participation rate, are known in economics with the name of *equity premium puzzle*.

The equity premium is defined as the difference in returns between risky (stocks) and risk-free (treasury bills) assets. Mehra and Prescott (1985) estimate the equity premium to be in the range of 5% to 8% per year, and these results are in line with other recent papers (e.g.: Mehra (2007)). The puzzle arises because the large difference in percentage returns implies an unreasonably high level of risk aversion among non-stockholders. Economists attempt to explain (at least partially) the equity premium puzzle proposing different solutions in the past years: stock market participation costs (Vissing-Jorgensen (2002)), financial literacy (Van Rooij, Lusardi, and Alessie (2011)) or alternative behavioural approaches (e.g.:Benartzi and Thaler (1995)) are only some of the possible explanations proposed by researchers. A rather unexplored element that can contribute to explain the equity premium puzzle is the time horizon. Indeed, stock returns have shown to be a mean-reverting process, implying that stocks are safer in the long term, but may display negative results in short periods. Mehra and Prescott (1985) show that risky assets outperform treasury bills over long periods (10 years or more), but may show negative results in the short-run due to the volatility of the market (1 to 5 years). Two classic examples are the year of the dot com bubble (2001) or the financial crash (2008), in which the stock market lost more than 20% of its capitalization. In other words, an investment in risky assets that stands for less than ten years may cause capital losses because of the market volatility, while if the same risky investment has a longer time horizon, its returns outperform the safe assets performances. This evidence coincides with the common financial advice that the longer is the horizon of the investment, the higher should be the share of capital allocated in stocks. Veld-Merkoulova (2011) show that the behaviours of the private investors are in line with those indications: they allocate a larger proportion of their wealth in risky assets at the early stage of life, and this share is decreasing with the age of the individual. Therefore, a crucial choice in the portfolio allocation of households is the time horizon of the investments. This decision is complicated especially among the elderly, because of the higher uncertainty toward their remaining life. Then, subjective survival expectation should play a role in the portfolio allocation process of the elderly. Indeed, high survival beliefs imply a longer (expected) life horizon, which implies also a longer horizon for individual investments. In recent years, subjective survival expectations attracted an increasing interest because they influence several households' decisions. Van Solinge and Henkens (2010) use dutch data to demonstrate that older employees take into account subjective life expectancy in the choice of the retirement age. Their results show that agents who choose to retire later expect to live longer. On the other hand, Bloom et al. (2006) find that the length of the working life of couples is not affected by subjective survival probabilities. However, they show that the households' accumulated wealth for retirement increases in life expectancy in the US. Nivakoski (2020) finds similar results using Irish data. Other papers use subjective survival expectation to estimate life-cycle consumption and saving models (e.g.: M. Hurd, McFadden, and Gan (2008)) or retirement models (Bresser (2020)). These papers show that the precision of the model predictions increases using subject rather than objective survival expectations. Last, O'Dea and Sturrock (2021) study the role of survival optimism in the annuity puzzle. However, none of these papers considers the household portfolio allocation problem and the influence of subjective survival probability in the investment decision process.

This paper studies the relationship between stock market participation of the elderly and their subjective survival expectations. Agents with pessimistic survival expectations may think to live not long enough to benefit from the equity premium and decide to not participate in the stock market, missing good opportunities.
Using the English Longitudinal Study of Ageing (ELSA) dataset, I use direct questions about individuals' life expectations to estimate the subjective survival curves and subjective life expectancy (in years), following the approach of O'Dea and Sturrock (2021) on annuities. I compare the estimated survival curves and objective survival probabilities[1], obtaining a survival optimism index. In line with the previous findings in the literature (Elder (2013), Gan, M. D. Hurd, and McFadden (2007)), the index shows that individuals underestimate their survival chances in their 50s, 60s and 70s, while they tend to be optimistic in their 80s and at older ages. Therefore, survival pessimism is dominant among the elderly and most pessimistic individuals may decide to not invest in stocks. This suggests that the non-participation decision may not be only a consequence of their risk aversion or stock market entry costs, but may derive also from subjective life expectation, as individuals think they perceive that they will live not long enough to benefit from the equity premium.

[1] I use the survival probabilities provided by the Office of National Statistics (ONS) life tables.

I test this hypothesis by studying the relationship between the survival optimism index and stock market participation of singles, controlling for a series of demographics such as age and individual health conditions. The probit estimates show that survival optimism and subjective life expectancy have positive and significant effects on agents' stock market participation.

This paper contributes to the literature providing evidence of additional explanations for the equity premium puzzle. I study the role played by subjective survival optimism in determining stock market participation. Stock market participation of individuals may be affected by pessimistic survival expectations because they imply a shorter expected time horizon. If this horizon is less than fifteen-twenty years, the agent thinks to have not enough time to benefit from the equity premium. The results show that the subjective survival horizon matters for the individual portfolio choices, going against the classical assumption of the standard constant-portfolio theory (e.g.:Merton (1969) and Samuelson (1975)).

The rest of the paper is organized as follow: Section 3.2 presents the ELSA data and the sample selection procedure, Section 3.3 estimates the individual survival curves and compares subjective and objective survival expectation. Section 3.4 presents the empirical results and Section 3.5 concludes.

3.2 ELSA data

ELSA is a longitudinal survey that collects data from a representative sample of English people aged 50 years and above. It is a biennial survey (first wave in 2002) that aims to gather data to study the aspects of the ageing process, like social care, retirement, pension policies and social participation in England. The original sample of ELSA (first wave) was selected from the Health Survey for England (HSE[2]) respondents in the period 1998-2001. After the first survey in 2002, younger age groups of ELSA are refreshed to balance the panel over time.

[2]More information about HSE at `http://healthsurvey.hscic.gov.uk`.

This paper works with Wave 8 data of ELSA, which collects data about 8,445 individuals interviewed between May 2016 and June 2017.

3.2.1 Life expectation

The ELSA survey includes a specific module about respondents expectations. This module asks individuals about their subjective beliefs concerning certain possible events in the future. For example, the questions ask about the probability of leaving an inheritance, of being a worker at a certain age and of changing residence in the next years. The following statement opens the expectation section of the survey:
"Now I have some questions about how likely you think various events might be. When I ask a question, I'd like you to give me a number from 0 to 100, where 0 means that you think there is absolutely no chance an event will happen, and 100 means that you think the event is absolutely certain to happen."

The interest of this study is in subjective survival probabilities. The survey has a question that asks individuals about their survival beliefs at a certain age. Specifically, the question asks:
"What are the chances that you will live to be age X or more?"
where the age X depends on the current age of the respondent as Table 3.1 shows.

Table 3.1: Life expectancy questions: *"What are the chances that you will live to be age X or more?"*.

respondent age	target age (X)
< 66	75
66 to 69	80
70 to 74	85
75 to 79	90
80 to 84	95
85 to 99	100

Additionally, individuals aged less than 70 years were asked a second question concerning their subjective survival chance to age 85, if their answer to the previous question is greater than 0. Therefore, younger respondents answer two different survival questions, while older respondents answer only one.

