UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M. FANNO"

CORSO DI LAUREA MAGISTRALE IN ECONOMICS AND FINANCE

TESI DI LAUREA

"Subjective survival expectations and equity market participation: a study conducted across fifteen different countries"

RELATORE:

CH.MA PROF./SSA

Chiara Dal Bianco

LAUREANDO: Alberto Rossini

MATRICOLA N. 2036092

ANNO ACCADEMICO 2022 – 2023

Dichiaro di aver preso visione del "Regolamento antiplagio" approvato dal Consiglio del Dipartimento di Scienze Economiche e Aziendali e, consapevole delle conseguenze derivanti da dichiarazioni mendaci, dichiaro che il presente lavoro non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere. Dichiaro inoltre che tutte le fonti utilizzate per la realizzazione del presente lavoro, inclusi i materiali digitali, sono state correttamente citate nel corpo del testo e nella sezione 'Riferimenti bibliografici'.

I hereby declare that I have read and understood the "Anti-plagiarism rules and regulations" approved by the Council of the Department of Economics and Management and I am aware of the consequences of making false statements. I declare that this piece of work has not been previously submitted – either fully or partially – for fulfilling the requirements of an academic degree, whether in Italy or abroad. Furthermore, I declare that the references used for this work – including the digital materials – have been appropriately cited and acknowledged in the text and in the section 'References'.

Firma (signature) ……Alberto Rossini……

Abstract

Is the underestimation of lifetime horizon a reason for the low participation in stock markets?
Using wave 6 of SHARE, I analyse this empirical relationship across 15 different countries. Interestingly, it emerges a great heterogeneity in how participants self-evaluate their lifetime horizon. I find that stock market participation and underestimation are negatively correlated, the more you underestimate your life the less you participate in the market. Precisely, a 1 standard deviation increase in the underestimation lowers by about 1% the stock market participation. This effect seems to be stronger for the younger participants in my sample (50-64 years), where the reduction in stock market participation associated with underestimation is almost 2%.
I extend the analysis to the equity share invested in the market. Accounting for endogenous sample selection, I study the effect of underestimation on equity share. In opposition to what found by Spaenjers and Spira (2015), in my sample underestimation does not affect the equity share.
Finally, I analyse whether there are differences in the effect of underestimation on both share invested and stock market participation in respondents with bequest motives. In both cases the effect goes in the expected direction but it is not significant, as opposed to what found by Spaenjers and Spira (2015) for the US.

Index

Acknowledgments

This paper uses data from SHARE Waves 1, 2, 3, 4, 5, 6, 7, 8 and 9 (DOIs: 10.6103/SHARE.w1.800,10.6103/SHARE.w2.800, 10.6103/SHARE.w3.800, 10.6103/SHARE.w4.800, 10.6103/SHARE.w5.800, 10.6103/SHARE.w6.800, 10.6103/SHARE.w7.800, 10.6103/SHARE.w8.800, 10.6103/SHARE.w8ca.800, 10.6103/SHARE.w9ca800) see Börsch-Supan et al. (2013) for methodological details.(1) The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

Chapter 1

1.1 Introduction

Time horizon matters in financial decisions. Two strands of the existing economic literature, one related to the myopic loss aversion and the other related to the mean reverting characteristics of asset returns (they will be discuss in detail in section 2) testify that individuals take into account or they are strongly influenced by the time horizon when they have to make economic or financial choices, in particular a longer time horizon tends to make assets appear safer and it allows individuals to mitigate their risk aversion over that extended period.

Lot of people in recent times increased their interests toward the financial world but it still remains unexplained why such a low stock market participation persists. For example, whereas Nordic countries such as Denmark, Switzerland and Sweden show an higher stock market participation, data suggests that in Mediterranean countries, such as Italy and Spain, only 10% hold any stocks, fifteen percent of Germans invest directly in the stock market and seventeen percent in the Netherlands Kaustia et al. (2022). This low stock market participation became even more difficult to explain if I think about to the famous equity premium puzzle. The large premium in favour of stocks should reflect a higher participation in financial markets but the stock market participation is way distant from the optimum and this is still particularly puzzling. Many researchers have tried to explain this low stock market participation phenomenon making use of cultural factors[1] (lack of trust, lack of financial literacy) and risk aversion. Another potential factor, rather unexplored in the literature, is individual (perceived) time horizon. In this thesis, I focus on the role of time horizon on the financial and economic decision making process of individuals and, in particular on their stock market participation.

This argument leads us to the most important question of this thesis: is there a relationship between individuals' inclination to participate in the stock market and their longevity expectations?

While there is evidence in the literature that the horizon affects economic choices it's worth noting that the literature investigating the relationship between stock market participation and one's lifetime horizon is quite limited.

[1] Van Rooij, Lusardi, Alessie (2012). Guiso, Sapienza, Zingales (2008).

One of the most challenging sides of this work is to find a measure to proxy individuals' perceived life-time horizon. One of the most used to measure the residual life span are the subjective probability to survive to a certain age or the subjective life expectancy.

I study the empirical relationship between subjective survival probabilities and stock market participation in the European context. One of the advantages in working with a large set of countries is to take benefits from their heterogeneity. As we will see in the part of the descriptive analysis, subjective survival expectations display a significative heterogeneity across countries and one of the aims of this thesis is to understand if this heterogeneity explains the heterogeneity in the stock market participation.
Spaenjers and Spira (2015) investigate the same relationship for the US context. They use data from U.S. households, and their main empirical analysis tests the relationship between subjective life horizon and equity holdings. The main finding of their analysis is that longer subjective life horizons are associated with higher equity shares.

I use data from SHARE (Survey of Health Ageing and Retirement in Europe) and more precisely from wave 6 (the fieldwork of the Sixth Wave of SHARE was completed in November 2015).
I focus on people between 50 and 90 years old, from 15 different countries. I study the patterns of subjective survival probabilities in each country, separately by gender and age, I then compare the subjective probabilities with objective probabilities that I calculated through the life table of each country. Across countries, I find a significance heterogeneity in terms of subjective and objective probabilities. Data show a marked pattern of correlation between stock market participation and subjective survival probabilities which I investigate more formally through a regression model.
It emerges that in countries such as Denmark and Sweden, market participation is high and this coincides with a tendency to overestimate survival probabilities, while in countries such as the Czech Republic and Estonia, market participation is among the lowest in my sample and coincides also with the lowest subjective survival probabilities.
From the empirical analysis emerges that subjective survival probabilities play a role in stock market participation, the more you underestimate your survival probabilities, the less you invest in the market. Differently from Spaenjers and Spira (2015), the subjective survival probabilities do not seem to influence the share of wealth invested in equity market.

The research thesis is organized as follows: the remaining of this chapter presents the related literature. Chapter 2 reports details on the data used and provides descriptive analyses.

In chapter 3 I present estimation results to answer the research question. First, I focus on the relationship between stock market participation and subjective survival. Then, I investigate whether subjective survival also affects the share of wealth invested in risky assets (accounting for selection into stock market participation). Finally, I investigate whether the presence of bequest motives attenuates the role of subjective survival on stock market participation.

1.2 Related Literature

Stock market participation is puzzling around all the world, there is a large equity premium in favour of stocks but people still prefer to not invest in the stock market or they invest but way less with respect to the optimum allocation. Mehra and Prescott (1985) were the first in explaining this puzzle to the world demonstrating that since 1926 the annual real return of stocks was 7% compared with 1% in treasury bills, this high equity premium is difficult to be explained by a plausible level of investors risk aversion. Academic research in this field tried to explain such an interesting phenomenon, for both the United States and the European countries, for which we have a considerable equity premium estimation. Dimson, Marsh and Staunton (2011) reported a global evidence of equity risk premium in 19 different countries, in a period that goes from 1900 to 2011 and while they've found considerable variations across countries something remains stable: equity risk premium was sizeable everywhere. For example, in the period analysed from Dimson et al. (2011), the geometric mean of the equity risk premium relative to bills was 6 percent for France, 5.8 percent for Italy, 5.9 percent for Germany, 2.8 percent for Denmark, 3.2 percent for Spain with a general value for the Euro area of 3.8 percent. I deliberately reported premiums from these countries because they're also part of the analysis of this research paper.

Despite this equity premium in favour of stocks there is still a great heterogeneity in the stock market participation that is way distant from the optimum in all the Europe. Kaustia et al. (2022) analysed what drives stock market participation, they analysed the stock market participation in 19 European countries using a sample from the first four waves of SHARE data (from 2004 to 2013). They find large heterogeneity.

Considering total participation (defined as direct participation plus indirect participation through mutual funds ira's) only Nordic countries such as Sweden, Denmark, Switzerland show highest rates above fifty percent. In comparison countries from the Mediterranean area as Spain and Italy show only ten percent rate while France, Germany and Austria respectively 30 percent, 30 percent and 20

percent. Typically, about half of total participation comes from indirect stock investments. When considering direct stock market participation alone, very low rates are documented.

Of course, one of the first cultural factors that could be a reason of low stock market participation is the lack of financial education. Lusardi and Mitchell (2007) reveal that many households are not aware of basic economic and financial concepts and this translates into questionable choices when it comes to making important economic decisions for their future.
The limited financial literacy and education can deter individuals from participating in the stock market, lack of understanding about how the stock market works and the potential benefits of long-term investing can discourage people from investing. It's experimentally proved Billari, Favero and Saita (2023) that a literacy intervention program increased the financial and demographic survival literacy of the participants pushing them towards seeking more information and becoming more active in financial decisions. The paper is about a financial program the "finlife" implemented in an employer-based pension fund in Italy. The program's effects are assessed through an experiment that consists in an online seminar and a randomized experimental design to evaluate the short-term impact of the online treatment on financial and survival literacy and to see what are the subsequent investment decisions within the pension fund. The results of the program are a statistically significant increase in financial and survival literacy which consist in a consequent increase of behaviours toward information related to financial planning and a higher probability of changing investment lines within the pension fund. Overall, the experiment shows that the program can have a significant impact on knowledge and behaviour.
Programs that help to increase the financial knowledge can lead to a higher and conscious equity market participation.

From the empirical and economic point of view rather than the cultural, one of the most interesting possible explanations to the stock market participation puzzle is the myopic loss aversion by Benartzi and Thaler (1995), the explanation they proposed is based on two concepts from the psychology of the decision making process. The initial concept to consider is loss aversion, which refers to being more sensitive to losses with respect to gains. This concept is of great significance in the descriptive theory of decision making under uncertainty of Kahneman and Tversky (1979).
In order to well understand this loss aversion topic a small example could help. Consider this old problem posed by the nobel prize winner Samuelson (1968): "you have to flip a coin if you get head you win 200$, if you get tail you lose 100$. Do you accept the bet?" It turns out that most people were not willing to take the bet because for many of them the pain of losing 100$ is greater than the pleasure of winning 200$, this sentiment is the intuition that stands behind the concept of loss

aversion. Losses hurt twice the emotional impact of the pleasure derived from gains even when potential loss is small and does not pose much risk.

Another side of the myopic loss aversion is the evaluation period and the effective investment horizon: an investor with an evaluation of one year so an investor that checks his gain/losses every year behaves as if he had an investment horizon of the same length even if he invests to retire in 30 years. So artificially this investment horizon is way lower and the fact of always checking the account it makes you perceive the market much more volatile than it is in the long term. So it could be that this equity premium was so large and investors irrationally stayed out of the equity market because they look at stock price too much frequently and they conclude that stock market is too much volatile for them, without knowing that the volatility in the long term is much more reduced and stock returns are much safer.

This logic is a first suggestion that investment time horizon matters in the participation in the stock market.

The longer an investor plans to hold an asset without evaluating it frequently, the more interesting the asset will look. In other words, an investor is not willing to accept the risk of equity market participation because of to two issues: loss aversion and short evaluation period. Benartzi and Thaler refers to this combination as myopic loss aversion.

Other reasons to the equity premium puzzle are related to liquidity constraints and ageing effects. It could be that young people are more informed and attracted by the stock market so that they would like to invest in stock also because in the long run stocks are safer than in the short run and they have more time with respect to old people, but usually young people cannot invest consistently in the market due to liquidity constraints. Therefore, stocks are accessible mainly to old people with a shorter residual expected life horizon and for that they invest less in the market.

In order to understand why the lifetime horizon matters in the participation in the market of risky assets some evidence about the benefits of investing with a longer horizon should be discussed.

There is a large literature that testifies the predictability of asset returns, for example Barberis et al (2000) and Campbell and Viceira (2002). They not only have demonstrated that asset returns are predictable but more specifically that asset returns are mean reverting. Mean reversion in the context of stock returns refers to the tendency of a stock's price or returns to move back toward its historical average or mean over time. In other words, when stock returns are mean-reverting, they exhibit a pattern where, after experiencing periods of above-average returns or below-average returns, they are likely to return to their long-term average levels. When stock returns are mean-reverting, they tend to oscillate around a central value or average. If a stock has a historically high return compared to its average, it is more likely to have lower returns in the future, and vice versa.

Mean reversion implies that stocks are safer in the long run i.e. their volatility reduces as time increases; Barberis (2000) using a basic VAR framework for U.S. stock returns demonstrates that the 10-year returns standard deviation is very different from that of monthly returns: in the first case it is 27.3 percent and in the second it is 45.2 percent. Also Campbell and Viceira (2002) showed that U.S. stock returns change their standard deviation from 18 percent value at a one year horizon to 14 percent value at a twenty-five year horizon. These are all papers about the U.S. stock market but there is also evidence of a similar pattern in Europe, in fact Bec and Gollier (2005) tried to extend the above findings to France. They use quarterly data from 1970 to 2006 to perform a VAR analysis of real asset returns on French financial markets. They follow the same analysis of Campbell and Viceira (2002) and they find very similar conclusions. Stock returns exhibit mean reversion with an intensity that is stronger with respect to the one of U.S. data. More precisely, the French stock returns cuts their annualized standard deviation from 22 percent for a one-year horizon to only 2.8 percent for a twentyfive year investment horizon.

Considering the above reported evidence, mean-reversion implies that stocks are safer in the long run and this suggests that a long horizon agent should have a larger demand for stocks.

Summing up, two main arguments in the recent research are telling the importance of horizon in the optimal risky asset allocation: loss aversion (for a loss-averse individual the optimal risk exposure increases with the investment horizon) and mean reverting asset returns (since stocks are mean reverting the longer the time horizon the safer the stocks).

Since the horizon has been defined as something crucial for the stock market participation and the literature does not particularly focus on the effect of lifetime horizon on economic choices, it would be interesting to study if there is an effective relationship between the expected subjective survival probabilities and the stock market participation.

There are couple of papers standing on how portfolios vary with age, for example Fagereng et al (2013) study asset market participation and portfolio choice over life cycle for a large sample of Norwegian households, or Poterba and Samwick (2001) study the relationship between age and portfolio structure for households in United States, but they do not investigate the pure effect of the remaining life span on the household financial decision making processes.

First of all, it's useful to revise the existing literature on subjective probabilities in order to understand if they're reliable and if they're a good proxy in predicting future participants characteristics.

For example, Hurd (2009) illustrates that there has been a growing trend in the recent years towards an increased reliance on subjective probabilities. Now they are widely used in household questionnaires to understand the decision making processes. Furthermore, the intriguing aspect is that when comparing subjective probabilities to real-life outcomes, it becomes evident that these probabilities possess a significant predictive capacity, particularly in situations where individuals

possess substantial private information, such as those related to survival and retirement. In this case, it seems that subjective probabilities are more effective in predicting mortality with respect to other situations.

Another confirms about the precision and the reliability of subjective survival probabilities comes from Perozek (2008). Numerous mortality predictions rely on historical data but they do not take into account information about potential changes in the factors due to fatality hazards over time. Perozek claims that individuals, through personal knowledge about their health status, the health status of their family members and through their own risk assessment, are the only ones who can predict their remaining life with certainty. Additionally, they can help policymakers studying and designing mortality rates by sharing this information.

There is a great advantage in the author's approach which is to take into consideration personal characteristics such as diet, smoking habits, physical condition. Given that lot of this information are private, only individuals can evaluate how these factors will influence their personal mortality risk, which is a function of their medical history, current health status, and family history.

Finally, an interesting paper testifying that people are particularly precise in predicting their own remaining life span.

This research conducted by Smith et al (2011) aims to assess how accurate and reliable are people's expectations about their own remaining life span. They did the analysis using four waves of the Health and Retirement Survey (HRS) and they tested whether longevity expectations were in line with actual mortality.

This research expresses that individuals are quite accurate in predicting their own demise. Moreover, the individuals who eventually experienced mortality were those who expressed the lowest life expectancies. Another interesting finding is that people who experienced mortality had a decline in subjective survival expectations, while among the survivors, subjective survival expectations remained stable. Finally, it has been discovered a negative relationship between subjective survival expectations and health shocks or increases in individual functional limitations. This means that an individual's longevity expectation is a fairly accurate index of personal survival probability, both in its responsiveness to events that experts would suggest increase the odds of death and as a prediction of future mortality.

Now that I reported evidence of the subjective survival probabilities as a good measure in predicting how people will live, it's essential to mention some paper that analysed the relationship between this subjective survival probabilities and economic choices. For example, Bloom et al (2006) studied the relationship between the effect of subjective survival probabilities on retirement and wealth in the United States. They investigate the idea that expected longevity affects retirement decisions and accumulated wealth using data from the Health and Retirement study for the United States. To

evaluate expected longevity, they rely on individual's personal belief of the probability to reach age 75. They find a positive relationship between life expectancy and wealth accumulation among couples. Couples with higher subjective survival probabilities have higher wealth accumulation.

While existing research has uncovered evidence regarding a relationship between subjective probabilities and economic choices, these papers do not consider portfolio choice. This is one of the reasons that makes distinctive the research question of the present thesis.

Chapter 2

Descriptive analysis

This study uses data on households from the SHARE (Survey of Health Ageing and Retirement in Europe) datasets. Share is a research infrastructure for studying the effects of health, social, economic and environmental policies over the life-course of European citizens and beyond.
SHARE is a longitudinal study started in 2004, conducted every two years on a sample of individuals aged 50 and above. In this analysis I consider data from the wave 6 that has been conducted in 2015. The following countries were part of the survey: Austria, Belgium, Croatia, Czech Republic, Denmark, Estonia, France, Germany, Greece, Israel, Italy, Luxembourg, Poland, Portugal, Spain, Switzerland, Sweden, and Slovenia. The SHARE interview consists of various thematic blocks or modules covering information on demographics and network, social network, children, assets, household income, consumption, mental health, financial transfers, expectations and activities.

One of the most relevant for this thesis is the section dedicated to the expectations in which questions related to future aspects of an individual life could be found. This section contains questions related to longevity expectations.
In the following section I describe the expectation variable and the other variables that I will use in the empirical analysis in Chapter 3.

2.1 Variables

The first variable I want to describe is of significance importance, the one related to the subjective survival expectations. The variable is included in the expectation section, and it's called *ex009_*. In *ex009_* baseline respondents are asked "What are the chances that you will live to be age <fill> or more?". The <fill> used in this question is a function of the age of each respondent.

Respondents' age:	Fill in ex009_:
age < 65	75
age > 64 and age < 70	80
age > 69 and age < 75	85
age > 74 and age < 80	90
age > 79 and age < 85	95
age > 84 and age < 95	100
age > 94 and age < 100	105
age > 99 and age < 105	110
age > 104	120

The <fill> used in each interview is stored in the variable *ex009age*, while the substantive answer to the question is stored as *ex009_*.
For example, to an individual that is 52 years old it is asked "what are the chances that you will live to be age 75 or more?" And the individual response is in a scale between 0 and 100 representing the probability to be alive at that target age.
To an individual that is 73 years old it is asked "what are the chances that you will live to be age 90 or more?" and so on with all the other ages.

In rare cases (e.g. if age was ex post corrected due to an interviewer remark) the target age might deviate from this rule. For example, a 65 year old individual is not asked for the target age of 80 (as it should be according to the rule) but for another age. In order to have the clearest dataset I dropped these rare cases in which the target age questioned to the individuals does not respect the rule. Precisely, I dropped 7,110 observations. Most of them (about 5,000) are related to the 75 and 80 target age.

***Table 1: Definitions**. This table presents definition of control and dependent variables used in this thesis.*

Variable	Description	Values
age	Age	
age2	Age squared	
Male	Sex of the respondent	Male = 1 ; Female = 0
couple	Is the respondent in a couple?	Couple = 1 ; Single = 0
child	Does the respondent have at least one children?	yes = 1 ; no = 0
emp	Employment status	Employee or self employed = 1 ; otherwise = 0
educ	Education level grouped in four categories: 1. very low educ (isced code 0-1) ; 2. low educ (isced code 2) ; 3. medium educ (isced code 3-4) ; 4. high educ (isced code 5-6)	
hhquart	Generated quartile household income, considering the total household income, each household assigned to the respective quartile	
adl	Activities of daily living	Fom 1 to 10, the higher the rate the higher the limitations
Austria	Living in Austria	
Germany	Living in Germany	
Sweden	Living in Sweden	
Spain	Living in Spain	
Italy	Living in Italy	
France	Living in France	
Denmark	Living in Denmark	
Switzerland	Living in Switzerland	
Belgium	Living in Belgium	
Israele	Living in Israele	
Czech Republic	Living in Czech Republic	
Poland	Living in Poland	
Luxembourg	Living in Luxembourg	
Portugal	Living in Portugal	
Estonia	Living in Estonia	
fin_resp	The individual is the financial respondent	yes = 1 ; no = 0
optimism	"How often do you feel that life is full of opportunities?" "How often do you feel that the future looks good for you?"	"Often" to at least one of the two questions = 1 ; otherwise = 0
cognition	Combination of numeracy and world list learning test	The higher the score the higher is cognition
underestimation	Respondent underestimation of life expectancy with respect to objective survival probabilities	The higher the value the higher the underestimation
smp	Does the respondent have a non-zero investment in bond, stock or mutual funds?	yes = 1 ; no = 0
equityshare	Amount invested in bond, stock, or mutual funds / (household net financial assets + household real assets)	

Table 1 presents a description of the variables used across this research thesis.

The variable *male* is equal one for males and zero for females. The dummy variable named *couple* equals one if the participant has a partner and zero otherwise. The variable *emp* is one if the respondent is employed or self-employed and zero otherwise. I created a dummy variable *child* equal to one if the number of children is greater or equal than one, zero otherwise.

For income, I generated household income quantiles, separately for each country.

I regrouped the ISCED[2] variable, and created a variable named *educ* with takes value 1 for ISCED 0-1 (very low education), value 2 for ISCED 2 (low education), value 3 for ISCED 3-4 (medium education), and value 4 for ISCED 5-6 (high education).

The variable *cognition* captures the cognitive abilities of the individual. In creating the variable, I considered a set memory tests included in SHARE: the number of words recalled in the first trial of the memory test, the number of words recalled in the second trial of the memory test, and the score in the numeracy test (subtractions of 7 from 100, 5 times). I use principal component analysis and define as measure of cognition the score derived from the first component.

Adl (activities of daily living) is an objective measure of health. The *adl* variable assigns a score from 1 to 10 to participants, the higher the score the higher the limitations in activities of daily living of the individual.

A key variable in the analysis is stock market participation (*smp*). The variable *smp* takes value 1 if the individual invests in bond, stocks or mutual funds, zero otherwise.

The dummy variable for the financial respondent (*fin_resp)* identifies people that answers to specific financial questions. The financial respondent answers the modules related to household finances on behalf of the couple. In case of a one-person household or a respondent living as single, the respondent is always the financial respondent. For individuals having a partner, one member of the couple is designated as financial respondent.

In a specific section there are questions about life satisfaction and the future, precisely "How often do you feel that life is full of opportunities?" and "How often do you feel that the future looks good for you?". The possible answers are "often" "sometimes" "rarely" "never". I used these variables to define an indicator of optimism. The *optimism* dummy takes value 1 if the respondent answers "often" to at least one of the two questions, zero otherwise.

The variables *underestimation* and *equityshare* will be described in detail in Chapter 3.

For the objective survival probabilities computation, there is an accurate description in the appendix section A. This information, that I collected for a set of countries participating in SHARE wave 6, is

[2] ISCED 0 = Early childhood education; ISCED 1 = Primary Education; ISCED 2 = Lower Secondary Education; ISCED 3 = Upper Secondary Education; ISCED 4 = Post-secondary non-Tertiary Education; ISCED 5 = Short-cycle tertiary education; ISCED 6 = Bachelors degree or equivalent tertiary education level

used to compare subjective survival probabilities with the subjective ones. Other details will follow in Appendix section A.

In Table 2 descriptive statistics are reported. My dataset is composed by a sample of 43415 individuals from 15[3] different countries (Austria, Belgium, Czech Republic, Denmark, Estonia, France, Germany, Israel, Italy, Luxembourg, Poland, Portugal, Spain, Switzerland, and Sweden), the household age has been restricted to a range between 50 years old and 90 years old to avoid the use of very selected individuals (due to selection into mortality). The average age of the sample is 67. 44% of individuals in the survey are males, 73% have a partner, 90% have at least one child, 38% are employed or self-employed and 24% have a college degree. Among the individuals, 20 percent have limitations in daily activities and 50 percent can be defined optimists given my definition of optimism. Finally, the average stock market participation in my sample is 24 percent.

Table 2: statistics. *This table presents descriptive statistics (number of observations and mean) for the control and dependent variables used in this thesis. All variables are defined in Table 1.*

Variable	N	Mean
age	43415	66.95
male	43415	0.44
couple	43415	0.73
emp	43415	0.38
child	43415	0.90
educ1	43415	0.21
educ2	43415	0.16
educ3	43415	0.39
educ4	43415	0.24
adl	43415	0.20
optimism	43415	0.50
cognition	43415	0.07
underestimation	43415	0.62
smp	43415	0.24
equityshare	10236	0.19

The table 2A describe the sample size divided by country, and therefore how each country contributes to the formation of the sample. Countries with among the highest numbers of observations are Italy, Estonia, and Belgium, where the latest represents almost the 11 percent of the sample size alone. On the other side, countries with the lowest number of observations are Israel and Luxembourg, representing only almost 3 percent each.

[3] SHARE dataset is composed by 18 countries. I exclude Croatia, Slovenia, and Greece because there are no life tables in the necessary form to calculate objective survival probabilities for year 2015.

In the section B of the appendix, there are tables presenting descriptive statistics for each country and a specific table in which only the financial respondent is considered.

***Table 2A**. This table presents number of observations for each country.*

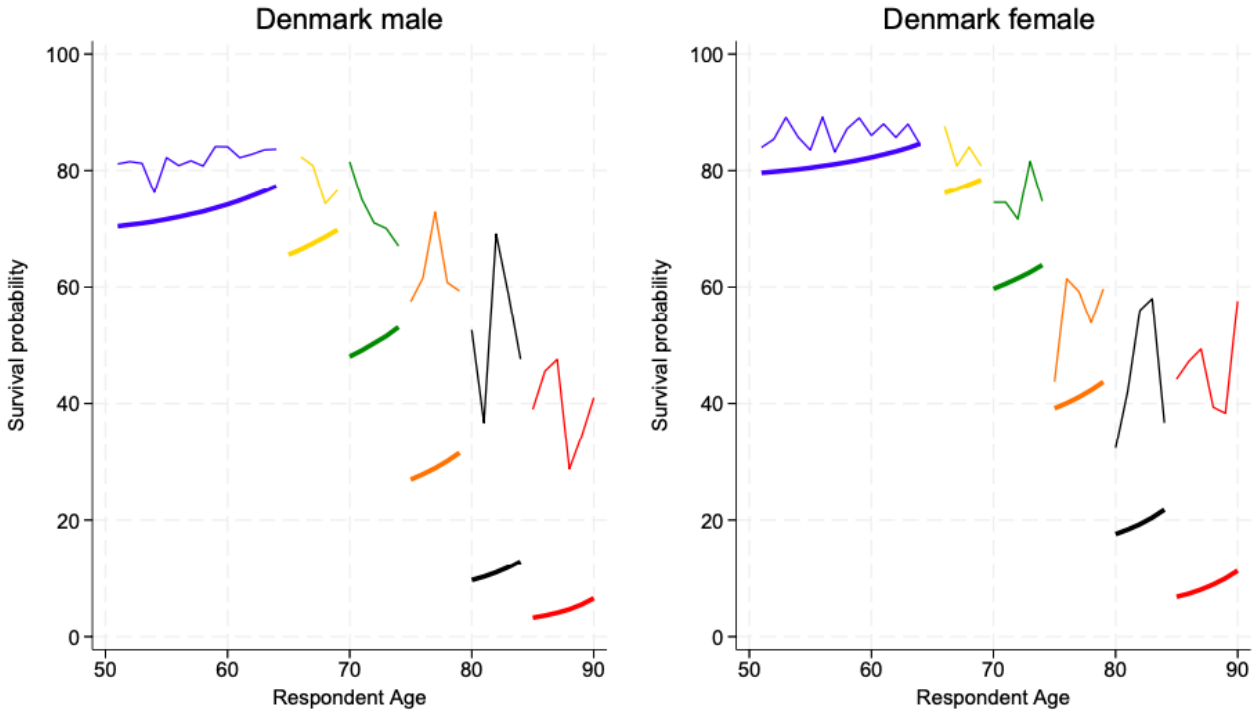
Country dummies	Freq.	Percent	Cum.
Austria	2,583	5.95	5.95
Germany	3,731	8.59	14.54
Sweden	3,108	7.16	21.7
Spain	3,709	8.54	30.25
Italy	4,142	9.54	39.79
France	3,073	7.08	46.86
Denmark	3,143	7.24	54.1
Switzerland	2,298	5.29	59.4
Belgium	4,729	10.89	70.29
Israel	1,222	2.81	73.1
Czech Republic	3,400	7.83	80.94
Poland	1,349	3.11	84.04
Luxembourg	1,256	2.89	86.94
Portugal	1,231	2.84	89.77
Estonia	4,441	10.23	100
Total	43,415	100.00	

2.2 Heterogeneity in subjective survival probabilities across countries

From an initial descriptive analysis, an interesting heterogeneity emerges in the way of evaluating one's survival probabilities. I did this analysis for all fifteen countries but in this section, I describe three of the most representative cases.
In section B in the appendix, I report the graphs of all the countries.

Denmark, Italy, and Czech Republic

***Figure 1**: subjective and objective survival probabilities for Denmark by age and gender*



***Legend.** Thinnest lines: subjective survival probabilities; Thickest lines: objective survival probabilities. Each color represents a different target age: blue 75 years, yellow 80 years, green 85 years, orange 90 years, black 95 years, red 100 years.*

Figure 1 represents Denmark survival probabilities analysis. It is divided by gender due to the different life expectancy and to evaluate potential differences in self-evaluations by gender. In the y-axis there are the survival probabilities expressed in a scale ranging between 0 and 100. In the x-axis there is the respondent age, expressed in a range from 50 years old to 90 years old. I divided the whole

pattern by target age where each target age is represented by a different colour[4]. Recalling the subjective probabilities description, the target age is the age for which is asked the probability to be still alive, and it is a function of the respondent age. In this representation there are age-ranges in which it is highlighted the pattern of survival probabilities.

For each colour there are two different lines: the thinnest represents the subjective survival probabilities and the thickest the objective survival probabilities.

There exist differences in the variability between objective and subjective.

Objective probabilities, computed from life tables, demonstrates higher stability and a consistent upward trend within a specified age range. For example, consider the age range from 50 to 65, where objective is always increasing. This happens because as individuals move towards the limit of the age range (65), their proximity to the target age in question (75) naturally leads to an increased probability of survival. However, when there is a shift in the target age (e.g. a change to age 80) it is expected to observe a drop in the objectives. This decline is expected: individuals' age increases which in turn leads to higher mortality rates.

The subjective probabilities display more variability. This variability arises from individuals' private information about health or diverse personal experiences and personal judgements. In addition, an increasing variability in the pattern with age is due to smaller sample size and increased noise in the average expectations computed from the data.

When examining the graph, the first observation is about the consistent differences between the subjective and objective probabilities, more precisely in the Denmark case, the consistent overestimation of survival probabilities (note that subjective probabilities are always above the objectives.) People tend to believe they will live longer with respect to actual mortality rates. Not in every country there is an overestimation, but a common pattern emerges across all countries: despite the initial condition that could be an overestimation or underestimation of life-expectancy, moving on with the age leads to an overestimation, that is from 75s onward there is a constant average overestimation for almost all countries. In some countries it is more pronounced, as in the case of Denmark. In fact for a Danish male aged 85, the average subjective probability to be still alive at 100 is around 60%, while the actual is only around 10%.