3.2.2 Subjective reports

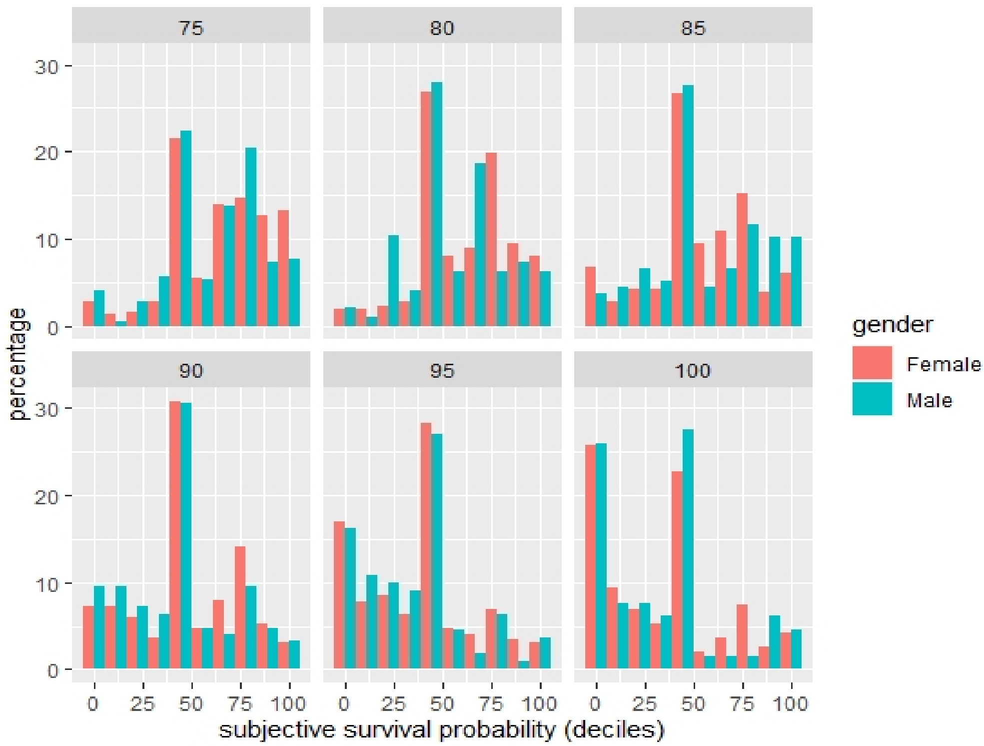
This Section analyses the survival probability answers of ELSA participants. Many papers, such as Tversky and Kahneman (1983) and Johnson et al. (1993), show that the concept of probability is complicated, and systematic biases affect agents' perception of probabilities. Therefore, a relevant point of the study is to assess whether individuals completely understood the questions of the expectation module and their probabilistic nature.

The life expectation questions accept only integer answers from 0 to 100, adding the possibility of *Don't know* answers. Only 3.6% of singles choose this option, revealing a high willingness to answer that specific question.

Figure 3.2.1 presents the distribution of subjective survival probabilities of ELSA respondents by target age, comparing females and males. Each sub-figure shows the distribution of subjective survival probability of females and males in bins of 10 percentage points (0-10, 11-20,..., 91-100). They show that the higher is the considered target age (reported at the top of each histogram), the lower are the average survival probabilities of respondents. Note that the target age increases with respondents' age, and the time horizon of agents is decreasing in target age.
Figure 3.2.1 also shows that almost a quarter of respondents choose a probability between 40% to 50% as the answer to the (first) survival question, with no gender differences. Further investigation reveals that 20.8% of agents report *50%* as their subjective survival belief. This focal answer may signal those who want to answer but have not understood the question and its probabilistic nature. I compare these answers and the other expectation questions to test this hypothesis. First, there are no respondents who answer *50%* to all questions of the expectation module. Secondly, the individuals who initially answer *50%* are not significantly more likely than others to pick *50%* again in the rest of the ELSA expectation module. Therefore, there is

no evidence to exclude those agents from the analysis.

Figure 3.2.1: Subjective survival probabilities distributions by target age and gender.



A non-negligible fraction of respondents answer *0%* and *100%* to the life-expectation questions, respectively 8.5% (185) and 6.5% (141). This might signal a lack of understanding of the concept of probability, where these answers express certainties. Therefore, I decide to exclude those respondent from the sample. An additional explanations for *0%* probability of survival answers is that those individuals might be in very bad health conditions or have received a terminal diagnosis. I find evidence that health is correlated with subjective survival chances comparing the sickness index values of those reporting *0%* survival chances and the rest of the sample. The former group has an average value of *2.76* while the rest of the sample index value is *0.56*, therefore, those reporting *0%* survival chances have worst health conditions[3].

[3]The sickness index is an indicator of the health status of the agent: the higher is the index, the

I account for implausible answers of agents to the survival expectation questions. In particular, when individuals are asked two survival questions, they may report a higher chance of survival to the older age than to the younger age. Such answers indicate a fundamental misunderstanding of the question and the probability concept. Therefore, I remove these observations (24 individuals).

Table 3.2 shows the average subjective survival probabilities of females and males respondents by target age. Overall, women survival beliefs are higher than that of men, but these differences reduce with age and target age.

Table 3.2: Average subjective survival probability by target age and gender.

target age	subjective survival probability	
	Females	Males
75	66.1%	62.2%
80	62.5%	57.2%
85	54.5%	53.8%
90	51.1%	44.3%
95	43.0%	39.1%
100	42.6%	43.2%

3.2.3 Sample selection

The ELSA survey categorizes individuals in three main financial unit types: single (2392 respondents), couple with separate finances (1080 respondents) and couple with joint finances (4965 respondents). For this study, I select only *single* agents. The aim is to capture the effect of subjective survival probability on financial choices, and the portfolio allocation process of couples may be affected by the preferences of both members. Studying the aggregation of spouses' beliefs in their financial decisions process is beyond the scope of this paper and will be investigated in future research.

worst are the health conditions of the individual (see Dal Bianco (2020) for more details). Sample mean is *0.75* and standard deviation is *2.99*.

The selection includes single respondents aged 50 to 89 years, with non-negative labour/pension income, total wealth and investments in risky assets. The final sample is composed of 1807 individuals, 1217 females and 590 males. Table 3.3 reports the basic statistics of the sample by gender:

Table 3.3: Sample descriptive statistics.

	Female	Male
obs	1217	590
age	72.0	69.9
median age	72	70
sd age	9.2	9.2
work	18,9%	23,2%
wealth: mean (thousand £)	267	375
wealth: median (thousand £)	190	197
wealth: sd (thousand £)	361	1.747
weekly income: mean (£)	281	331
weekly income: median (£)	246	287
weekly income: sd (£)	174	234
investments: mean (thousands £)	22.2	46.5
stock market participation	45,2%	49,3%

As expected, females are older, on average. Women live longer than men, as the ONS life tables show. Moreover, among singles, there may be a large number of widows who survived to their husbands. It is not surprising that a higher percentage of men is still in the job market. Indeed, the average retirement age in England is around 65 years and given that males are younger in this sample, there is a higher proportion of them aged less than 65 years. Wealth is higher among men, but the median values of males and females distribution show little difference[4]. Similarly, males average labour/pension income is 50£ per week higher than females income, but the medians reduce the difference to 30£ per week circa. As for wealth, there might be males with very high salaries whose values significantly affect the moments

[4]Wealth includes both financial wealth and housing wealth. Then, there are men with extremely high wealth who sensibly affect the moments of the wealth distribution.

of the distribution. The gender pay gap may depends on multiple factors (e.g.: company size and type of occupation) and reduces since the introduction of the Equal Pay Act in 1975 in the UK, but still affects the English society. Last, males are more likely to participate in the stock market than females and invest a larger amount of their wealth in risky assets.

3.3 Accuracy of subjective reports

Including subjective survival expectations in the individual decision process is a relatively new approach in economics and integrate the large literature about individual beliefs, as Manski (2004) explains.
This Section compares the mortality rates from the life tables. Then, I derive the subjective survival curves and life expectations of each respondent using the answers of ELSA. I use these estimates to assess the importance of individual survival beliefs in the portfolio allocation choice.

Overall, the subjective probabilities highlight that young respondents (less than 70 years) have pessimistic expectations about their survival chances, while older agents show increasing optimism. These findings are consistent with the literature that studies the accuracy of self-reported survival expectations. M. D. Hurd and McGarry (1995) is the first study that analyses the accuracy of subjective survival probabilities, using the Health and Retirement Study (HRS) data. They show that the respondents' expectations in the first wave of the HRS are consistent with objective survival probabilities, on average. However, men tend to underestimate their survival chances to younger ages, while women tend to overestimate their survival chances to older ages. M. D. Hurd and McGarry (2002) built on this work and study the evolution of subjective survival probabilities and their ability to predict the actual mortality rate. They show that health shocks (new diagnosis) affect individual survival beliefs and that the agents that survive for longer periods report 50% higher survival chances than those who died. Peracchi, Perotti, et al. (2010)

compare cross-country subjective survival probabilities in Europe using the Survey of Health, Ageing and Retirement in Europe (SHARE) data. Their results show that male expectations are close to the life table survival probability, whereas females tend to underestimate their survival chances at any age. Moreover, they find a positive relationship between subjective survival expectations and agents demographics such as education, income and wealth.