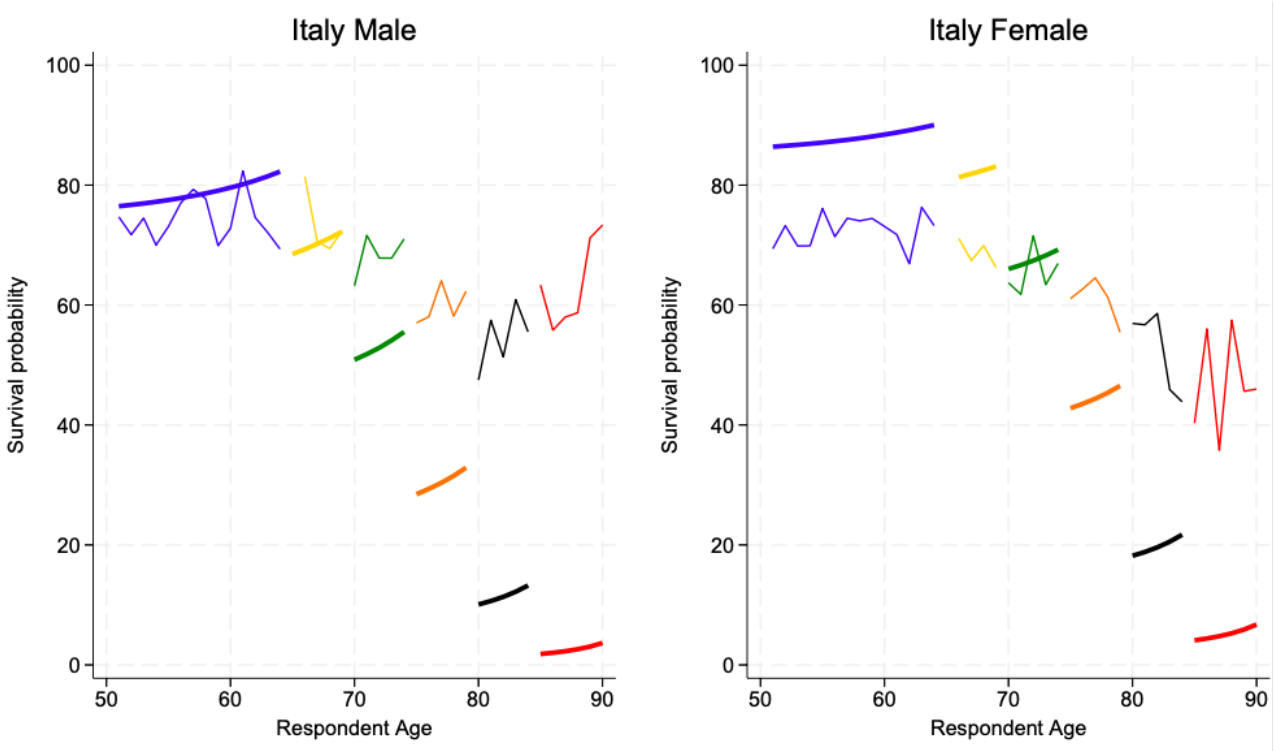
Figure 2 describes survival probabilities in Italy: the pattern is different. Italy initially exhibits higher objectives survival probabilities in the first age range (Italy has a reputation for having a higher life

[4] blue: target age 75, yellow: target age 80, green: target age 85, orange: target age 90, black: target age 95, red: target age 100

expectancy on global scale) and then they stabilize at the same level of Denmark. Both Italy and Denmark present some of the highest survival probabilities among the countries.

However, the subjective pattern in Italy is quite different. It starts with a slight underestimation with some values aligning with objectives in the first age range. Subsequently, it follows the common pattern previously described. Overestimation increasing with advancing age. When a Danish male reaches his early 50s, he believes he has an 80 percent chance of reaching the age of 75, while an Italian male estimates this probability to be around 70 percent. Furthermore, when analysing the 50-65 years old female survival figures for Italy and Denmark, an interesting result is evident, Italian females consistently underestimate their life expectancy, with values around 75 percent, while Danish females tend to overestimate theirs, with values around 85 percent.

***Figure 2**: subjective and objective survival probabilities for Italy by age and gender*



***Legend.** Thinnest lines: subjective survival probabilities; Thickest lines: objective survival probabilities. Each color represents a different target age: blue 75 years, yellow 80 years, green 85 years, orange 90 years, black 95 years, red 100 years.*

After reporting cases with higher survival probabilities, it is also important to discuss situations where the lowest levels of survival probabilities have been recorded.

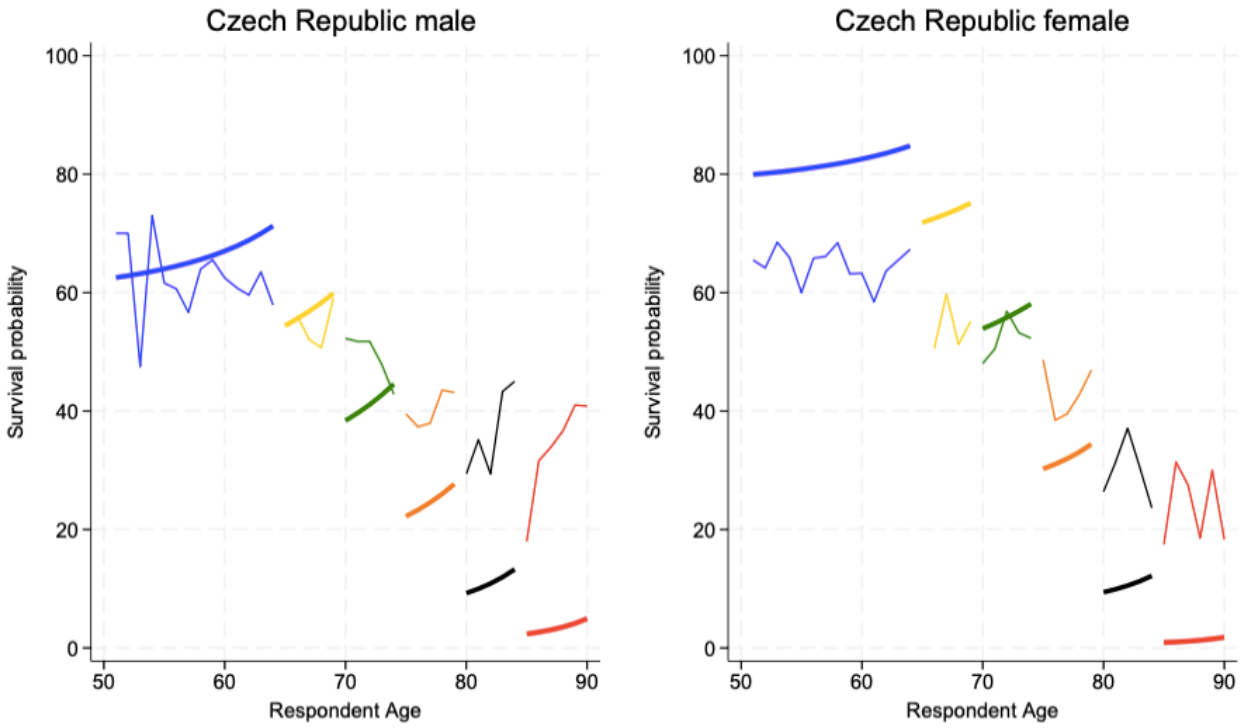
Figure 3 presents Czech Republic case, where significant differences are observed when compared to Italy and Denmark. Specifically, in case of males in their 50s, objective survival probabilities are

slightly above 60 percent, which is lower with respect to what I have described in previous countries. This lower benchmark in Czech Republic, probably reflects the overall standard of living and quality of life in the country.

Until 70s, subjective survival probabilities remain below the objectives resulting in a rather pessimistic self-assessment of remaining life span for individuals in the country. Overall, they display a similar pattern, beginning with an underestimation to end up in overestimation from 75s onwards. For example, a Czech Republic male in his 50s oscillates between 50 and 60 percent chances to be alive when he will be 75, that is lower with respect to other countries.

Comparing Czech Republic males with females is even more interesting. In fact, despite the significant difference in the benchmark (females have quite higher objective survival probabilities) the female self-evaluated survival probabilities are in line with their male counterparties.

***Figure 3**: subjective and objective survival probabilities for Czech Republic by age and gender*



***Legend.** Thinnest lines: subjective survival probabilities; Thickest lines: objective survival probabilities. Each color represents a different target age: blue 75 years, yellow 80 years, green 85 years, orange 90 years, black 95 years, red 100 years.*

Regarding gender differences, it is not surprising to see a consistent positive difference in the objectives in favour of females. Literature showed that in almost all countries around the world men

consistently live shorter lives than women (The male disadvantage in life expectancy: can we close the gender gap? , Muhammad Zakir Hossin 2021)
To further describe the significant heterogeneity in survival probabilities identified in this analysis I analyse the disparities in subjective estimates among females across these countries. For example, in the Czech Republic, for females in their 50s, the probability to be still alive at 75 years old is close to 60 percent, for Italian females close to 70 percent, for Danish females close to 85 percent.

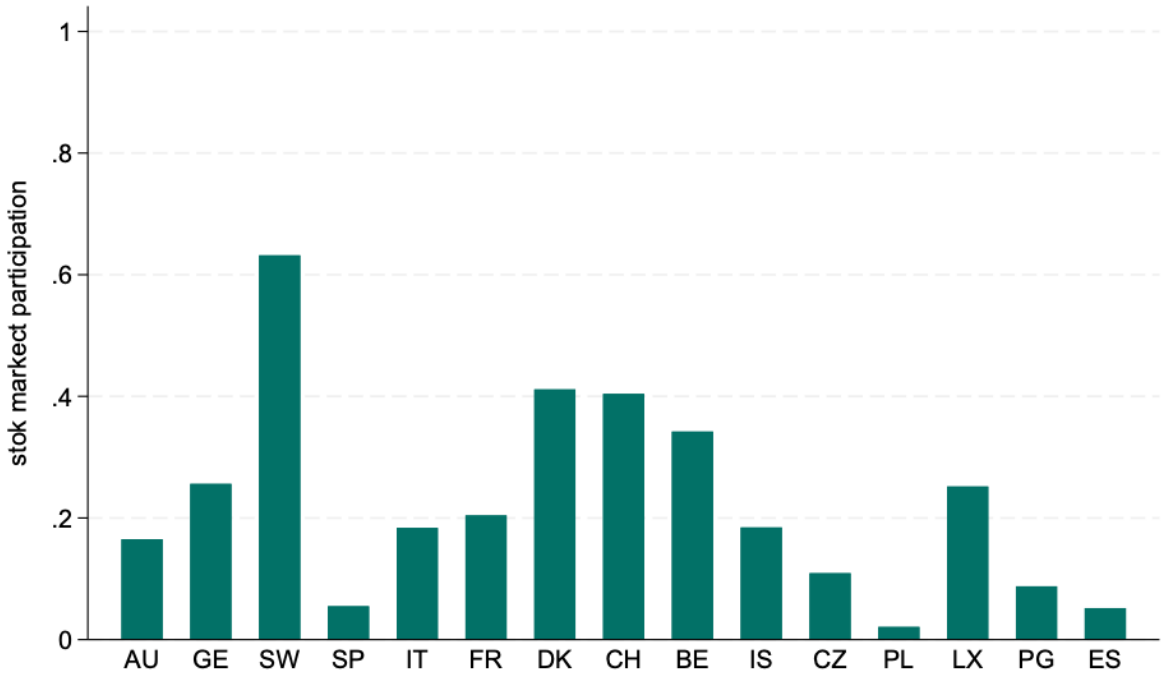
I will explain empirically the factors that influence subjective survival probabilities in chapter 3.
But in this section, it is interesting to discuss some insights about what drives formation of survival beliefs. As we have described before, private information about health and diverse personal experience could have a role in the variability. But another possible explanation comes from a research paper by Angelini et. al (2014). They analyse how self-reported life satisfaction vary across countries[5], a phenomenon that may be explained by the different scale and benchmark that people use to evaluate themselves. Precisely, they try to answer to the question "do differences in the level of life satisfaction across countries depend on how respondents interpret subjective survey questions?". Applying vignette methodology, they find that when scale differences are taken into account ranking of life satisfaction among different countries significantly changes. For example, assuming that responses scales are constant across individuals they find that Danes are the most satisfied and Italians are the least satisfied with life, but accounting for scale differences, the self-reported life satisfaction gap between Danes and Italians disappears and they have been replaced by the Netherlands and Czech Republic respectively. Another interesting finding in this paper is related to how the distribution of self-assessed life satisfaction in a country changes when other countries response scales are used. For example, with the Danish scale, more than 95% of respondents in all countries rate themselves satisfied or very satisfied with their own lives.
So, there is evidence that heterogeneity in life satisfaction is explained by different scales and benchmark that people use to evaluate their selves, and if it's true it could be also that the heterogeneity in the subjective survival probabilities is driven, at least in part, by these factors.

[5] They use SHARE data of the 2006 to 2007 wave. The sample is composed by the following countries: Sweden, Denmark, Germany, the Netherlands, Belgium, France, Spain, Italy, Poland and the Czech Republic

2.3 Stock market participation

In this section I analyse how the stock market participation varies across countries, from a descriptive perspective. In the first chapter, making use of the recent literature, I highlighted that low stock market participation remains a prevalent trend in European countries. To understand if our sample shows the same conditions, figure 4 represents the average stock market participation for each country in my sample. Since stock market participation is a binary variable taking value of one if the household participates in the market and value 0 otherwise, the figure is revealing us that, for instance, more than 60 percent of Swedish households within my sample participate in the market. As expected, northern-European countries like Denmark and Sweden along with some other financial hubs like Switzerland and Luxembourg exhibit higher levels of stock market participation. However, they're still below the 50 percent value, except only for Sweden.
Poland, Spain, and Estonia exhibit the lowest stock market participation in our sample.

***Figure 4**: average stock market participation by country*



***Legend**: AU – Austria, GE – Germany, SW – Sweden, SP – Spain, IT – Italy, FR – France, DK – Denmark, CH – Switzerland, BE – Belgium, IS – Israel, CZ – Czech Republic, PL – Poland, LX – Luxembourg, PG – Portugal, ES – Estonia*

In order to explore the potential correlation between stock market participation and subjective survival probabilities, it may be useful to plot a first descriptive graph between these two variables to understand if there is room for exploration.

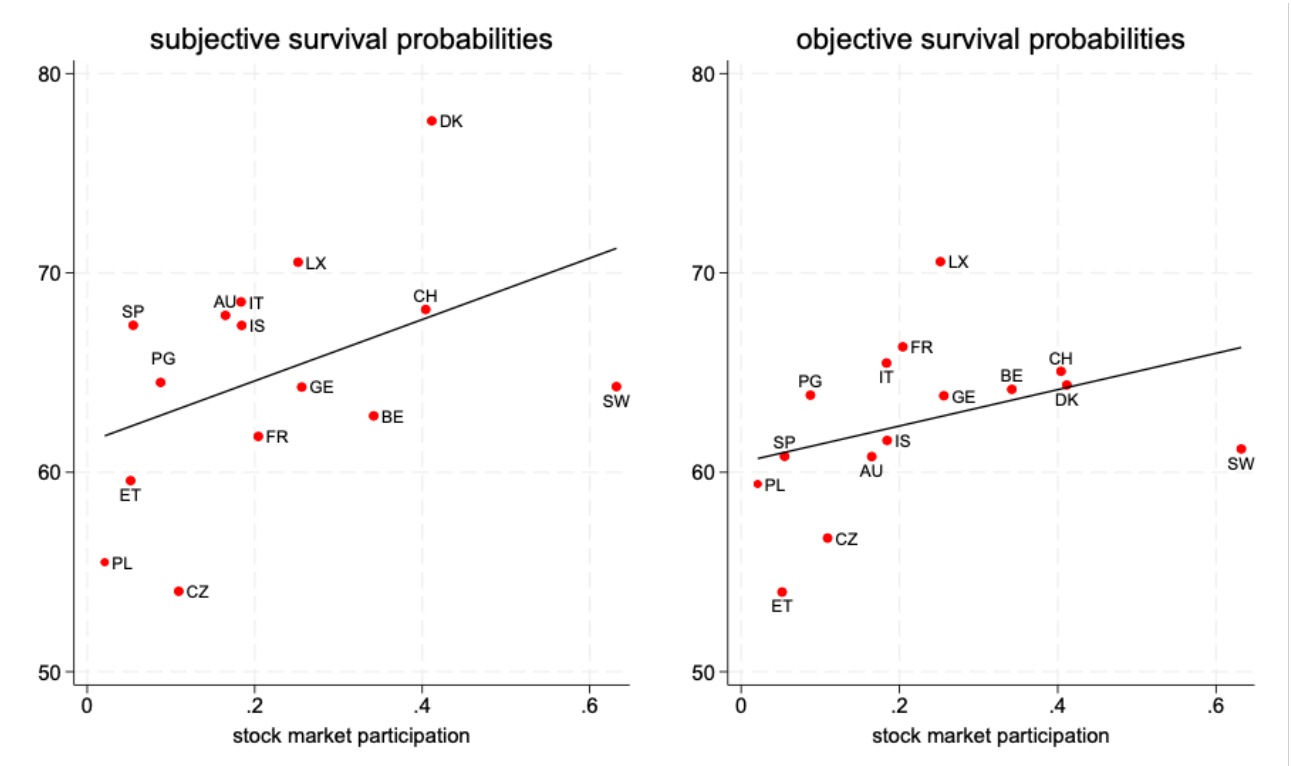
I present, with figure 5 a first analysis of the relationship between survival probabilities and stock market participation. The figure suggests that there is a positive relationship between the variables that invites further investigation.

The left-side panel is related to the subjective survival probabilities, the right hand one to the objective survival probabilities. On the y-axis there is average survival probability whereas in the x-axis the average stock market participation. In this way each dot represents the combination value between survival probabilities and stock market participation for each country.

Interesting to notice: Denmark, Switzerland and Luxembourg seem to have the highest subjective survival probabilities, and this is connected with also the highest level of stock market participations. Even more, countries with the lowest average levels of subjective survival probabilities seems to be also the one with the lowest stock market participations.

The panel of objectives interestingly follows a similar pattern but with a less strong correlation. When observing these two variables, a pattern emerges, making it worthwhile to analyse their relationship.

***Figure 5**: Survival probabilities and stock market participation, by country.*



***Legend**: AU – Austria, GE – Germany, SW – Sweden, SP – Spain, IT – Italy, FR – France, DK – Denmark, CH – Switzerland, BE – Belgium, IS – Israel, CZ – Czech Republic, PL – Poland, LX – Luxembourg, PG – Portugal, ET – Estonia*

Chapter 3

3.1 Underestimation

The main empirical relationship I want to analyse is between subjective survival probabilities and stock market participation.

The figure showed at the end of chapter 2 suggests this relationship direction: countries with the lowest subjective survival probabilities are the one with the lowest stock market participation.

To account for longevity expectations (Chapter 1 suggests that horizon matters in financial decisions), I created a variable that focuses on the underestimation of survival probabilities (i.e. subjective survival probabilities are lower with respect to objectives).

At this point, a possible question observing data could be: is the underestimation of survival probabilities reducing stock market participation?

Since I am interested in participants that underestimate their survival life probability, I created an accuracy measure as follows:

$$accuracy = \begin{cases} \dfrac{subjective_i}{objective_i}, & if\ \dfrac{subjective_i}{objective_i} < 1 \\ 1, & if\ \dfrac{subjective_i}{objective_i} \geq 1 \end{cases}$$

The variable is created in such a way that if participants underestimate their survival probabilities (subjective < objective → ratio < 1) it is equal to the ratio between subjective and objective, if instead participants do not deviate from the objective or they overestimate their survival probabilities (subjective >= objective → ratio >= 1) the variable takes value 1.

The lower the objective with respect to the subjective, the lower will be the accuracy measure (closer to zero).

Starting from *accuracy*, I define my variable of interest in the following way:

Equation 1

$$underestimation = \frac{1 - accuracy}{\sigma_{accuracy}}$$

The variable *underestimation* is zero for every participant that is accurate or overestimate survival probabilities with respect to the life tables. It is value if participant underestimates, more precisely: the higher the underestimation, the higher the value. Note that the underestimation variable is standardized in such a way that standard deviation is equal to one.

3.2 Optimism

Many factors affect underestimation or overestimation (and in general the estimation) of survival probabilities. For example, Subjective survival probabilities covary with factors such as health status, health behaviour or income. It is reasonable to assume that if participants are free from illnesses and avoid unhealthy behaviours like smoking, they may expect to have a longer life expectancy compared to someone who engages in these unhealthy behaviours.
The literature has not only highlighted the significance of these topics but has also demonstrated that optimism and cognitive factors play a crucial role in shaping beliefs and influencing economic decisions. Puri and Robinson (2006) demonstrated that optimism is related to numerous work/life choices: more optimistic people work harder, expect to retire later, are more likely to remarry, invest more in individual stocks, and save more. An optimistic outlook induces individuals to overestimate the probability that positive events happen. Coherently, Angelini and Cavapozzi (2015) show that Dispositional optimism is found to be a relevant predictor of the ownership of stocks as well as of the share of gross financial wealth invested in this asset. They also take in consideration that individuals with higher cognitive skills might be more accurate in estimating their survival probabilities.

A very interesting paper by Grevenbrock et. al (2021) studies the roles of psychological biases in the differences between subjective survival probabilities and objective survival probabilities, in the U.S. population. They identify two factors: cognition and optimism. Their regression analyses confirm that these factors play important roles in the formation of survival beliefs. First, in their paper there is an interesting mention of the “flatness bias”, i.e. the trend that previous literature showed about the deviations between subjective survival probabilities and objective survival probabilities. Respondents of age 50-70 express underestimation while respondents of age 75 onwards express overestimation, on average. This is very interesting compared to what I found. If it is true that in some cases the pattern presents this flatness bias, in others it deviates by presenting a constant overestimate.

They consider relative optimism and, as a cognitive factor, the likelihood insensitivity. Likelihood insensitivity refers to a cognitive weakness according to which people cannot distinguish well among the likelihood of different events.
Their findings suggest that although optimism influences the formation of beliefs, age-increasing overestimation of objective survival probabilities is not due to increasing optimism as one may expect. It is rather a consequence of age-increasing insensitivity to objective survival probabilities.
Therefore, the tendency to overestimate survival probabilities as one gets older could be attributed at least in part to cognitive limitations.
With these findings, the literature helps in understanding the importance of optimism and cognitive abilities as driving factors for survival beliefs.
Since I want to understand the pure effect of the underestimation, i.e. the time horizon component of underestimation, in my main regression I added as controls the variables *optimism* and *cognition.*

3.3 Explaining underestimation

Table 3 presents OLS estimates for a linear regression of my underestimation measure on several controls. The aim of these estimates is to illustrate what drives the underestimation of subjective survival.
To account for the possible correlation between observations within the dataset (violation of independence assumption) standard error need to be adjusted. In particular, SHARE data are collected for more than one individual per household, therefore in the first column standard errors are clustered at the household level. In the second column, I simply selected the financial respondent for each household, i.e. the household member that answers to the economic and financial questions.
The coefficients for optimism and cognition are negative and statistically significant. This result suggests a strong correlation between these two indicators and underestimation. As expected, a more optimistic attitude makes participants underestimate less their survival probabilities, the same is true for cognition: the higher the cognitive abilities the less survival probabilities are underestimated, suggesting a better ability in forming probabilistic expectations.
All coefficients are statistically significant except for: *educ*, *couple*, *age* and *age2* (in column two *age* is statistically significant at 10% level).
Although the age and age squared variables appear not to be significant, an F-test indicates that the two variables are jointly significantly different from zero.

***Table 3**: Explaining variations in Underestimation. Model estimated using OLS. In the first column standard errors are clustered at the household level, in the second column financial respondent is selected.*

VARIABLES	1 underestimation	2 underestimation
optimism	-0.322***	-0.321***
	(0.00986)	(0.0120)
cognition	-0.0571***	-0.0614***
	(0.00543)	(0.00619)
male	-0.124***	-0.126***
	(0.00890)	(0.0118)
Age	-0.00468	-0.0142*
	(0.00778)	(0.00815)
age2	-0.0000251	-0.0000385
	(5.76e-05)	(5.85e-05)
Couple	-0.0131	-0.00243
	(0.0135)	(0.0143)
child	-0.0426**	-0.0377**
	(0.0171)	(0.0182)
emp	-0.154***	-0.168***
	(0.0120)	(0.0158)
educ2	0.00724	0.00707
	(0.0174)	(0.0198)
educ3	0.00300	0.0150
	(0.0162)	(0.0182)
educ4	0.000058	0.00219
	(0.0178)	(0.0205)
hhquart = 2	-0.0546***	-0.0558***
	(0.0155)	(0.0160)
hhquart = 3	-0.0711***	-0.0726***
	(0.0160)	(0.0179)
hhquart = 4	-0.113***	-0.120***
	(0.0161)	(0.0193)
adl	0.162***	0.155***
	(0.00967)	(0.00792)
Country identifier = 11, Austria	-0.147***	-0.148***
	(0.0243)	(0.0299)
Country identifier = 13, Sweden	0.0112	0.0101
	(0.0241)	(0.0288)
Country identifier = 15, Spain	-0.253***	-0.273***
	(0.0232)	(0.0298)
Country identifier = 16, Italy	-0.180***	-0.190***
	(0.0233)	(0.0284)
Country identifier = 17, France	0.0869***	0.0878***
	(0.0246)	(0.0289)
Country identifier = 18, Denmark	-0.254***	-0.245***
	(0.0208)	(0.0287)
Country identifier = 20, Switzerland	-0.0357	-0.0515*
	(0.0235)	(0.0312)
Country identifier = 23, Belgium	-0.00166	-0.0167
	(0.0217)	(0.0259)
Country identifier = 25, Israel	-0.0835**	-0.0658*
	(0.0345)	(0.0394)
Country identifier = 28, Czech Republic	0.153***	0.144***
	(0.0259)	(0.0284)
Country identifier = 29, Poland	0.152***	0.124***
	(0.0352)	(0.0380)
Country identifier = 31, Luxembourg	-0.0665**	-0.0642*
	(0.0304)	(0.0377)
Country identifier = 33, Portugal	-0.165***	-0.194***
	(0.0348)	(0.0418)
Country identifier = 35, Estonia	0.0444*	0.0297
	(0.0242)	(0.0263)
Constant	1.452***	1.805***
	(0.261)	(0.283)
Observations	43.415	30.455
R-squared	0.091	0.090

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the analysis also emerges that if participants are employed or self-employed, they tend to underestimate less their life survival probabilities, and this is true also for participants that have children. The same relationship direction is true for income quintiles. As the household income increases the underestimation decreases.

Country dummies are all statistically significant except for Switzerland, Belgium, and Sweden. More interesting, they seem to show the pattern illustrated in chapter two. For example, regression analysis shows that in Denmark there is the tendency to underestimate less the life survival probabilities with respect to the country reference category: Germany. In fact, as we remember from chapter 2 Denmark showed a constant overestimation.

All these variables enter with the expected sign and together explain almost 10% of the variation in the underestimation in my sample, suggesting that the underestimation variable is not entirely explained by these factors.

3.4 Underestimation and stock market participation

The estimating equation is defined as follows

$$y_i = \beta_0 + \beta_1 Underestimation_i + \gamma_1 X_i + \mu_i$$

The outcome y_i is the stock market participation variable. Stock market participation is regressed on a constant term, the underestimation variable, a set of regressors X_i and an error term μ_i.

The variable included in the set of regressors were described in chapter two: *male, age, age2, couple child, emp, educ, hhquart, adl and country dummies*.

The model is estimated by ordinary least squares, given the binary nature of the outcome variable I am estimating a linear probability model which has heteroskedastic standard errors. I therefore report in the estimation results heteroskedasticity robust standard errors.

Table 4 reports results for the regression. The first column considers all observations and therefore I use as before cluster robust standard error to account for the correlation between same household members. The second column considers the financial respondent for each household.

Let's focus the result discussion on the second column. As expected, the main finding indicates that individual's underestimation of life survival probabilities has a negative impact on their participation

Table 4: *Underestimation and stock market participation. Model estimated using OLS. In the first column standard errors are clustered at the household level, in the second column financial respondent is selected. Heteroskedasticity robust standard errors.*

	1	2
VARIABLES	smp	smp
underestimation	-0.00708***	-0.00795***
	(0.00179)	(0.00195)
cognition	0.0292***	0.0311***
	(0.00206)	(0.00231)
optimism	0.0365***	0.0389***
	(0.00418)	(0.00469)
male	0.00241	0.0400***
	(0.00268)	(0.00464)
Age	0.0235***	0.0232***
	(0.00295)	(0.00302)
age2	-0.000145***	-0.000141***
	(2.11e-05)	(2.16e-05)
Couple	0.0185***	0.00739
	(0.00528)	(0.00535)
child	-0.0503***	-0.0494***
	(0.00768)	(0.00722)
emp	-0.00591	0.00436
	(0.00536)	(0.00613)
educ2	0.0249***	0.0233***
	(0.00588)	(0.00650)
educ3	0.0693***	0.0653***
	(0.00587)	(0.00641)
educ4	0.122***	0.128***
	(0.00724)	(0.00784)
hhquart = 2	0.0415***	0.0425***
	(0.00572)	(0.00551)
hhquart = 3	0.0778***	0.0789***
	(0.00665)	(0.00663)
hhquart = 4	0.169***	0.159***
	(0.00761)	(0.00756)
adl	-0.0136***	-0.0134***
	(0.00205)	(0.00223)
Country identifier = 11, Austria	-0.0988***	-0.0917***
	(0.0123)	(0.0115)
Country identifier = 13, Sweden	0.378***	0.368***
	(0.0135)	(0.0131)
Country identifier = 15, Spain	-0.118***	-0.106***
	(0.0108)	(0.0104)
Country identifier = 16, Italy	0.00542	0.0167
	(0.0119)	(0.0114)
Country identifier = 17, France	-0.0121	-0.0151
	(0.0125)	(0.0117)
Country identifier = 18, Denmark	0.150***	0.150***
	(0.0140)	(0.0133)
Country identifier = 20, Switzerland	0.135***	0.124***
	(0.0150)	(0.0142)
Country identifier = 23, Belgium	0.112***	0.0989***
	(0.0121)	(0.0113)
Country identifier = 25, Israel	-0.0528***	-0.0448***
	(0.0157)	(0.0155)
Country identifier = 28, Czech Republic	-0.109***	-0.0931***
	(0.0113)	(0.0106)
Country identifier = 29, Poland	-0.171***	-0.148***
	(0.0106)	(0.0100)
Country identifier = 31, Luxembourg	0.0361**	0.0413***
	(0.0167)	(0.0159)
Country identifier = 33, Portugal	-0.0630***	-0.0528***
	(0.0143)	(0.0138)
Country identifier = 35, Estonia	-0.175***	-0.157***
	(0.00985)	(0.00921)
Constant	-0.792***	-0.808***
	(0.102)	(0.105)
Observations	43.415	30.455
R-squared	0.208	0.213

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

in the stock market. Specifically, a one standard deviation increase in underestimation results in approximately 1 percent reduction in stock market participation.

To address the issue that deviations between subjective and objective probabilities reflect the fact that individuals have more information about their health behavior, or that they overestimate (underestimate) their life survival probabilities because they are optimist (pessimist), I controlled for a set of regressors that helps to highlight the pure effect of underestimation on the stock market participation.

All coefficients are statistically significant except for *emp* and *couple.*

Optimism and *cognition* enters with the expected sign: an optimistic household tend to invest more in the market, as in the case of a household with higher cognitive abilities.

Child coefficient suggests individuals prefer not to invest when they have children. Precisely, they invest 5% less in the stock market.

Stock market participation is increasing with the education level and increasing with the household income, while it has a negative relationship with limitations of activity daily living.

There is evidence of a gender gap: stock market participation is higher for male. Precisely, a male is 4 percent more likely to invest in the market.

Country dummies are all statistically significant except for Italy and France. It's interesting to notice that differences in stock market participation found in the descriptive analysis also show up in the empirical analysis. For example, countries like Denmark and Sweden have a higher participation in the market with respect to Germany (the reference category), respectively 15 percent and 38 percent more. Also, countries in which stock market participation is lower, like Estonia and Czech Republic, show a stock market participation that is 10 and 15 percent lower with respect to Germany.

These results suggest that a heterogeneity analysis can highlight some interesting patterns.

3.5 Heterogeneity analysis: country, age, education level

Since my sample is composed of different countries, it is interesting to estimate the same model dividing my sample in subgroups representing different geographical areas.

Table 5 reports estimation results. The first column reproposes the baseline model estimated in the full sample, while the other columns show estimates for sub samples of countries. In the second column denominated "northern Europe" the model is estimated only for Sweden, Denmark and Estonia. In this case, the reference category for the country dummy variables is Sweden. For example, in the sub sample of northern countries, households living in Estonia have a 50 percent lower stock market participation with respect to Sweden. In the third column denominated "eastern Europe" the model is estimated only for Poland and Czech Republic. In this case, the reference category for country dummies is Czech Republic. In the fourth column denominated "southern Europe" the model is estimated only for Portugal, Italy and Spain and the reference category for the country dummies is Spain: households in Italy participate in the stock market 12 percent more with respect to their counterparties in Spain.

The last column is denominated "western Europe", and the model is estimated in a subsample composed by: Germany, Luxembourg, Belgium, Switzerland, France, Israel and Austria. In this case, the reference category is Austria.

In Comparing the subsamples, underestimation coefficient is significantly different from zero only when I estimate the model for western Europe and northern Europe countries. These two subsamples seems to behave similarly in terms of underestimation effect on stock market participation and they do not particularly deviate from the full sample. The parameter is smaller in magnitude for southern Europe countries, (almost half with respect to northern and western Europe) and the lack of significance might be driven by the low sample size. The effect is instead zero for eastern Europe countries.