3.3.1 Comparing subjective and objective survival expectation

A life table is a demographic tool used to analyse death rates and calculate life expectancies at various ages. It presents data for males and females separately because of their different mortality patterns.
The Office for National Statistics (ONS) provides the life tables of the United Kingdom. It collects information about actual and projected mortality data for the population by age, cohort and gender. These data become the objective survival probability, i.e. the benchmark against which I assess the subjective expectations of the ELSA sample. However, as O'Dea and Sturrock (2021) show, the ELSA sample is not comparable to the English residents because the survey excludes the institutionalized fraction of the population. In other words, ELSA excludes those who live in residential care and other similar structures. Therefore, the mortality rate of ELSA is slightly lower than that of the English population. Using the administrative death records linked to ELSA, O'Dea and Sturrock (2021) re-scale the ONS life tables to adjust the difference in mortality rate implied by the ELSA sample selection. They show that the ELSA mortality rates for men and women are 71% and 69% of the ONS original level, respectively. In other words, those in the ELSA sample are less likely to die (or have higher life expectancies) than the English population. According to these evidence, I re-scale the ONS life table reducing the reported mortality rates by 71% and 69% for men and women. These re-scaled survival curves become the objective survival measure in the rest of the paper, used as benchmark in the

empirical analysis.

3.3.2 Survival optimism index

This Section studies whether ELSA respondents show optimistic or pessimistic beliefs about their survival probabilities, using the re-scaled ONS objective survival curve. I differentiate the respondents by age and sex and I compare the average subjective survival beliefs of each group with the correspondent objective survival probability.

Figure 3.3.2 shows the results of this comparison. It represents the average subjective and objective survival probabilities grouped by target age and gender. The individual reports (solid lines) represent the weighted average of ELSA Wave 6, 7 and 8 survival answers (respectively run in 2012, 2014 and 2016), while the objective probabilities (solid-dot lines) use the ONS life tables of 2012, 2014 and 2016[5]. Consistently with the literature about subjective beliefs, the figure shows a clear pattern: agents underestimate their survival chances when young and overestimate them as they become older. An example of similar results is represented by the work of Elder (2013). The author shows that people aged 80 years or more underestimate their mortality rates.

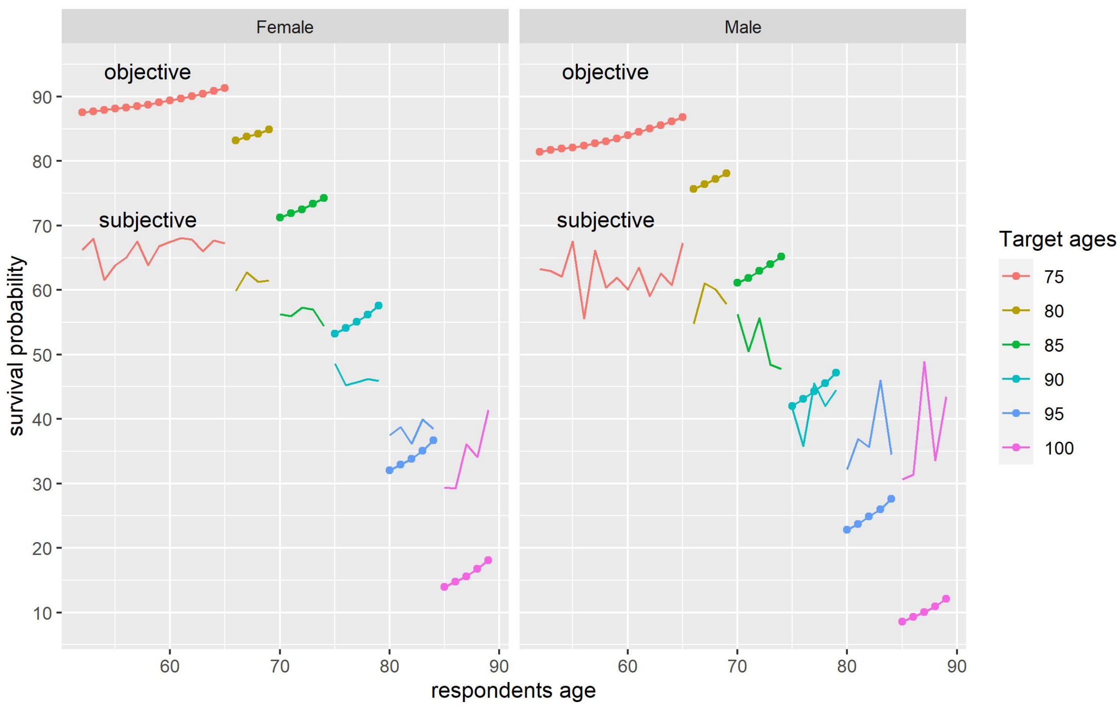
Following these evidence, I construct a survival optimism index. This index represents the proportion of under or overestimation of the survival probability of the agent. I define it as follow:

$$o_{i,t} = \frac{S_{i,t}^{subjective}}{S_{i,t}^{objective}} \tag{3.1}$$

where $S_{i,t}^{subjective}$ is the subjective survival report (ELSA) and $S_{i,t}^{objective}$ is the

[5] I use three ELSA waves subjective survival expectations to reduce the effects of possible outliers. Therefore, the subjective survival probability of each age-sex group is a weighted average of the three waves reports, while the objective survival probabilities are the weighted average of ONS life tables in years 2012, 2014 and 2016. The weights are proportional to the number of observation of each wave in every age-sex group.

Figure 3.3.2: Subjective and objective survival probabilities by target age and gender. ELSA waves 6, 7 and 8 and ONS life table 2012, 2014 and 2016[a].



[a]The objective survival rate (dot lines) are increasing because they represents the survival probability of agents of *increasing* ages to a single, fixed target age (one target age for each color). Therefore, the higher is the agent actual age, the closer he/she is to the target age, the higher is her/his probability to survive to at least that age.

objective survival probability (ONS) of the agent i to the correspondent target age t. Note that $S_{i,t}^{objective}$ depends on two factors: target age and actual age of agent i[6]. Then, if the index $o_{i,t} < 1$, the agent i has pessimistic expectation about his/her own survival probability, otherwise he/she has optimistic expectations.

Table 3.4 shows the average optimism index by target age and gender, and also compares the average subjective and objective survival probability by target age. Survival optimism is increasing in respondents' age, especially among the elderly. Relatively to the objective values, men become optimistic about survival beliefs at

[6]Section A of the Appendix describes the computation of objective survival probabilities.

younger ages than females, but overall there are no large differences in the optimistic behaviours of males and females.

Table 3.4: Survival optimism by target age and gender

	Females			Males		
target age	survival probability subjective	survival probability objective	optimism	survival probability subjective	survival probability objective	optimism
75	66,1%	89,7%	0,73	62,2%	84,5%	0,74
80	62,5%	84,3%	0,74	57,2%	77,3%	0,75
85	54,5%	73,1%	0,75	53,8%	63,5%	0,85
90	51,1%	55,7%	0,91	44,3%	45,0%	0,99
95	43,0%	33,7%	1,28	39,1%	25,0%	1,57
100	42,6%	15,9%	2,71	43,2%	10,4%	4,18

3.3.3 Survival curves

This Section describes the estimation procedure used to compute individual survival curves from subjective survival expectation reports. The procedure needs specific assumptions about the functional form of the individual's survival curve. Indeed, infinite solutions coherent with subjective beliefs are possible if no restrictions are imposed. I follow the work of O'Dea and Sturrock (2021), which assume that the individual survival beliefs follow a Weibull distribution. The Weibull distribution is largely used in survival analysis, epidemiology and ageing process estimation (e.g.: Bissonnette, M. D. Hurd, and Michaud (2017))[7].

The Weibull distribution has two-parameters, λ_i and k_i, the scale parameter and the shape parameter, respectively. The Weibull distribution allows to compute the survival probability $S_i(\alpha)$ of an individual i of age z to the target age α as:

$$S(\alpha) = exp\Big[- \Big(\frac{\alpha - z}{\lambda_i}\Big)^{k_i}\Big] \quad : \quad \lambda_i, k_i > 0 \tag{3.2}$$

[7]A second widely used distribution in these literature is the Gompertz distribution. In their study, Bissonnette, M. D. Hurd, and Michaud (2017) show that the Weibull and the Gompertz distributions lead to similar results.

In other words, Equation 3.2 represents the probability of survival to at least age α of an agent i who is z years old.

To estimate the two Weibull parameters λ_i, k_i of each individual i, O'Dea and Sturrock (2021) make the weak additional assumption that individuals are almost certain not to live beyond age 110 years. In particular, they assume that the agents' survival probability at age 110 years is the one provided by ONS life tables[8]. This assumption implies that individuals aged 70 years or more have two survival points, while those aged 69 years or less have three survival points. This is because the former group answers to only one survival expectation question, while the latter answers to two survival expectation questions.

In the first case, I estimate λ_i and k_i Weibull parameters solving the following system of two equation in two unknowns:

$$\begin{cases} S(\alpha)_{i,z} = exp\Big[-\Big(\frac{\alpha-z}{\lambda_i}\Big)^{k_i}\Big] \\ S(110)_{i,z} = exp\Big[-\Big(\frac{110-z}{\lambda_i}\Big)^{k_i}\Big] \end{cases} \tag{3.3}$$

where α is the target age of agent i, z is the actual age of agent i, $S(\alpha)_{i,z}$ is the subjective survival probability of agent i (ELSA report) and $S(110)_{i,z}$ is the ONS rescaled survival probability of agent i to target age 110 years. Therefore, the only two unknowns variable of the system are the scale and shape parameters of the Weibull distribution.

For those individuals aged 69 or less, I have three different survival points. I follow O'Dea and Sturrock (2021), which estimate λ_i and k_i Weibull parameters with nonlinear least squares. In particular, the estimation procedure minimizes:

$$(\hat{\lambda}_i, \hat{k}_i) = arg\ min_{\lambda_i, k_i} \sum_{\alpha \in A_i} \left(S_i^{subjective}(\alpha) - exp\Big[-\Big(\frac{\alpha - z}{\lambda_i}\Big)^{k_i}\Big]\right)^2 \tag{3.4}$$

where $S_i^{subjective}(\alpha)$ is the subjective survival probability and $A_i = [75, 85, 110]$ is

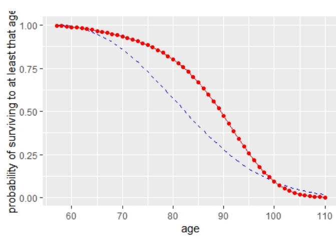
[8] $S_{i,110}^{objective} \in [0.001, 0.003]$.

the vector composed by the three target ages.

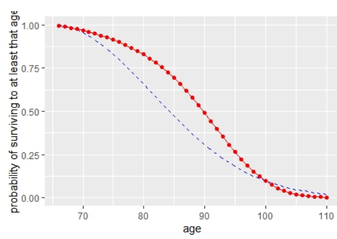
Figures 3.3.3 and 3.3.4 compare the subjective survival curves constructed using the estimated Weibull distribution and the ONS objective survival curves. Each fitted survival curve refers to a specific age and gender group and uses the median value of λ_i, k_i and age of that group. The different target ages of the ELSA questions determine the age intervals used in Figures 3.3.3 and 3.3.4. Because males and females mortality rates differ significantly, the two Figures report them separately.

Figure 3.3.3: Females: objective (dot line) and subjective (dashed line) survival curves. For individuals aged less than 70 years, I use the λ and k Weibull parameters estimated with nonlinear least squares (Equation 3.4). For individuals aged 70 years or more, I use the value of the parameters computed solving the system of two equations (Equation 3.3).

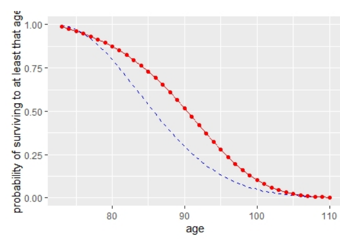
(a) Females aged 50 to 59



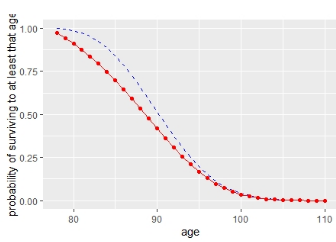
(b) Females aged 60 to 69



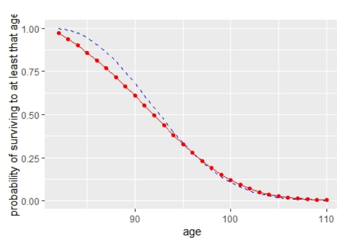
(c) Females aged 70 to 74



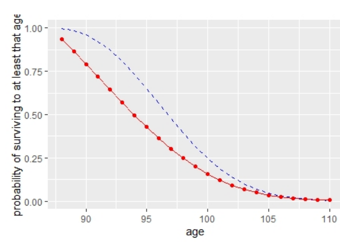
(d) Females aged 75 to 79



(e) Females aged 80 to 84

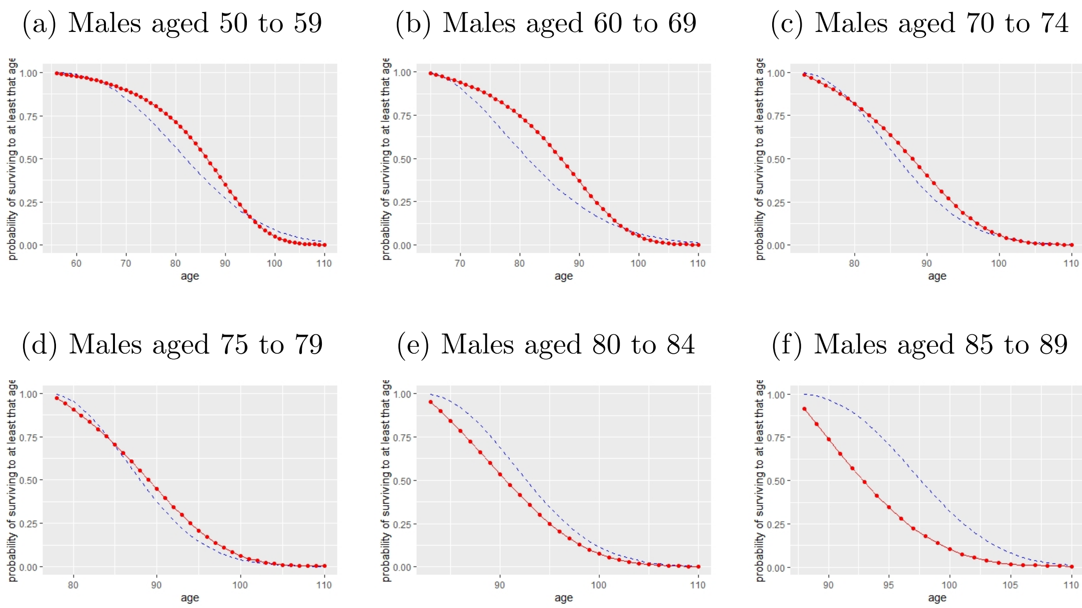


(f) Females aged 85 to 89



A clear pattern emerges from Figures 3.3.3 and 3.3.4: pessimism dominates among young respondents aged 50 to 75 years, while optimism prevails among older agents with no gender differences. Moreover, the subjective survival curves of younger agents (50 to 70 years) show that they underestimate their survival chances until age

Figure 3.3.4: Males: objective (dot line) and subjective (dashed line) curves. For individuals aged less than 70 years, I use the λ and k Weibull parameters estimated with nonlinear least squares (Equation 3.4). For individuals aged 70 years or more, I use the value of the parameters computed solving the system of two equations (Equation 3.3).