Table 5: *Heterogeneity analysis by country*

VARIABLES	full sample smp	northern europe smp	eastern europe smp	southern europe smp	western europe smp
underestimation	-0.00708***	-0.0100***	-0.000333	-0.00487	-0.0108***
	(0.00179)	(0.00340)	(0.00358)	(0.00348)	(0.00319)
cognition	0.0292***	0.0306***	0.0181***	0.0385***	0.0262***
	(0.00206)	(0.00425)	(0.00453)	(0.00420)	(0.00330)
optimism	0.0365***	0.0169*	0.0311***	0.0220***	0.0503***
	(0.00418)	(0.00870)	(0.00940)	(0.00845)	(0.00675)
male	0.00241	0.00809	0.0105	-0.0122***	0.00371
	(0.00268)	(0.00561)	(0.00636)	(0.00457)	(0.00456)
Age	0.0235***	0.0161***	0.00827	0.0166***	0.0331***
	(0.00295)	(0.00594)	(0.00704)	(0.00536)	(0.00485)
age2	-0.000145***	-8.18e-05*	-5.02e-05	-0.000101***	-0.000215***
	(2.11e-05)	(4.28e-05)	(5.02e-05)	(3.83e-05)	(3.48e-05)
Couple	0.0185***	0.0167	0.0189	-0.00456	0.0284***
	(0.00528)	(0.0122)	(0.0127)	(0.00911)	(0.00865)
child	-0.0503***	-0.0476***	-0.0782***	-0.0359***	-0.0540***
	(0.00768)	(0.0160)	(0.0265)	(0.0133)	(0.0121)
emp	-0.00591	-0.00320	0.0119	0.00801	-0.00967
	(0.00536)	(0.0112)	(0.0130)	(0.00953)	(0.00889)
educ2	0.0249***	0.0531***	-0.0268**	-0.000459	0.0418***
	(0.00588)	(0.0176)	(0.0108)	(0.00942)	(0.0105)
educ3	0.0693***	0.0746***	0.0172*	0.0324**	0.104***
	(0.00587)	(0.0171)	(0.00935)	(0.0128)	(0.00915)
educ4	0.122***	0.114***	0.107***	0.0563***	0.159***
	(0.00724)	(0.0183)	(0.0200)	(0.0169)	(0.0110)
hhquart = 2	0.0415***	0.0326**	0.00825	-0.000420	0.0705***
	(0.00572)	(0.0133)	(0.0128)	(0.00893)	(0.00957)
hhquart = 3	0.0778***	0.0823***	-0.00218	0.0363***	0.112***
	(0.00665)	(0.0159)	(0.0157)	(0.0107)	(0.0112)
hhquart = 4	0.169***	0.133***	0.0516***	0.155***	0.224***
	(0.00761)	(0.0175)	(0.0193)	(0.0135)	(0.0125)
adl	-0.0136***	-0.0127***	-0.00321	-0.00580*	-0.0201***
	(0.00205)	(0.00454)	(0.00448)	(0.00309)	(0.00359)
Country identifier = 11, Austria	-0.0988***				
	(0.0123)				
Country identifier = 13, Sweden	0.378***				
	(0.0135)				
Country identifier = 15, Spain	-0.118***				
	(0.0108)				
Country identifier = 16, Italy	0.00542			0.122***	
	(0.0119)			(0.00907)	
Country identifier = 17, France	-0.0121				0.0935***
	(0.0125)				(0.0125)
Country identifier = 18, Denmark	0.150***	-0.220***			
	(0.0140)	(0.0151)			
Country identifier = 20, Switzerland	0.135***				0.229***
	(0.0150)				(0.0149)
Country identifier = 23, Belgium	0.112***				0.214***
	(0.0121)				(0.0119)
Country identifier = 25, Israel	-0.0528***				0.0468***
	(0.0157)				(0.0155)
Country identifier = 28, Czech Republic	-0.109***				
	(0.0113)				
Country identifier = 29, Poland	-0.171***		-0.0866***		
	(0.0106)		(0.00986)		
Country identifier = 31, Luxembourg	0.0361**				0.140***
	(0.0167)				(0.0166)
Country identifier = 33, Portugal	-0.0630***			0.0434***	
	(0.0143)			(0.0115)	
Country identifier = 35, Estonia	-0.175***	-0.558***			
	(0.00985)	(0.0119)			
Country identifier = 12, Germany					0.0946***
					(0.0125)
Constant	-0.792***	-0.196	-0.200	-0.587***	-1.281***
	(0.102)	(0.205)	(0.245)	(0.186)	(0.168)
Observations	43.415	10.692	4.749	9.082	18.892
R-squared	0.208	0.308	0.077	0.119	0.126
Robust standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

Table 6 presents results for a different heterogeneity analysis: I estimated the model by level of education subsamples. As before, the first column is the estimated model in the full sample. In the second column denominated "educ2" I estimated the model only for individuals with code 0-1-2 of ISCED level of education. For the column "educ3" the model is estimated only for individuals with code 3-4 of ISCED level of education. For the last column only individuals with high education are considered, that is code 5-6 of ISCED level of education.

Coefficients of subjective survival are statistically significant for the subsamples with ISCED code 3-4 and ISCED code 5-6 (in this last case at a 10% significance level). It seems that individuals with a medium level of education (ISCED 3-4) the underestimation of survival probabilities matters more for the stock participation with respect to lower and highest level respectively.

Finally, table 7 presents results of the heterogeneity by age. I estimated the same model in three different subsamples divided by age. Precisely, the second column presents results only for individuals aged between 50 and 64, third column for individuals aged between 65 and 74 and the last column for individuals 75 or more.

Results show that for individuals in pre-retirement age (50-64) underestimation of survival probabilities matters more in the stock market participation. Precisely, for this subsample a 1 standard deviation increase in underestimation reduces the stock market participation by almost 2 percent.

Table 6: *Heterogeneity analysis by education level*

VARIABLES	full sample smp	educ2 smp	educ3 smp	educ4 smp
underestimation	-0.00708***	-0.00245	-0.0138***	-0.00830*
	(0.00179)	(0.00227)	(0.00296)	(0.00480)
cognition	0.0292***	0.0268***	0.0279***	0.0403***
	(0.00206)	(0.00267)	(0.00338)	(0.00499)
optimism	0.0365***	0.0359***	0.0325***	0.0445***
	(0.00418)	(0.00594)	(0.00657)	(0.00958)
male	0.00241	-0.00319	-0.00794	0.0228***
	(0.00268)	(0.00448)	(0.00532)	(0.00770)
Age	0.0235***	0.0224***	0.0180***	0.0234***
	(0.00295)	(0.00380)	(0.00484)	(0.00686)
age2	-0.000145***	-0.000146***	-0.000105***	-0.000135***
	(2.11e-05)	(2.69e-05)	(3.54e-05)	(5.01e-05)
Couple	0.0185***	0.0192***	0.0107	0.0302**
	(0.00528)	(0.00668)	(0.00858)	(0.0125)
child	-0.0503***	-0.0436***	-0.0512***	-0.0606***
	(0.00768)	(0.00984)	(0.0126)	(0.0162)
emp	-0.00591	0.0150*	-0.0176**	-0.00937
	(0.00536)	(0.00781)	(0.00827)	(0.0123)
educ2	0.0249***			
	(0.00588)			
educ3	0.0693***			
	(0.00587)			
educ4	0.122***			
	(0.00724)			
hhquart = 2	0.0415***	0.0234***	0.0591***	0.0479***
	(0.00572)	(0.00672)	(0.00947)	(0.0157)
hhquart = 3	0.0778***	0.0598***	0.0880***	0.0943***
	(0.00665)	(0.00822)	(0.0105)	(0.0163)
hhquart = 4	0.169***	0.146***	0.184***	0.183***
	(0.00761)	(0.0111)	(0.0116)	(0.0166)
adl	-0.0136***	-0.00358	-0.0240***	-0.0329***
	(0.00205)	(0.00234)	(0.00383)	(0.00707)
Country identifier = 11, Austria	-0.0988***	-0.0756***	-0.116***	-0.0870***
	(0.0123)	(0.0187)	(0.0158)	(0.0240)
Country identifier = 13, Sweden	0.378***	0.401***	0.365***	0.395***
	(0.0135)	(0.0237)	(0.0192)	(0.0218)
Country identifier = 15, Spain	-0.118***	-0.0842***	-0.152***	-0.155***
	(0.0108)	(0.0171)	(0.0194)	(0.0253)
Country identifier = 16, Italy	0.00542	0.0311*	-0.00954	0.00517
	(0.0119)	(0.0182)	(0.0187)	(0.0307)
Country identifier = 17, France	-0.0121	-0.000115	-0.0329*	0.0257
	(0.0125)	(0.0190)	(0.0173)	(0.0253)
Country identifier = 18, Denmark	0.150***	0.214***	0.156***	0.133***
	(0.0140)	(0.0275)	(0.0190)	(0.0218)
Country identifier = 20, Switzerland	0.135***	0.118***	0.125***	0.180***
	(0.0150)	(0.0282)	(0.0182)	(0.0311)
Country identifier = 23, Belgium	0.112***	0.0878***	0.123***	0.156***
	(0.0121)	(0.0196)	(0.0178)	(0.0203)
Country identifier = 25, Israel	-0.0528***	-0.0398*	-0.0608**	-0.0361
	(0.0157)	(0.0222)	(0.0247)	(0.0277)
Country identifier = 28, Czech Republic	-0.109***	-0.0716***	-0.136***	-0.0741***
	(0.0113)	(0.0180)	(0.0145)	(0.0274)
Country identifier = 29, Poland	-0.171***	-0.0982***	-0.198***	-0.249***
	(0.0106)	(0.0168)	(0.0132)	(0.0279)
Country identifier = 31, Luxembourg	0.0361**	0.00704	0.0499**	0.0902**
	(0.0167)	(0.0223)	(0.0247)	(0.0389)
Country identifier = 33, Portugal	-0.0630***	-0.0573***	-0.0715*	-0.0238
	(0.0143)	(0.0187)	(0.0407)	(0.0520)
Country identifier = 35, Estonia	-0.175***	-0.0944***	-0.195***	-0.179***
	(0.00985)	(0.0169)	(0.0121)	(0.0191)
Constant	-0.792***	-0.728***	-0.511***	-0.753***
	(0.102)	(0.133)	(0.164)	(0.233)
Observations	43.415	16.15	16.782	10.483
R-squared	0.208	0.190	0.177	0.163

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: *Heterogeneity analysis by age*

VARIABLES	full sample smp	age 50-64 smp	age 65-74 smp	age > 74 smp
underestimation	-0.00708***	-0.0174***	0.0000329	-0.00475*
	(0.00179)	(0.00316)	(0.00358)	(0.00269)
cognition	0.0292***	0.0260***	0.0289***	0.0335***
	(0.00206)	(0.00311)	(0.00385)	(0.00379)
optimism	0.0365***	0.0345***	0.0303***	0.0469***
	(0.00418)	(0.00600)	(0.00783)	(0.00849)
male	0.00241	-0.00649	0.00790	0.00660
	(0.00268)	(0.00411)	(0.00594)	(0.00647)
Age	0.0235***	0.0369	0.0689	-0.0132
	(0.00295)	(0.0241)	(0.0941)	(0.0344)
age2	-0.000145***	-0.000260	-0.000456	7.99e-05
	(2.11e-05)	(0.000208)	(0.000671)	(0.000210)
Couple	0.0185***	0.0152*	0.0105	0.0145
	(0.00528)	(0.00775)	(0.0101)	(0.00986)
child	-0.0503***	-0.0672***	-0.0426***	-0.0281**
	(0.00768)	(0.0112)	(0.0143)	(0.0131)
emp	-0.00591	0.00313	-0.0138	0.0169
	(0.00536)	(0.00658)	(0.0114)	(0.0209)
educ2	0.0249***	0.0267***	0.0192*	0.0171
	(0.00588)	(0.00899)	(0.0114)	(0.0108)
educ3	0.0693***	0.0657***	0.0624***	0.0689***
	(0.00587)	(0.00885)	(0.0109)	(0.0110)
educ4	0.122***	0.114***	0.119***	0.128***
	(0.00724)	(0.0108)	(0.0135)	(0.0140)
hhquart = 2	0.0415***	0.0363***	0.0618***	0.0371***
	(0.00572)	(0.00854)	(0.0110)	(0.00979)
hhquart = 3	0.0778***	0.0647***	0.112***	0.0839***
	(0.00665)	(0.00949)	(0.0124)	(0.0127)
hhquart = 4	0.169***	0.165***	0.209***	0.150***
	(0.00761)	(0.0100)	(0.0148)	(0.0178)
adl	-0.0136***	-0.0124***	-0.0185***	-0.00871***
	(0.00205)	(0.00422)	(0.00427)	(0.00276)
Country identifier = 11, Austria	-0.0988***	-0.0927***	-0.0890***	-0.116***
	(0.0123)	(0.0180)	(0.0216)	(0.0226)
Country identifier = 13, Sweden	0.378***	0.345***	0.385***	0.413***
	(0.0135)	(0.0210)	(0.0225)	(0.0251)
Country identifier = 15, Spain	-0.118***	-0.125***	-0.128***	-0.100***
	(0.0108)	(0.0152)	(0.0200)	(0.0208)
Country identifier = 16, Italy	0.00542	0.000318	0.0221	-0.00887
	(0.0119)	(0.0163)	(0.0218)	(0.0232)
Country identifier = 17, France	-0.0121	0.00531	-0.0204	-0.0380
	(0.0125)	(0.0175)	(0.0236)	(0.0232)
Country identifier = 18, Denmark	0.150***	0.103***	0.195***	0.239***
	(0.0140)	(0.0187)	(0.0252)	(0.0293)
Country identifier = 20, Switzerland	0.135***	0.0672***	0.202***	0.175***
	(0.0150)	(0.0210)	(0.0264)	(0.0291)
Country identifier = 23, Belgium	0.112***	0.108***	0.168***	0.0640***
	(0.0121)	(0.0162)	(0.0231)	(0.0240)
Country identifier = 25, Israel	-0.0528***	-0.0608***	0.00488	-0.106***
	(0.0157)	(0.0233)	(0.0291)	(0.0256)
Country identifier = 28, Czech Republic	-0.109***	-0.101***	-0.113***	-0.113***
	(0.0113)	(0.0163)	(0.0195)	(0.0213)
Country identifier = 29, Poland	-0.171***	-0.171***	-0.183***	-0.154***
	(0.0106)	(0.0145)	(0.0195)	(0.0206)
Country identifier = 31, Luxembourg	0.0361**	0.0613***	0.0149	-0.0341
	(0.0167)	(0.0221)	(0.0310)	(0.0336)
Country identifier = 33, Portugal	-0.0630***	-0.0424**	-0.0802***	-0.0901***
	(0.0143)	(0.0206)	(0.0257)	(0.0279)
Country identifier = 35, Estonia	-0.175***	-0.181***	-0.178***	-0.161***
	(0.00985)	(0.0137)	(0.0181)	(0.0191)
Constant	-0.792***	-1142	-2483	0.675
	(0.102)	(0.695)	-3298	-1405
Observations	43.415	20.403	12.573	10.249
R-squared	0.208	0.175	0.231	0.262

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.6 Share invested in risky asset

Once an important relationship between underestimation and stock market participation has been found, it would be interesting to understand if there is the same relationship between underestimation and equity portfolio shares.

To investigate this relationship, I define a new variable: *equityshare.* For each respondent in my sample, I know if there is equity market participation.

Once I selected only individuals that participate in the market, for each of them I calculated the equity share. That is, for each respondent, I divided what they have invested in the market by the net household wealth: the result will be the equity share invested in risky assets.

$$y_i = \beta_0 + \beta_1 Underestimation_i + \gamma_1 X_i + \mu_i$$

The model is defined as before, but now *equityshare* is the dependent variable y_i , and the sample is limited only to those households that participate in the market.

The model is estimated through Ordinarily Least Squares, and only the financial respondent for each family has been selected.

Table 8 presents regression results. The sign of underestimation coefficient is in the expected direction, however the coefficient is not statistically different from zero: it seems that in my sample, underestimation does not affect the share invested in risky asset. As opposed to what found by Spaenjers and Spira (2015). In fact, they found that risky asset share is a positive function of subjective life horizon: a one year longer horizon increases by 0.07 percent more the risky asset allocation.