(a) Males aged 50 to 59
(b) Males aged 60 to 69
(c) Males aged 70 to 74
(d) Males aged 75 to 79
(e) Males aged 80 to 84
(f) Males aged 85 to 89

95-100 years and then switch to a moderate optimism. Comparing Figures 3.3.3 and 3.3.4, men are less pessimistic than women at a younger age, but females become optimistic about survival probabilities earlier.

Using the estimated Weibull parameters λ_i, k_i, I derive a second survival optimism index similar to the one in Section 3.3.2 as:

$$o^{w}_{i,t} = \frac{S^{Weibull}_{i,t}}{S^{objective}_{i,t}} \qquad (3.5)$$

where $S^{Weibull}_{i,t}$ is the individual subjective survival probability of Equation 3.2 derived from the estimated individual parameters.

3.3.4 Subjective life expectancy

This Section describes the procedure that estimates subjective life expectations using the subjective survival reports.

Life expectation e_i is defined as the average remaining period to live at exact age z and is comparable to the *mean time to failure* (MTTF). The MTTF is a maintenance metric used in engineering to compute the average amount of time a non-repairable asset operates before it fails. Because MTTF is relevant for the equipment that cannot or should not be repaired, it describes the average lifespan of that equipment. Therefore, this concept can be adopted in the current survival analysis and interpreted as the life expectancy of each respondent.
In Section 3.3 I assume that the functional form of the agents' survival probability follows a Weibull distribution. This distribution has a closed form solution for the MTTF:

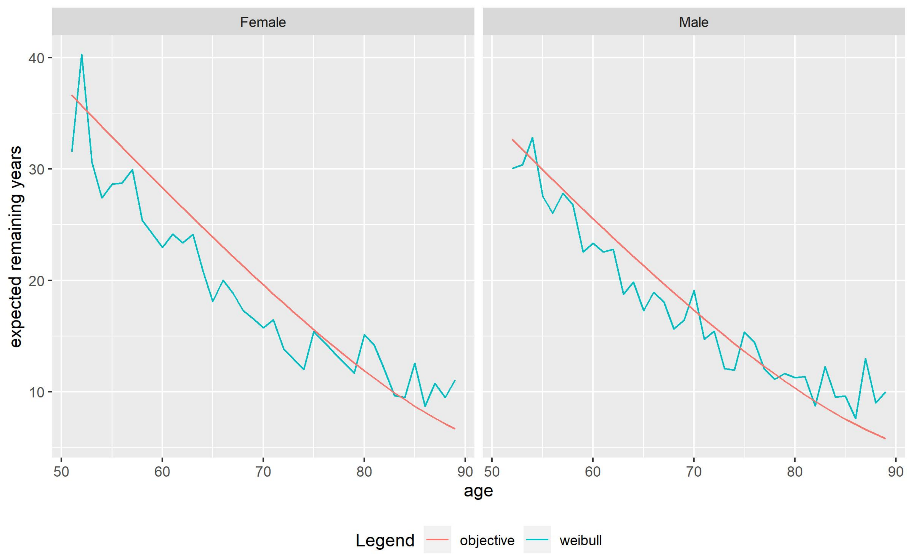
$$MTTF_i = E[X] = \lambda \cdot \Gamma(1 + \frac{1}{k}) \tag{3.6}$$

where X is a generic Weibull random variable and Γ is the gamma function $\Gamma(y) = (y-1)!$. λ and k are the scale and shape parameters of the Weibull distribution.
Then, for each agent i I substitute the estimated Weibull parameters (λ_i, k_i) and I compute e_i subjective life expectancy as:

$$e_i = \lambda_i \cdot \Gamma(1 + \frac{1}{k_i}) \tag{3.7}$$

Figure 3.3.5 compares the ONS objective life expectancy and the median subjective life expectancy by age and gender, where the subjective estimates follow Equation 3.7. The pattern confirms the results of the previous Sections and is coherent with the findings concerning subjective survival probabilities. Indeed, young agents underestimate their life expectancy, while older agents believe that they will live longer than their objective possibilities. Moreover, both males and females become optimistic about survival expectation around 80 years, as widely confirmed in the literature.

Figure 3.3.5: Median subjective vs objective life expectancy by gender and age - Weibull MTTF estimates.



3.4 Survival optimism and stock market participation

This Section presents the empirical results of the paper. I investigate the impact of subjective survival optimism on individual (singles) stock market participation using a probit model, controlling for a set of demographics such as education, age and wealth.

First, I present the econometric model and the estimates using the original survival optimism index, computed as the ratio between subjective survival probabilities and the re-scaled ONS survival chances (Section 3.3.2), and discuss the determinants of individuals stock market participation. Then, two robustness checks follow: the first uses the survival optimism index derived from Weibull estimates (Section 3.3.3) and

the second relies on the estimated subjective life expectancy (Section 3.3.4).
Overall, the estimates show that subjective survival beliefs play a relevant role in individuals portfolio decisions and stock market participation increases in survival optimism.

3.4.1 Results

Individual stock market participation is a dummy variable that identifies whether the individuals hold risky assets[9] or not. Therefore, I use a probit regression to study the determinants of the individual probability of holding risky assets.
The probit models include a set of demographic controls such as a second order age polynomial, gender, income and wealth quartiles dummies, job market participation, education (in three categories, low, mid and high, that depends on the number of years in school), a sickness index[10], numeracy (in four categories: no numeracy, low, mid and high numeracy[11]) and agents subjective survival probabilities. Numeracy can be interpreted as a proxy of cognitive abilities, therefore it may be relevant for household and individual decision making. Indeed, agents with higher cognitive skills might be more precise in their survival report. In other words, the heterogeneity in cognitive skills may explain a portion of the variability of survival probabilities misperception . Moreover, higher numeracy is also associated with better financial choices, as Angelini and Cavapozzi (2017) show.

Table 3.5 reports the probit estimates that relies on the original survival reports of

[9]Risky assets includes shares, bonds, stocks and shares ISAs or life insurance ISAs. ISAs (Individual Saving Account) is a class of retail investment arrangement available to residents of the United Kingdom, with favorable tax condition. They offer four types of account: cash ISA, innovative finance ISA (IFISA), stocks & shares ISA and lifetime ISA.

[10]For further information, see Dal Bianco (2020).

[11]Numeracy is based on a set of questions based on the solution of simple math exercises, like computing percentages, fractions, additions and subtractions. In particular, I use 5 questions that ask to compute subtractions: respondents have to subtract 7 from 100, and then 7 from the previous result and so on, five times. I compute individual numeracy score as the sum of correct answers of the respondent. I aggregate the scores from 2 to 4, obtaining a total of 4 possible categories for each partner: no numeracy skills, low numeracy, medium numeracy, high numeracy.

ELSA participants. The two econometric specifications include the survival optimism index derived using the subjective answers of the survey (see Section 3.3.2). Then, the survival optimism is the ratio between subjective survival reports and ONS rescaled survival probabilities. Column (1) Table 3.5 uses this optimism index as a continuous variable, while in Column (2) Table 3.5 the optimism index is categorized in quartiles dummies.

The estimates of Table 3.5 show that individual stock market participation is increasing in income, wealth and numerical skills (numeracy). On the other hand, low education has negative effects on stock market participation. These findings are in line with the literature that studies household portfolio allocation, its determinants and its puzzles. The assumption of stock market participation costs (see Vissing-Jorgensen (2002)) may partially explain why the wealthier risk more in financial terms. Indeed, those costs do not depend on the returns of the investments but reduce both investment returns and household wealth. Therefore, the wealthier is the household, the lower is the weight of these costs on the total wealth and potential returns. Numeracy approximates the individual cognitive skills. Assuming that agents with low numeracy need more time to improve their financial knowledge means that they have to pay higher entry costs. Therefore, the potential returns of the investments decrease for individuals with low numeracy skills. Then, individuals with low cognitive abilities show a higher rate of non-participation. Similarly, low education negatively affects stock market participation. The sickness index negatively affects the probability of holding risky assets. Poor health conditions reduce the life horizon of the agent and also the investments horizon. Therefore, the agent may decide to not hold stocks because the equity premium manifests for investments of 20 years or more and his/her subjective life horizon is shorter. Last, Column (1) shows positive and significant effects of the (continuous) optimism index, while Column (2) highlights that the optimism effects are probably non-linear. The positive relationship between stock market participation and survival optimism (Column (1)) implies that as subjective survival beliefs decrease, the stock market participation rate decreases too. Table 3.6 helps with the interpretation of the nonlinearity of the optimism index in Table 3.5 Column (2), showing the average age, subjective and objective life expectancy by optimism index quartiles. The fourth quartile of the optimism index is non-significant because it gathers older agents whose expected life is not long enough to gain from the equity premium, even if they show very high optimistic expectation.

Table 3.7 presents the estimates of two robustness checks, which use the subjective survival probabilities estimated from the Weibull parameters and distribution. Column (1) uses the optimism index derived in Section 3.3.3 using the Weibull survival probabilities as numerator, while Column (2) includes the subjective life expectation (the MTTF) computed as described in Section 3.3.4.

Table 3.7 confirms the results of the previous analysis: Column (1) shows positive and significant effects of the (continuous) optimism index, and the coefficient is comparable in magnitude to the one of Column (1) Table 3.5. In Column (2), the positive effect of life expectancy confirms that individual survival beliefs play a role in portfolio decisions and stock market participation. In this case, the variable identifies the number of years that the individual expects to live.

3.5 Conclusion

This paper studies the effect of subjective survival expectation on individual portfolio allocation, proposing a new possible explanation of the well-known "equity premium puzzle". I use the ELSA survey, a longitudinal survey that interviews English people aged 50+, that provides a direct measure of respondents' subjective survival probabilities. Assuming that survival beliefs follow a Weibull distribution, I estimate subjective survival curves and life expectancy, following the work of O'Dea and Sturrock (2021).

The results show that survival beliefs play a relevant role in the portfolio allocation process. Indeed, stock market participation is increasing in subjective survival probabilities and life expectancy. In other words, an agent with a (subjective) long time horizon is more likely to participate in the stock market. The relationship between survival optimism and stock market participation is coherent with the well-established trend that risky assets outperform treasury bills in the long run (20 years).

These results are relevant for the policy maker, financial institutions and financial consultants. Many individuals refuse to hold risky assets while they approach their retirement, and one of the reasons may be the misperception of individual survival probabilities. Correct information from the government or other authorities such as financial institutions may convince some of the non-stockholders to rethink their long term portfolio allocation. Therefore, there is space for policy intervention to correct individual perception of retirement's duration and life expectancy.

This paper spreads new light on the determinants of the portfolio decision process, however, further work is needed especially about the reliability of the data on subjective survival expectations. Indeed, these reports may be heavily affected by heaping and rounding problems. The ELSA survey explicitly asks for integer answers from 0 to 100, and I already discussed that almost 20% of the sample pick the comfortable *50%* answer. Moreover, research such Tversky and Kahneman (1983) and Johnson et al. (1993) show that the probabilities are complicated and systematic biases affect agents' perception. A second point relates to individual expectations: it might be that those who show pessimistic survival expectations are pessimistic in general, and therefore they may have pessimistic expectations about stock returns. Then, these agents may decide not to invest in risky assets because of their overall pessimistic view, because of their survival pessimism or because of a combination of the two. It is a promising agenda for future research the consideration of variables that capture the attitudes of agents, such as personality traits, or expectations about financial returns.

Last, this analysis does not consider bequest reason, that may affects the portfolio choice of the elderly.

Table 3.5: Probit estimates - Individual stock market participation. Subsample: includes those who reported survival chances $s \in (0, 100)$.

participation	optimism: continuous	optimism: quartiles
	(1)	(2)
age	0.220^{***}	0.178^{***}
	(0.060)	(0.057)
age^2	-0.002^{***}	-0.001^{***}
	(0.0004)	(0.0004)
female	-0.021	-0.047
	(0.073)	(0.073)
2^{nd} income quartile	0.018	0.017
	(0.096)	(0.096)
3^{rd} income quartile	0.117	0.126
	(0.096)	(0.096)
4^{th} income quartile	0.417^{***}	0.423^{***}
	(0.101)	(0.101)
2^{nd} wealth quartile	0.729^{***}	0.728^{***}
	(0.104)	(0.104)
3^{rd} wealth quartile	1.092^{***}	1.090^{***}
	(0.106)	(0.106)
4^{th} wealth quartile	1.564^{***}	1.557^{***}
	(0.113)	(0.113)
work	-0.016	-0.019
	(0.106)	(0.106)
low education	-0.407^{***}	-0.401^{***}
	(0.106)	(0.106)
mid education	-0.082	-0.083
	(0.101)	(0.101)
sickness index	-0.038^{***}	-0.036^{***}
	(0.013)	(0.013)
low numeracy	0.378^{**}	0.382^{**}
	(0.179)	(0.180)
mid numeracy	0.405^{**}	0.406^{**}
	(0.186)	(0.187)
high numeracy	0.566^{***}	0.564^{***}
	(0.172)	(0.173)
optimism	0.092^{**}	
	(0.043)	
2^{nd} optimism quartile		0.176^{*}
		(0.101)
3^{rd} optimism quartile		0.244^{**}
		(0.103)
4^{th} optimism quartile		0.101
		(0.112)
Constant	-9.364^{***}	-8.200^{***}
	(2.153)	(2.053)
Observations	1,807	1,807
Akaike Inf. Crit.	1,967.75	1,970.49

Table 3.6: Mean age by quartile of optimism index as of Section 3.3.2.

optimism index quartile	mean age	life expectancy (years) subjective	objective
$1^{st} = 0.55$	73.4	6.5	16.4
$2^{nd} = 0.83$	67.6	17.0	21.1
$3^{rd} = 1.13$	67.3	22.7	21.5
$4^{th} = 11.0$	78.0	18.5	13.1

Table 3.7: Probit estimates - Individual stock market participation. Subsample: includes those who reported survival chances $s \in (0, 100)$.

participation	Weibull: optimism	Weibull: life expectancy
	(1)	(2)
age	0.221^{***}	0.219^{***}
	(0.060)	(0.060)
age^2	-0.002^{***}	-0.001^{***}
	(0.0004)	(0.0004)
female	-0.021	-0.052
	(0.073)	(0.073)
2^{nd} income quartile	0.018	0.016
	(0.096)	(0.096)
3^{rd} income quartile	0.117	0.120
	(0.096)	(0.096)
4^{th} income quartile	0.417^{***}	0.421^{***}
	(0.101)	(0.101)
2^{nd} wealth quartile	0.729^{***}	0.732^{***}
	(0.104)	(0.104)
3^{rd} wealth quartile	1.092^{***}	1.091^{***}
	(0.106)	(0.106)
4^{th} wealth quartile	1.564^{***}	1.562^{***}
	(0.113)	(0.113)
work	-0.017	-0.021
	(0.106)	(0.106)
low education	-0.407^{***}	-0.400^{***}
	(0.106)	(0.106)
mid education	-0.082	-0.082
	(0.101)	(0.101)
sickness index	-0.038^{***}	-0.035^{***}
	(0.013)	(0.013)
low numeracy	0.378^{**}	0.385^{**}
	(0.179)	(0.180)
mid numeracy	0.406^{**}	0.415^{**}
	(0.186)	(0.186)
high numeracy	0.567^{***}	0.574^{***}
	(0.172)	(0.173)
Weibull optimism	0.093^{**}	
	(0.043)	
life expectancy		0.015^{**}
		(0.007)
Constant	-9.411^{***}	-9.893^{***}
	(2.157)	(2.251)
Observations	1,807	1,807
Akaike Inf. Crit.	1,967.605	1,967.779

Appendix
Subjective survival expectation and portfolio decision

A ONS life tables and ELSA target ages

A.1 Number of survivors and probability of surviving

As stated above, l_x represents the number of people alive at exact age x. Generally, the l_x represents a hypothetical population and not a precise population estimate, therefore the l_0 (initial population) is an arbitrary number that the ONS sets to 100,000.