Table 8: *Underestimation and equity share. Model estimated using OLS. Only financial respondents are considered.*

VARIABLES	equityshare
underestimation	-0.00548
	(0.0160)
optimism	-0.0403
	(0.0291)
cognition	-0.00912
	(0.0155)
male	0.0649**
	(0.0275)
Age	-0.0248
	(0.0201)
age2	0.000183
	(0.000144)
Couple	-0.0713**
	(0.0355)
child	0.0351
	(0.0421)
emp	0.00850
	(0.0368)
educ2	-0.00284
	(0.0601)
educ3	0.0205
	(0.0514)
educ4	0.0562
	(0.0532)
hhquart = 2	-0.148***
	(0.0450)
hhquart = 3	-0.170***
	(0.0466)
hhquart = 4	-0.184***
	(0.0488)
adl	0.00580
	(0.0307)
Country identifier = 11, Austria	0.426***
	(0.0783)
Country identifier = 13, Sweden	-0.0216
	(0.0545)
Country identifier = 15, Spain	-0.0469
	(0.106)
Country identifier = 16, Italy	0.0182
	(0.0692)
Country identifier = 17, France	-0.100
	(0.0702)
Country identifier = 18, Denmark	-0.0249
	(0.0579)
Country identifier = 20, Switzerland	0.0885
	(0.0628)
Country identifier = 23, Belgium	0.000881
	(0.0553)
Country identifier = 25, Israel	0.0558
	(0.0985)
Country identifier = 28, Czech Republic	-0.0800
	(0.0814)
Country identifier = 29, Poland	-0.0481
	(0.263)
Country identifier = 31, Luxembourg	-0.0672
	(0.0836)
Country identifier = 33, Portugal	0.00931
	(0.140)
Country identifier = 35, Estonia	-0.101
	(0.0980)
Constant	1143
	(0.695)
Observations	6.933
R-squared	0.014
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

3.6.1 Heckman selection model

The analysis of equity share presents issues of endogenous sample selection. In fact *equityshare* is observed for a subset of my sample, composed only by those individuals that participate in the stock market. This is an example of incidental data truncation: certain variables (*equityshare*) are observed only if other variables take on particular values (stock market participation).
Running a regression with this subsample would result in possible biased estimates.
To address this selection problem, I use the approach proposed by Heckman (1979).
The Heckman two-step procedure involves the estimation of a probit regression (selection equation) in which the probability of observing a positive equity share (participation in the financial market) is estimated. Then, a second regression is estimated through OLS (outcome equation) and uses the information derived from the first stage to correct the selection bias.
The outcome equation incorporates the results from the selected equation through an additionally explanatory variable (inverse mills ratio) containing information to correct the selection bias.

The first stage (selection equation) is defined as follows:

$$Prob(s = 1|Z) = \Phi(Z\delta)$$

Where s is a dummy variable representing if the outcome variable is observed (in this case if *equityshare* is observed), Z is the vector containing the explanatory variables, δ is the vector of unknown parameters and Φ is the cumulative distribution function of the normal standard distribution.

The second stage (outcome equation) will be:

$$y = X\beta + u$$
$$E[y|X, s = 1] = X\beta + E[u|X, s = 1]$$
$$E[y|X, s = 1] = X\beta + \rho\sigma_u\lambda(Z\delta)$$

From the second line of the following equation, it is clear that the conditional expectation of y is related to the probability that y is observed (s=1). The third line assumes that the error terms from the

first and the second stage are jointly normal distributed. ρ is the correlation between the error terms, σ_u is the standard deviation of u and λ is the inverse Mills ratio, from the first stage estimation.
In other words, in the first stage the dependent variable is the stock market participation (equity share is observed if the individual participates in the market), in the second stage the dependent variable is *equityshare.*

Exclusion restrictions

For a correct model identification, an exclusion restriction is necessary: a variable that I can reasonably exclude from the outcome equation, i.e. it is contains in Z but not in X. Precisely, a variable that appears in the selection equation, but does not appear in the outcome equation.
I used two alternative exclusion restrictions.
First, the vector Z contains the following explanatory variables: age and age squared, dummy for gender, children, employment status, couple status, education level, household income, activity daily limitations, variables for optimism and underestimation, country dummies, cognition, and finally the variable *numeracy2*.
Numeracy2 is a proxy for financial literacy derived from a set of questions on numerical ability (such as computation of interest rates, computations of discounts,…).
In the vector X, same variables are included except for *numeracy2*: that is my exclusion restriction.
I aggregate the numeracy2 score in three different categories: low numeracy (1) which contains from 0 to 2 correct answers, medium numeracy (2) which contains from 3 to 4 correct answers, high numeracy (3) which contains 5 correct answers. For example, participants with high numeracy level assigned (3) are participants that answered correctly to 5 questions in the second numeracy test.
The assumption behind this exclusion restriction is the following: people with low numeracy scores have low financial literacy, which the literature shows to be negatively correlated with stock market participation (Lusardi and Mitchell, 2014).
This variable perfectly fits the definition of exclusion restriction: a variable that directly affects the stock market participation but does not affect the share invested.
As an alternative exclusion restriction I used *net wealth quintile dummies*.
This exclusion restriction is motivated by the related literature (Guiso et al. 2003, Spira 2015), and precisely is motivated by evidence on the non-linearity in the relationship between wealth and equity market participation. For example, stock market participation is less sensitive to changes in wealth at a very high (low) level of net wealth. Assuming that equity share does not show this non-linearity relationship makes the exclusion restriction valid.

Table 9: *Heckman selection model. This table shows results of a Heckman selection model that explains equity share accounting for the endogeneity of the equity market participation decision. The first column shows the OLS baseline estimation model. Second and third column shows respectively the outcome equation results for the first strategy (numeracy2 as exclusion restriction), and the second strategy (net wealth quintiles).*

outcome: equityshare	baseline	heckman 1	heckman 2
underestimation	-0.00548	-0.0158	-0.0147
	(0.0160)	(0.0180)	(0.0161)
IMR	-	0.344	0.405***
	-	(0.267)	(0.0529)
p-value exclusion restriction 1 (num)	-	0.000	
p-value exclusion restriction 2 (nwq)	-		0.000

*Note: Control variables for the second and third columns are the same as in the baseline estimation in table 8. For the selection equations are additionally included numeracy2 in column 2 and net wealth quintiles in column 3. The last two lines report the p-value of the test for the significance of the exclusion restrictions. *** significance at the 1% level. Standard errors in parenthesis.*

Table 9 presents results from the two Heckman selection models. First column indicates the baseline OLS model estimation with outcome *equityshare*. The second and third columns report results from the Heckman-selection model, respectively the first and the second strategy. The two variables *numeracy2* and *net wealth quintiles,* are confirmed as valid exclusion restrictions.

In both cases, Heckman correction method confirms what I found with the baseline estimation. Accounting for endogenous sample selection, underestimation still does not influence the share of wealth invested in the equity market.

3.7 Bequest motives: the impact of having children

Previous literature suggests that the time horizon, measure in this thesis by subjective survival probabilities, is likely to matter less in households with bequest motives. Having strong bequest motives makes the individual time horizon irrelevant.
For example, if we consider having at least one child as an indicator for bequest motives, we can assume that households with children consider differently their life horizon: they have the tendency to behave as if their time horizon is infinite.
Since the time horizon is different for households with bequest motives, it is reasonable to expect that the influence of underestimation on both stock market participation and equity share invested has less weight than for households without bequest motives (they do not consider their time horizon infinite).
To study this effect, I created a new interaction variable between children and underestimation: *child_underestimation.* When the *child* dummy variable is equal to one and therefore when the individual has at least one child, this interaction variable shows what is the additional effect of underestimation in the presence of children (bequest motives).

The estimated model is similar to the one in the main regression, but now my variable of interest is the new interaction variable *child_underestimation.*
Table 10 present results. In the first column the outcome is equity share whereas in the second is the stock market participation. The model is estimated through Ordinarily Least Squares and standard errors are adjusted to consider heteroskedasticity. For both regressions, only the financial respondent has been selected.
Coefficients for both models are in line with expectations, it seems that underestimation matter less when considering bequest motives. However, the two coefficients for *child_underestimation* are not statistically significant.

***Table 10**: the effect of bequest motives on equity shares and stock market participation.*

VARIABLES	equityshare	smp
child_underestimation	-0.0152	0.00566
	(0.0471)	(0.00640)
underestimation	0.00786	-0.0129**
	(0.0444)	(0.00603)
child	0.0426	-0.0533***
	(0.0481)	(0.00823)
optimism	-0.0402	0.0389***
	(0.0291)	(0.00465)
cognition	-0.00908	0.0311***
	(0.0155)	(0.00237)
male	0.0648**	0.0401***
	(0.0275)	(0.00452)
Age	-0.0248	0.0232***
	(0.0201)	(0.00312)
age2	0.000183	-0.000141***
	(0.000144)	(2.24e-05)
Couple	-0.0714**	0.00740
	(0.0355)	(0.00545)
emp	0.00862	0.00434
	(0.0368)	(0.00604)
educ2	-0.00280	0.0233***
	(0.0601)	(0.00757)
educ3	0.0206	0.0653***
	(0.0514)	(0.00695)
educ4	0.0562	0.128***
	(0.0532)	(0.00785)
hhquart = 2	-0.148***	0.0423***
	(0.0451)	(0.00610)
hhquart = 3	-0.169***	0.0788***
	(0.0467)	(0.00684)
hhquart = 4	-0.183***	0.159***
	(0.0488)	(0.00740)
adl	0.00562	-0.0134***
	(0.0307)	(0.00304)
Country identifier = 11, Austria	0.426***	-0.0918***
	(0.0783)	(0.0114)
Country identifier = 13, Sweden	-0.0213	0.368***
	(0.0545)	(0.0110)
Country identifier = 15, Spain	-0.0466	-0.106***
	(0.106)	(0.0114)
Country identifier = 16, Italy	0.0185	0.0168
	(0.0692)	(0.0109)
Country identifier = 17, France	-0.0998	-0.0152
	(0.0702)	(0.0111)
Country identifier = 18, Denmark	-0.0247	0.150***
	(0.0579)	(0.0110)
Country identifier = 20, Switzerland	0.0889	0.124***
	(0.0628)	(0.0119)
Country identifier = 23, Belgium	0.00114	0.0988***
	(0.0553)	(0.00990)
Country identifier = 25, Israel	0.0557	-0.0448***
	(0.0985)	(0.0151)
Country identifier = 28, Czech Republic	-0.0802	-0.0932***
	(0.0814)	(0.0109)
Country identifier = 29, Poland	-0.0483	-0.148***
	(0.263)	(0.0145)
Country identifier = 31, Luxembourg	-0.0666	0.0412***
	(0.0836)	(0.0144)
Country identifier = 33, Portugal	0.00940	-0.0530***
	(0.140)	(0.0160)
Country identifier = 35, Estonia	-0.102	-0.157***
	(0.0980)	(0.0100)
Constant	1135.000	-0.806***
	(0.696)	(0.108)
Observations	6.933	30.455
R-squared	0.014	0.213

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Conclusions

The literature suggests that subjective life horizon influences financial decisions. There are papers that studied how age affects portfolio choice but very few have studied the direct effect of subjective survival probabilities on stock market participation. Focusing on this relationship, I conducted an analysis on 15 different countries, finding great heterogeneity between how individuals evaluate their survival probabilities. Does this heterogeneity explain the heterogeneity in the equity market participation?

To answer this question, I use a regression model and my main result is that underestimation of survival probabilities leads to a reduction in stock market participation: 1 standard deviation increase in underestimation reduces by about one percent the stock market participation. The effect appears to be stronger for individuals between 50 and 64 years old, where the underestimation leads to a 2% reduction in stock market participation. Then, I investigate the relationship between underestimation and equity share invested: taking into account endogenous sample selection, underestimation does not seem to influence the share invested in equity market as was instead found by Spaenjers and Spira (2015). Finally, I analyze whether there are differences in the effect of underestimation on both share invested and stock market participation in respondents with bequest motives, proxied by the presence of children. In both cases I find non-significant effects, in opposition to what found by Spaenjers and Spira (2015).

APPENDIX

Section A: objective survival probabilities computation

Life Tables

An ordinary life table is a statistical tool that summarizes the mortality experience of a population and yields information about longevity and life expectation. Life tables are usually constructed separately for men and for women because of their substantially different mortality rates.

I use this tables to compute objective survival probabilities. Precisely, I use the life tables available in the Human Mortality Database[6] and in the national statistics office for each country.

I use only the data from life tables that are calculated by single year of age (complete life tables) focusing the age range from 50 years up to 90 years.

Definition of Life Tables functions

q_x: The probability that a person exact age x will die within one year.

l_x**:** The number of persons surviving to exact age x.

d_x**:** The number of deaths between exact ages x and x+1.

L_x**:** The number of person-years lived between exact ages x and x+1.

e_x**:** The average number of years of life remaining at exact age x.

[6] As done by Angelini and Cavapozzi (2015)

Computation

My function of interest is l_x representing the exact number of people alive at exact age *x.* Usually, the initial population sets by the national statistics offices l_0 is equal to 100,000.

I use the l_x value because it's determinant to calculate the objective survival probability from age *x* to age *x+1* in the following way:

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

where $s_{x,x+n}$ is defined as the objective survival probability of an individual of age *x* to age *x+n,* at a given year *y.* Note that, my subjective survival probabilities are from the wave 6 of SHARE conducted in 2015, therefore I calculated the objective probabilities using life tables from 2015.

Example

The following equation is used to calculate the objective survival probability for a Czech Republic female aged 55 years in 2015 surviving to age 75 years:

$$\frac{l_{75}}{l_{55}} * 100 = \frac{77{,}904}{96{,}390} * 100 = 80.82\ \%$$

By definition: a Czech Republic female aged 55 years in 2015 has a 80.82% chance of surviving at age 75 years.

Section B: figures and tables

Table 11: statistics. *This table presents descriptive statistics for Austria, Germany, Sweden, and Spain (number of observations and mean) for the control and dependent variables used in this thesis. All variables are defined in Table 1.*

AUSTRIA	N	Mean	GERMANY	N	Mean
age	2583	68.55	age	3731	65.85
male	2583	0.40	male	3731	0.47
couple	2583	0.64	couple	3731	0.78
emp	2583	0.25	emp	3731	0.47
child	2583	0.88	child	3731	0.89
educ1	2583	0.11	educ1	3731	0.01
educ2	2583	0.13	educ2	3731	0.10
educ3	2583	0.49	educ3	3731	0.58
educ4	2583	0.27	educ4	3731	0.32
adl	2583	0.17	adl	3731	0.18
optimism	2583	0.57	optimism	3731	0.58
cognition	2583	0.58	cognition	3731	0.39
underestimation	2583	0.48	underestimation	3731	0.62
smp	2583	0.16	smp	3731	0.26
equityshare	421	0.48	equityshare	964	0.19

SWEDEN	N	Mean	SPAIN	N	Mean
age	3108	69.76	age	3709	68.22
male	3108	0.46	male	3709	0.45
couple	3108	0.74	couple	3709	0.79
emp	3108	0.45	emp	3709	0.28
child	3108	0.93	child	3709	0.87
educ1	3108	0.20	educ1	3709	0.55
educ2	3108	0.13	educ2	3709	0.21
educ3	3108	0.34	educ3	3709	0.12
educ4	3108	0.32	educ4	3709	0.12
adl	3108	0.13	adl	3709	0.18
optimism	3108	0.59	optimism	3709	0.36
cognition	3108	0.31	cognition	3709	-0.79
underestimation	3108	0.59	underestimation	3709	0.51
smp	3108	0.64	smp	3709	0.06
equityshare	1972	0.18	equityshare	223	0.13

Table 12: statistics. *This table presents descriptive statistics for Italy, France, Denmark, and Switzerland (number of observations and mean) for the control and dependent variables used in this thesis. All variables are defined in Table 1.*

ITALY	N	Mean	FRANCE	N	Mean
age	4142	66.38	age	3073	66.79
male	4142	0.45	male	3073	0.43
couple	4142	0.77	couple	3073	0.67
emp	4142	0.29	emp	3073	0.35
child	4142	0.87	child	3073	0.91
educ1	4142	0.40	educ1	3073	0.30
educ2	4142	0.27	educ2	3073	0.09
educ3	4142	0.24	educ3	3073	0.38
educ4	4142	0.09	educ4	3073	0.24
adl	4142	0.15	adl	3073	0.21
optimism	4142	0.34	optimism	3073	0.53
cognition	4142	-0.41	cognition	3073	-0.03
underestimation	4142	0.58	underestimation	3073	0.76
smp	4142	0.19	smp	3073	0.21
equityshare	791	0.18	equityshare	639	0.08

DENMARK	N	Mean	SWITZERLAND	N	Mean
age	3143	64.64	age	2298	67.95
male	3143	0.47	male	2298	0.46
couple	3143	0.78	couple	2298	0.74
emp	3143	0.56	emp	2298	0.49
child	3143	0.92	child	2298	0.84
educ1	3143	0.07	educ1	2298	0.08
educ2	3143	0.09	educ2	2298	0.10
educ3	3143	0.38	educ3	2298	0.65
educ4	3143	0.45	educ4	2298	0.17
adl	3143	0.13	adl	2298	0.09
optimism	3143	0.71	optimism	2298	0.75
cognition	3143	0.51	cognition	2298	0.59
underestimation	3143	0.30	underestimation	2298	0.48
smp	3143	0.42	smp	2298	0.41
equityshare	1295	0.18	equityshare	925	0.27

Table 13: statistics. *This table presents descriptive statistics for Belgium, Israel, Czech Republic, and Poland (number of observations and mean) for the control and dependent variables used in this thesis. All variables are defined in Table 1.*

BELGIUM	N	Mean	ISRAEL	N	Mean
age	4729	65.36	age	1222	68.33
male	4729	0.45	male	1222	0.43
couple	4729	0.70	couple	1222	0.75
emp	4729	0.38	emp	1222	0.38
child	4729	0.88	child	1222	0.97
educ1	4729	0.14	educ1	1222	0.23
educ2	4729	0.22	educ2	1222	0.09
educ3	4729	0.28	educ3	1222	0.31
educ4	4729	0.36	educ4	1222	0.37
adl	4729	0.24	adl	1222	0.29
optimism	4729	0.59	optimism	1222	0.50
cognition	4729	0.20	cognition	1222	0.12
underestimation	4729	0.65	underestimation	1222	0.58
smp	4729	0.35	smp	1222	0.20
equityshare	1649	0.19	equityshare	239	0.23

CZECH REPUBLIC	N	Mean	POLAND	N	Mean
age	3400	67.68	age	1349	65.44
male	3400	0.40	male	1349	0.44
couple	3400	0.68	couple	1349	0.72
emp	3400	0.35	emp	1349	0.32
child	3400	0.96	child	1349	0.94
educ1	3400	0.11	educ1	1349	0.29
educ2	3400	0.25	educ2	1349	0.03
educ3	3400	0.49	educ3	1349	0.58
educ4	3400	0.15	educ4	1349	0.10
adl	3400	0.22	adl	1349	0.29
optimism	3400	0.35	optimism	1349	0.41
cognition	3400	0.27	cognition	1349	-0.46
underestimation	3400	0.87	underestimation	1349	0.92
smp	3400	0.12	smp	1349	0.02
equityshare	406	0.12	equityshare	29	0.08

Table 14: statistics. *This table presents descriptive statistics for Luxembourg, Portugal and Estonia (number of observations and mean) for the control and dependent variables used in this thesis. All variables are defined in Table 1.*

LUXEMBOURG	N	Mean	PORTUGAL	N	Mean
age	1256	64.33	age	1231	66.70
male	1256	0.46	male	1231	0.44
couple	1256	0.79	couple	1231	0.81
emp	1256	0.31	emp	1231	0.30
child	1256	0.89	child	1231	0.93
educ1	1256	0.31	educ1	1231	0.68
educ2	1256	0.11	educ2	1231	0.11
educ3	1256	0.43	educ3	1231	0.10
educ4	1256	0.16	educ4	1231	0.10
adl	1256	0.15	adl	1231	0.40
optimism	1256	0.63	optimism	1231	0.18
cognition	1256	0.29	cognition	1231	-0.70
underestimation	1256	0.57	underestimation	1231	0.70
smp	1256	0.26	smp	1231	0.09
equityshare	328	0.11	equityshare	113	0.20

.

ESTONIA	N	Mean
age	4441	67.67
male	4441	0.37
couple	4441	0.63
emp	4441	0.41
child	4441	0.91
educ1	4441	0.04
educ2	4441	0.20
educ3	4441	0.49
educ4	4441	0.27
adl	4441	0.28
optimism	4441	0.36
cognition	4441	0.01
underestimation	4441	0.78
smp	4441	0.05
equityshare	242	0.09

Table 15: statistics. *This table presents descriptive statistics of financial respondents (number of observations and mean) for the control and dependent variables used in this thesis. All variables are defined in Table 1.*

FIN_RESP	N	Mean
age	30455	67.33
male	30455	0.42
couple	30455	0.61
emp	30455	0.37
child	30455	0.89
educ1	30455	0.21
educ2	30455	0.16
educ3	30455	0.39
educ4	30455	0.25
adl	30455	0.21
optimism	30455	0.48
cognition	30455	0.09
underestimation	30455	0.64
smp	30455	0.23
equityshare	6933	0.20