The l_x value is of particular interest for the purpose of this study because it can be used to calculate the survival probability from age x to age $x + n$ as follows:

$$s_{x,x+n} = \frac{survivors\ at\ age\ x+n}{survivors\ at\ age\ x} * 100 = \frac{l_{x+n}}{l_x} * 100$$

where $s_{x,x+n}$ is the expected survival probability of an agent of age x to age $x + n$, at a given year y. Note that I used period life-tables, where the l_x at year y represents the number of surviving at exact age x years in the specific year y, under the projected assumptions for mortality rates in year y for ages up to age x.

An example The following equation is used to calculate the probability of a female aged 40 years in 2018 surviving to age 75 years:

$$\frac{l_{75}}{l_{40}} * 100 = \frac{80,277}{98,595} * 100 = 81.4\%$$

That is, a female aged 40 years in 2018 has a 81.4% chance of surviving to age 75 years.

I will refer to the survival probability computed as described in the example as *objective* survival probability, and to the answers to the ELSA questions as *subjective* survival probability.

Section 3.2.2 explains that the ELSA questions related to life expectancy rely on six different *target ages*, depending on the actual age of the respondent. In Table 1, I provide the objective survival probabilities by gender, age and target age, computed as Section A.1 explains.

Table 1: Objective survival probabilities (ONS) by age (from 40 to 89 years), ELSA target ages and gender in 2016. The survival probability is computed as $\frac{l_{x+n}}{l_x} * 100$, where $(x + n)$ is *target age* and x is *respondent age*.

		Objective survival probability	
Respondent age	Target age	Females	Males
50	75	82,6%	75,0%
51	75	82,8%	75,3%
52	75	82,9%	75,6%
53	75	83,2%	75,9%
54	75	83,4%	76,2%
55	75	83,6%	76,5%
56	75	83,9%	76,9%
57	75	84,2%	77,3%
58	75	84,5%	77,8%
59	75	84,9%	78,3%
60	75	85,3%	78,8%
61	75	85,8%	79,4%
62	75	86,2%	80,1%
63	75	86,8%	80,9%
64	75	87,4%	81,7%
65	75	88,0%	82,6%
66	80	77,0%	68,3%
67	80	77,7%	69,2%
68	80	78,4%	70,2%
69	80	79,2%	71,4%
70	85	61,8%	50,9%
71	85	62,6%	51,9%
72	85	63,4%	53,0%
73	85	64,4%	54,2%
74	85	65,6%	55,6%
75	90	40,2%	29,8%
76	90	41,1%	30,7%
77	90	42,1%	31,8%
78	90	43,4%	33,2%
79	90	44,8%	34,7%
80	95	18,0%	12,0%
81	95	18,7%	12,6%
82	95	19,6%	13,5%
83	95	20,6%	14,4%
84	95	21,8%	15,6%
85	100	5,0%	2,9%
86	100	5,3%	3,2%
87	100	5,8%	3,6%
88	100	6,5%	4,1%
89	100	7,2%	4,7%

Conclusion

The three chapters of this thesis represent an attempt to improve the understanding of how risk factors and risk behaviours influence households' choices and affect their life-cycle.

In Chapter 1, I study the earnings process of households as a source of uninsurable risk. I show that education explains part of the heterogeneity in the household earnings process. In particular, the persistence of past earnings is higher among the high-educated, i.e. high-educated households are less vulnerable to earnings shocks, especially the negative ones. Hence, the uninsurable risk implied by earnings is higher for low-educated households. I test these results by comparing the risk premium (measured as a percentage of consumption) of the high- and the low-educated implied by the distribution of the earnings shocks of each group. Estimates show that the low-educated are willing to pay a risk premium twice as high as the high-educated.

Chapter 2 focuses on the risk behaviours and attitudes of the household decision-makers and their effects on the portfolio allocation decision process. Using a collective approach, I show that the preferences of both spouses matter in the financial choices. In particular, I approximate the risk tolerance of the household as a weighted average of the decision-makers risk preferences, where the weights represent the bargaining power of each member in the decision process. The empirical estimates reveal that the household comprehensive risk tolerance affects the share of wealth allocated in risky assets only if the household has already decided to hold stocks. Further analysis shows that the proposed collective approach performs better than the standard unitary model, which describes the households as a single decision unit.

Last, in Chapter 3 I investigate the role of subjective survival beliefs in individual portfolio decisions. According to the literature, most young agents are pessimistic about their survival probability and become optimistic when they get older. I show that survival expectations are relevant for individual portfolio choices. The higher is the survival optimism of the agents, the longer is the (expected) life horizon. Consequently, also the time horizon of the (potential) investments increases. Then, the individuals (think to) have higher chances to gain from the equity risk premium, which is positive for long periods. The empirical results show a positive relationship between survival optimism and stock market participation, especially among the young agents in the sample.

Bibliography

Abdellaoui, Mohammed, Olivier l'Haridon, and Corina Paraschiv (2013). "Individual vs. couple behavior: an experimental investigation of risk preferences". In: *Theory and decision* 75(2), pp. 175–191.

Abowd, John M. and David Card (1989). "On the covariance structure of earnings and hours changes". In: *Econometrica* 57(2), pp. 411–445.

Addoum, Jawad M, Howard Kung, and Gonzalo Morales (2016). "Limited marital commitment and household portfolios". In: *Available at SSRN 2871296*.

Ameriks, John and Stephen P Zeldes (2004). *How do household portfolio shares vary with age*. Tech. rep. working paper, Columbia University.

Angelini, Viola and Danilo Cavapozzi (2017). "Dispositional optimism and stock investments". In: *Journal of economic psychology* 59, pp. 113–128.

Apps, Patricia F and Ray Rees (1997). "Collective labor supply and household production". In: *Journal of political Economy* 105(1), pp. 178–190.

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme (2017). "Earnings and consumption dynamics: a nonlinear panel data framework". In: *Econometrica* 85(3), pp. 693–734.

Arrow, K.J. (1965). *Aspects of the theory of risk-bearing*. Yrjö Jahnsson lectures.

Attanasio, Orazio P and Valérie Lechene (2014). "Efficient responses to targeted cash transfers". In: *Journal of political Economy* 122(1), pp. 178–222.

Banks, James and Sarah Smith (2000). "UK household portfolios". In: *Household portfolios*. Ed. by Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT press.

Benartzi, Shlomo and Richard H Thaler (1995). “Myopic loss aversion and the equity premium puzzle”. In: *The quarterly journal of Economics* 110(1), pp. 73–92.

Bertocchi, Graziella, Marianna Brunetti, and Costanza Torricelli (2014). “Who holds the purse strings within the household? The determinants of intra-family decision making”. In: *Journal of Economic Behavior & Organization* 101, pp. 65–86.

Bissonnette, Luc, Michael D Hurd, and Pierre-Carl Michaud (2017). “Individual survival curves comparing subjective and observed mortality risks”. In: *Health economics* 26(12), e285–e303.

Bloom, David E et al. (2006). *The effect of subjective survival probabilities on retirement and wealth in the United States.* Tech. rep. National Bureau of Economic Research.

Blundell, Richard W et al. (2016). “The dynamic effects of health on the employment of older workers”. In: *Michigan Retirement Research Center Research Paper*(2016-348).

Bonhomme, Stéphane and Jean-Marc Robin (2009). “Assessing the equalizing force of mobility using short panels: France, 1990–2000”. In: *The Review of Economic Studies* 76(1), pp. 63–92.

Bresser, Jochem de (2020). *The role of heterogeneous expectations in life cycle models.*

Browning, Martin et al. (1994). “Income and outcomes: A structural model of intrahousehold allocation”. In: *Journal of political Economy* 102(6), pp. 1067–1096.

Caliendo, Marco, Frank M Fossen, and Alexander S Kritikos (2009). “Risk attitudes of nascent entrepreneurs–new evidence from an experimentally validated survey”. In: *Small business economics* 32(2), pp. 153–167.

Cappellari, Lorenzo (2002). “Earnings dynamics and uncertainty in Italy: how do they differ between the private and public sectors?” In: *Labour Economics* 9(4), pp. 477–496.

Cappellari, Lorenzo (2004). “The Dynamics and Inequality of Italian Men’s Earnings Long-term Changes or Transitory Fluctuations?” In: *Journal of Human Resources* 39(2), pp. 475–499.

Chang, Yongsung, Jay H Hong, and Marios Karabarbounis (2018). “Labor market uncertainty and portfolio choice puzzles”. In: *American Economic Journal: Macroeconomics* 10(2), pp. 222–62.

Charness, Gary and Matthias Sutter (2012). “Groups make better self-interested decisions”. In: *Journal of Economic Perspectives* 26(3), pp. 157–76.

Cherchye, Laurens, Bram De Rock, and Frederic Vermeulen (2007). “The collective model of household consumption: a nonparametric characterization”. In: *Econometrica* 75(2), pp. 553–574.

Chiappori, Pierre-André (1992). “Collective Labor Supply and Welfare”. In: *Journal of Political Economy* 100(3), pp. 437–467. ISSN: 00223808, 1537534X. URL: http://www.jstor.org/stable/2138727.

Chiappori, Pierre-André (1988). “Rational Household Labor Supply”. In: *Econometrica* 56(1), pp. 63–90. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1911842.

Chiappori, Pierre-Andre and Maurizio Mazzocco (2017). “Static and intertemporal household decisions”. In: *Journal of Economic Literature* 55(3), pp. 985–1045.

Cooper, Russell and Guozhong Zhu (2016). “Household finance over the life-cycle: What does education contribute?” In: *Review of Economic Dynamics* 20, pp. 63–89.

Dal Bianco, Chiara (2020). “Disability insurance and the effects of return-to-work policies”. In: *(mimeo)*.

De Palma, Andre, Nathalie Picard, and Anthony Ziegelmeyer (2011). “Individual and couple decision behavior under risk: evidence on the dynamics of power balance”. In: *Theory and decision* 70(1), pp. 45–64.

Elder, Todd E (2013). “The predictive validity of subjective mortality expectations: Evidence from the health and retirement study”. In: *Demography* 50(2), pp. 569–589.

Fouarge, Didier, Ben Kriechel, and Thomas Dohmen (2014). "Occupational sorting of school graduates: The role of economic preferences". In: *Journal of Economic Behavior & Organization* 106, pp. 335–351.

Gan, Li, Michael D Hurd, and Daniel McFadden (2007). *12. Individual Subjective Survival Curves*. University of Chicago Press.

Geweke, John and Michael Keane (2000). "An empirical analysis of earnings dynamics among men in the PSID: 1968–1989". In: *Journal of econometrics* 96(2), pp. 293–356.

Gomes, Francisco, Michael Haliassos, and Tarun Ramadorai (2020). "Household finance". In: *Journal of Economic Literature, forthcoming*.

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

Gu, Ran, Cameron Peng, and Weilong Zhang (2021). "The Gender Gap in Household Bargaining Power: A Portfolio-Choice Approach". In: *https://doi.org/10.17863/CAM.69991*.

Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (2003). "Household stockholding in Europe: where do we stand and where do we go?" In: *Economic Policy* 18(36), pp. 123–170.

Guvenen, Fatih et al. (2019). *What do data on millions of US workers reveal about life-cycle earnings risk?* Tech. rep. Available at `https://fguvenendotcom.files.wordpress.com/2019/02/gkos_2017_rev3_submit.pdf`.

Haliassos, Michael and Carol C Bertaut (1995). "Why do so few hold stocks?" In: *The Economic Journal* 105(432), pp. 1110–1129.

Heckman, James J (1979). "Sample selection bias as a specification error". In: *Econometrica: Journal of the econometric society*, pp. 153–161.

Hurd, Michael, Daniel McFadden, and Li Gan (2008). "9. Subjective survival curves and life cycle behavior". In: *Inquiries in the Economics of Aging*. University of Chicago Press, pp. 259–310.

Hurd, Michael D and Kathleen McGarry (1995). “Evaluation of the subjective probabilities of survival in the health and retirement study”. In: *Journal of Human resources*, S268–S292.

Hurd, Michael D and Kathleen McGarry (2002). “The predictive validity of subjective probabilities of survival”. In: *The Economic Journal* 112(482), pp. 966–985.

Jappelli, Tullio and Mario Padula (2015). “Investment in financial literacy, social security, and portfolio choice”. In: *Journal of Pension Economics & Finance* 14(4), pp. 369–411.

Johnson, Eric J et al. (1993). “Framing, probability distortions, and insurance decisions”. In: *Journal of risk and uncertainty* 7(1), pp. 35–51.

Kelly, Morgan (1995). “All their eggs in one basket: Portfolio diversification of US households”. In: *Journal of Economic Behavior & Organization* 27(1), pp. 87–96.

Leeuwen, Bart van, Rob Alessie, and Jochem de Bresser (2020). “Household Composition and Preferences: A Collective Approach to Household Consumption”. In: *Review of Income and Wealth*.

Lusardi, Annamaria (2008). *Household saving behavior: The role of financial literacy, information, and financial education programs*. Tech. rep. National Bureau of Economic Research.

Manski, Charles F (2004). “Measuring expectations”. In: *Econometrica* 72(5), pp. 1329–1376.

Meghir, Costas and Luigi Pistaferri (2004). “Income Variance Dynamics and Heterogeneity”. In: *Econometrica* 72(1), pp. 1–32.

Mehra, Rajnish (2007). *The equity premium puzzle: A review*. Now Publishers Inc.

Mehra, Rajnish and Edward C Prescott (1985). “The equity premium: A puzzle”. In: *Journal of monetary Economics* 15(2), pp. 145–161.

Merton, Robert C (1969). “Lifetime portfolio selection under uncertainty: The continuous-time case”. In: *The review of Economics and Statistics*, pp. 247–257.

Mussida, Chiara and Matteo Picchio (2014). "The trend over time of the gender wage gap in Italy". In: *Empirical Economics* 46(3), pp. 1081–1110.

Nivakoski, Sanna (2020). "Wealth and the effect of subjective survival probability". In: *Journal of Population Economics* 33(2), pp. 633–670.

O'Dea, Cormac and David Sturrock (2021). "Survival pessimism and the demand for annuities". In: *forthcoming in Review of Economics and Statistics*.

Olafsson, Arna and Tomas Thornquist (forthcoming). "Bargaining over risk: The impact of decision power on household portfolios". In: *Review of Finance*.

Peracchi, Franco, Valeria Perotti, et al. (2010). "Subjective survival probabilities and life tables: Evidence from Europe". In: *Einaudi Institute for Economics and Finance (EIEF) Working Paper* 10, p. 16.

Poterba, James, Steven Venti, and David A Wise (2013). "Health, education, and the postretirement evolution of household assets". In: *Journal of Human Capital* 7(4), pp. 297–339.

Pratt, John W. (1964). "Risk Aversion in the Small and in the Large". In: *Econometrica* 32(1/2), pp. 122–136. ISSN: 00129682, 14680262.

Samuelson, Paul A (1975). "Lifetime portfolio selection by dynamic stochastic programming". In: *Stochastic Optimization Models in Finance*, pp. 517–524.

Tversky, Amos and Daniel Kahneman (1983). "Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment." In: *Psychological review* 90(4), p. 293.

Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie (2011). "Financial literacy and stock market participation". In: *Journal of Financial Economics* 101(2), pp. 449–472.

Van Solinge, Hanna and Kène Henkens (2010). "Living longer, working longer? The impact of subjective life expectancy on retirement intentions and behaviour". In: *European journal of public health* 20(1), pp. 47–51.

Veld-Merkoulova, Yulia V (2011). "Investment horizon and portfolio choice of private investors". In: *International Review of Financial Analysis* 20(2), pp. 68–75.

Vissing-Jorgensen, Annette (2002). *Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures.* Tech. rep. National Bureau of Economic Research.

Wachter, Jessica A and Motohiro Yogo (2010). "Why do household portfolio shares rise in wealth?" In: *The Review of Financial Studies* 23(11), pp. 3929–3965.

Zhang, Weiwei (2017). "Household risk aversion and portfolio choices". In: *Mathematics and Financial Economics* 11(3), pp. 369–381.