***Figure 4**: subjective and objective survival probabilities for Belgium by age and gender*

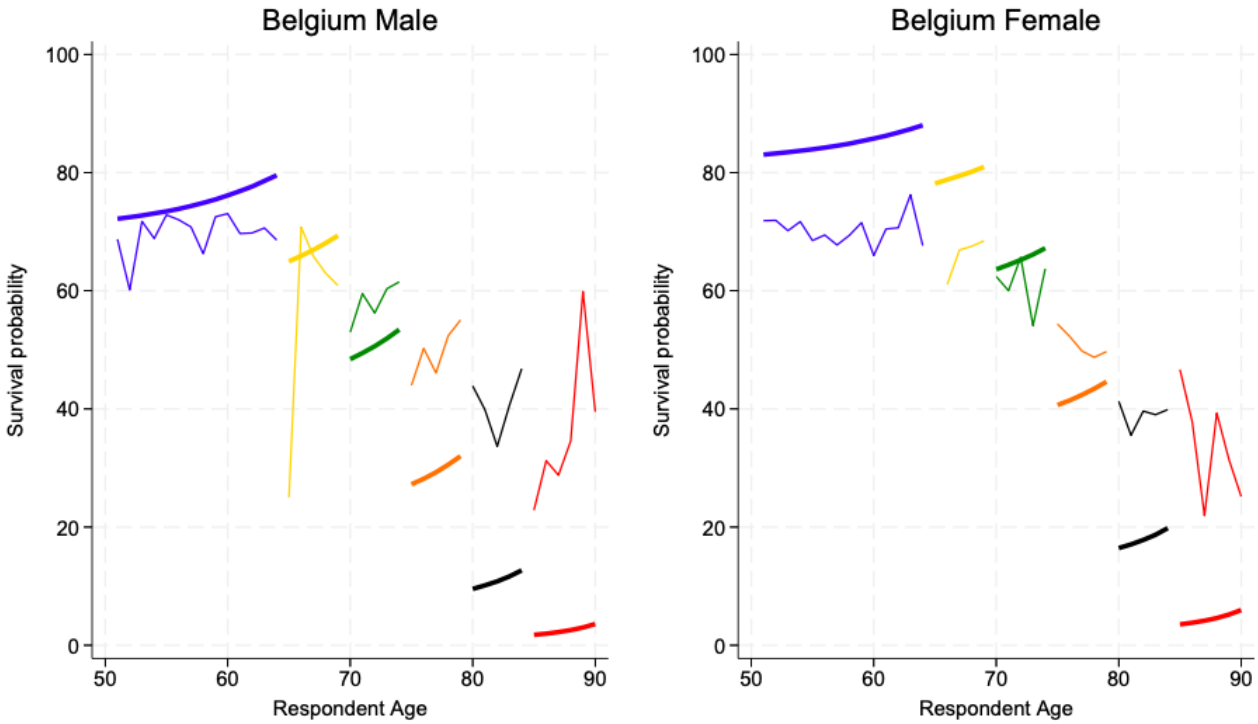


***Figure 5**: subjective and objective survival probabilities for Austria by age and gender*

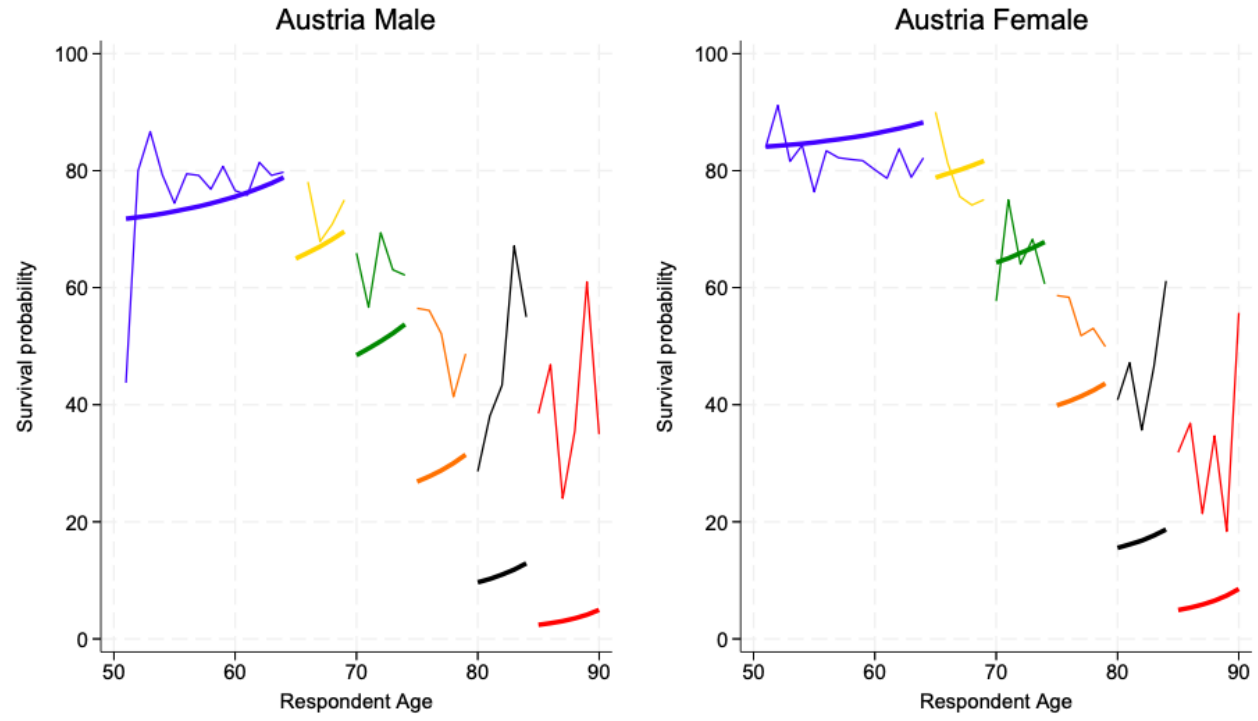


***Figure 6**: subjective and objective survival probabilities for Switzerland by age and gender*

***Figure 7**: subjective and objective survival probabilities for Estonia by age and gender*

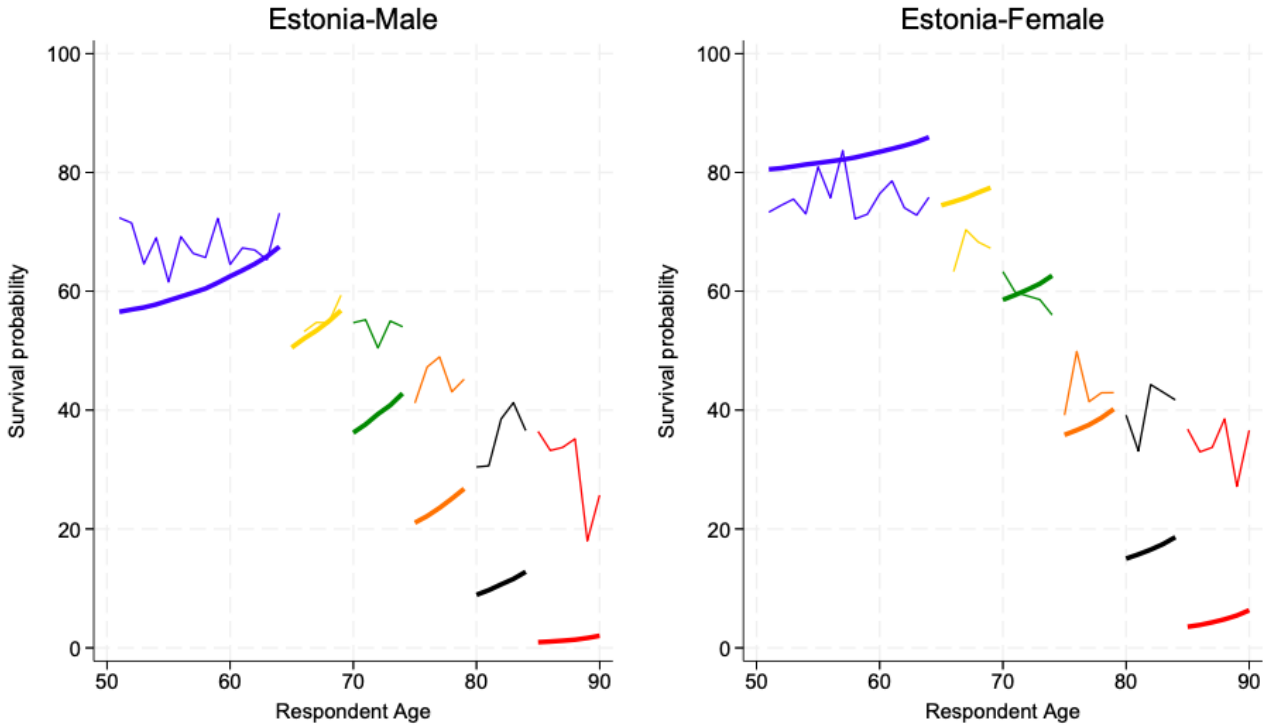


***Figure 8**: subjective and objective survival probabilities for Israel by age and gender*

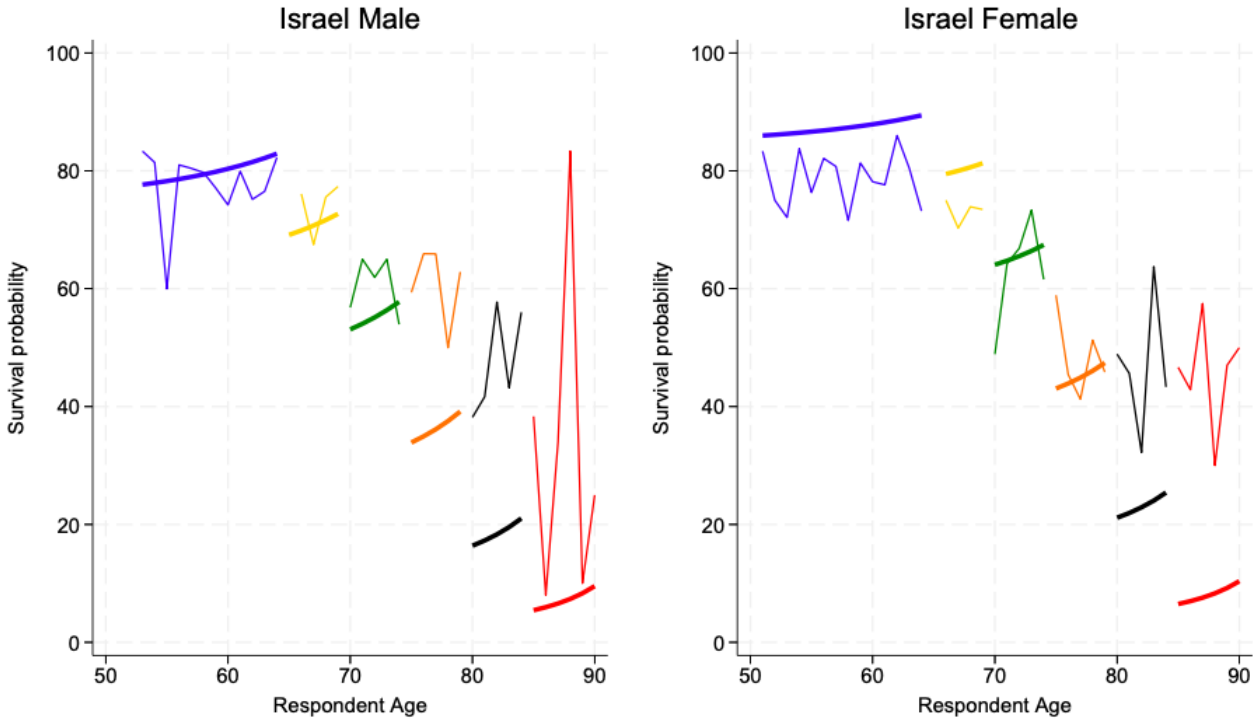


***Figure 9**: subjective and objective survival probabilities for France by age and gender*

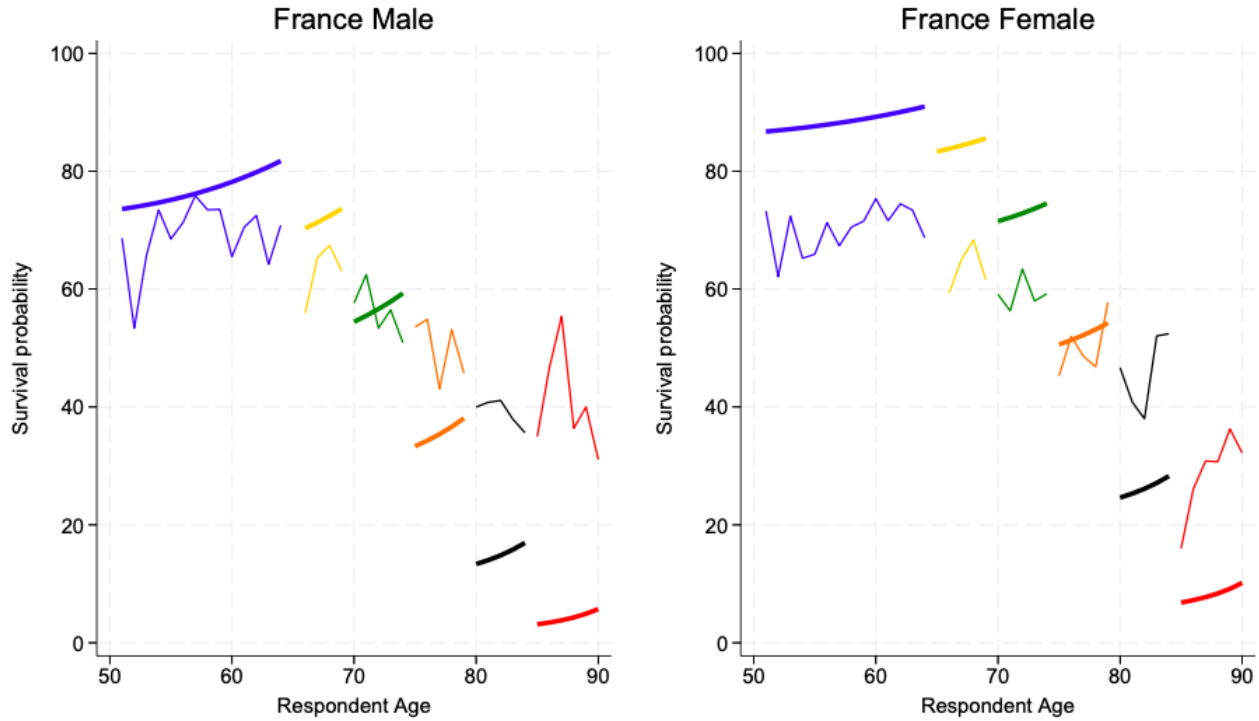


***Figure 10**: subjective and objective survival probabilities for Portugal by age and gender*

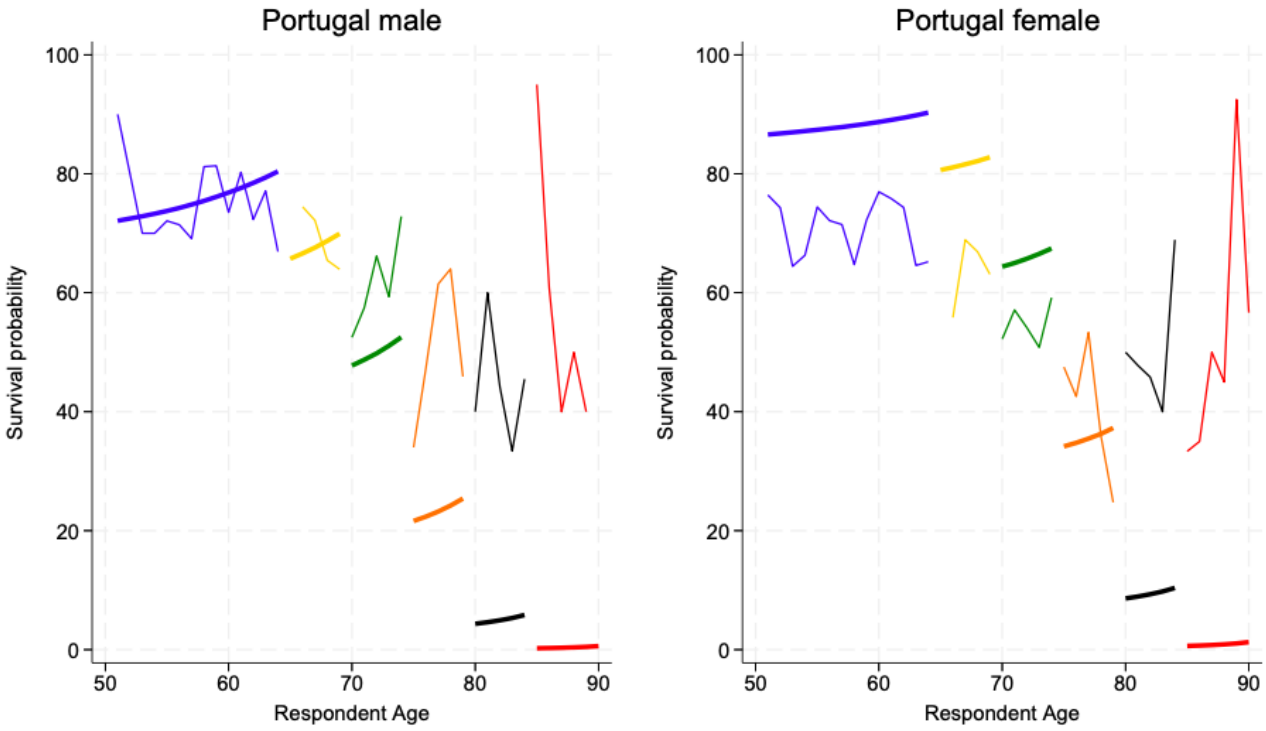


***Figure 11**: subjective and objective survival probabilities for Luxembourg by age and gender*

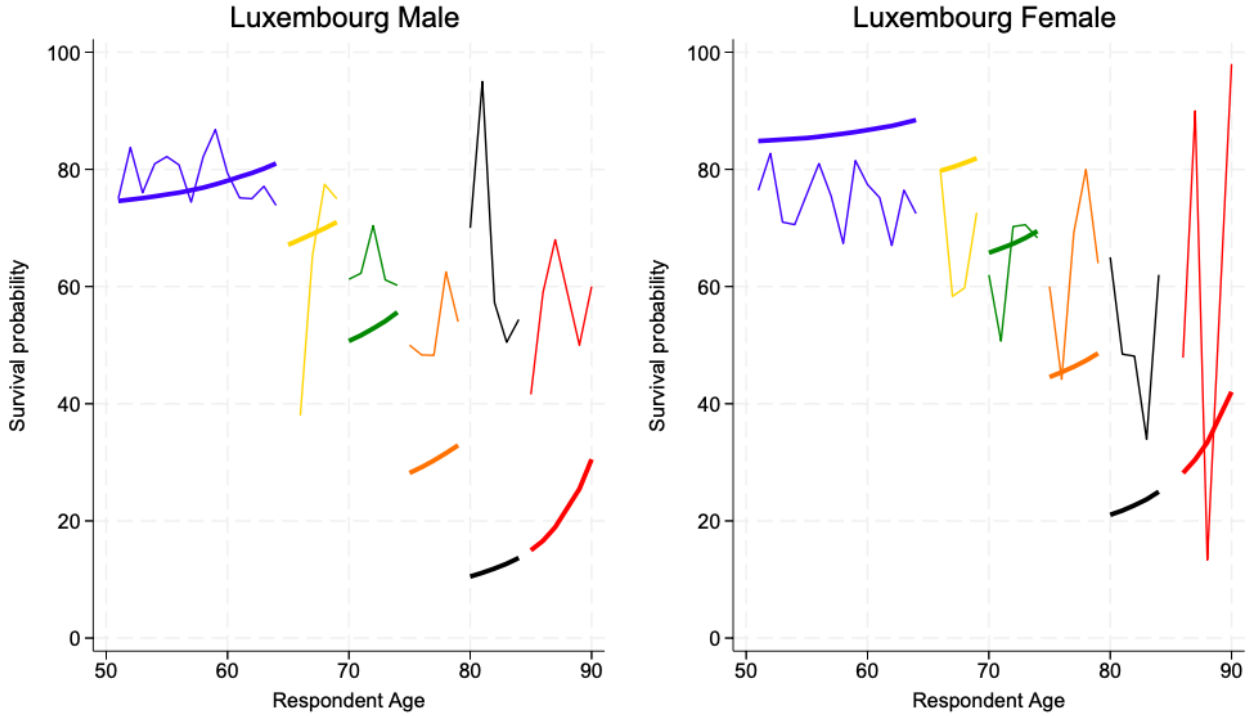


***Figure 12**: subjective and objective survival probabilities for Spain by age and gender*

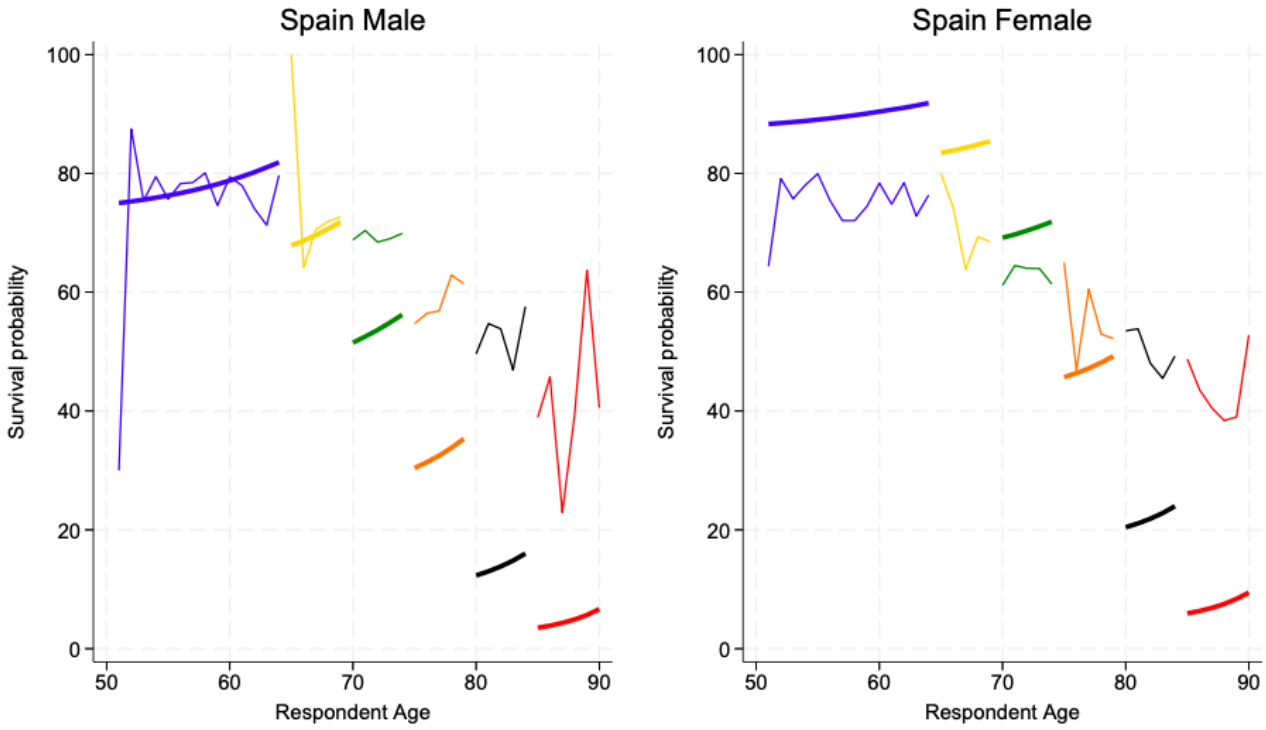


***Figure 13**: subjective and objective survival probabilities for Poland by age and gender*

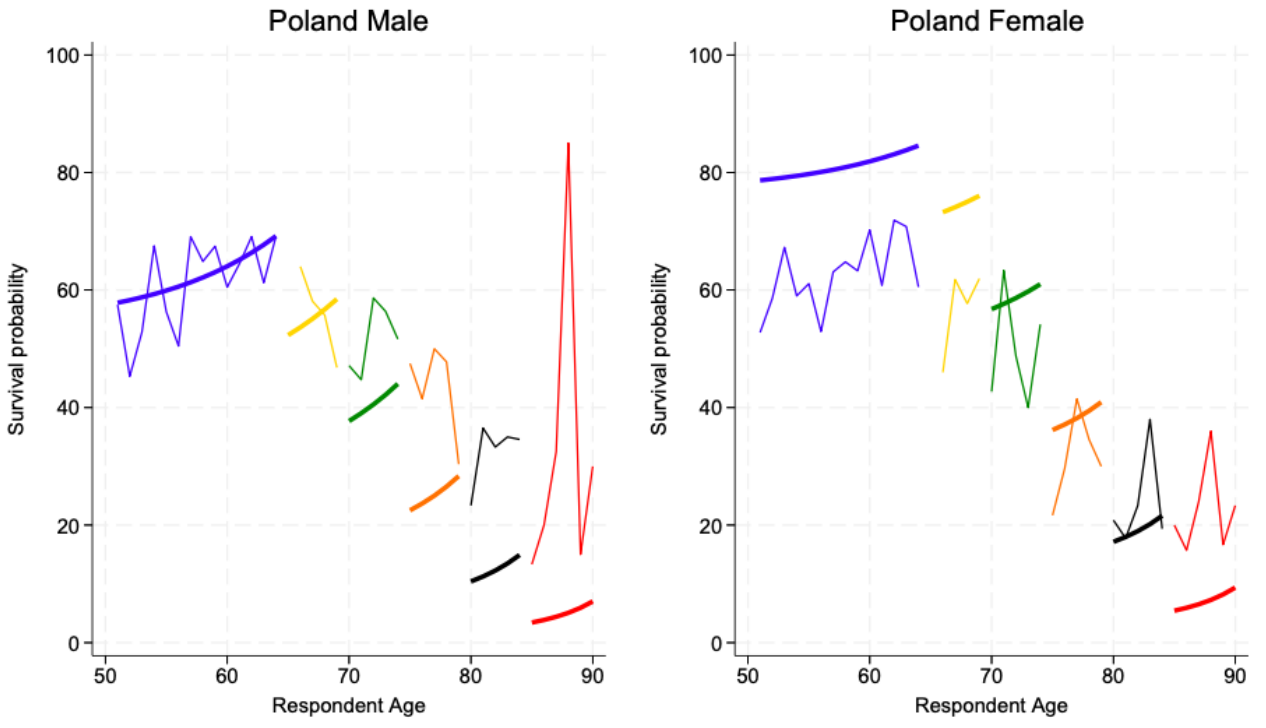


***Figure 14**: subjective and objective survival probabilities for Germany by age and gender*

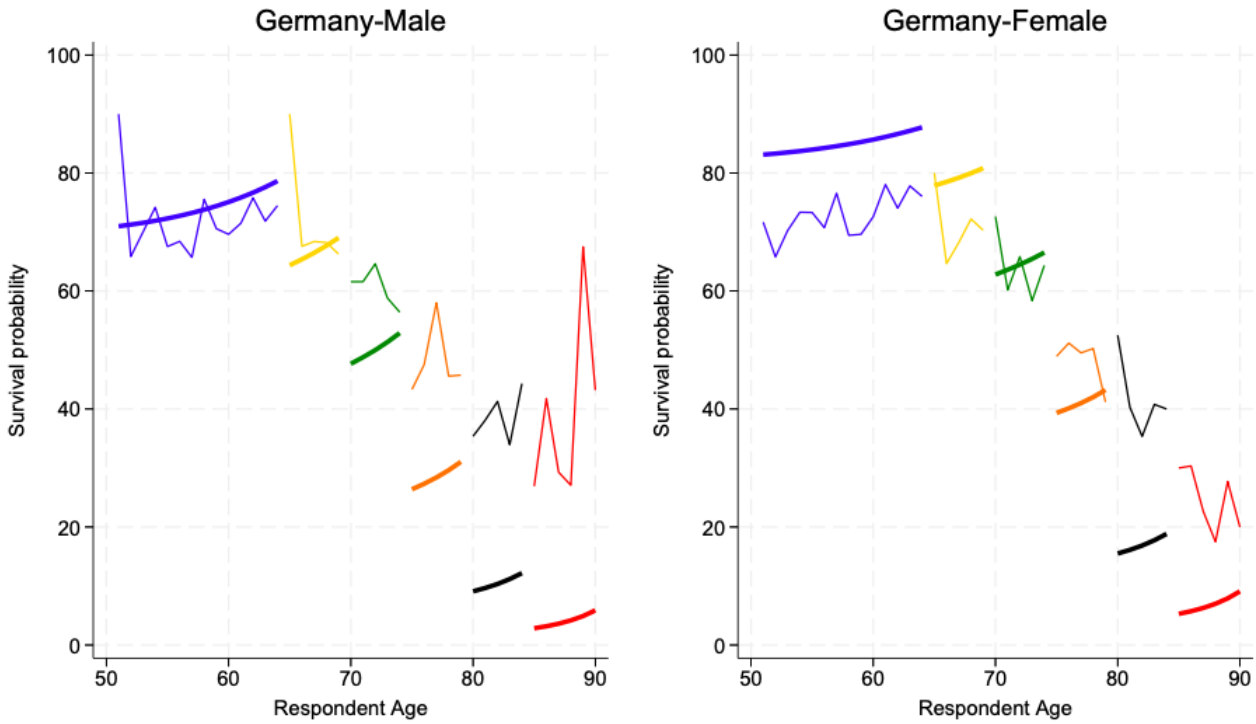
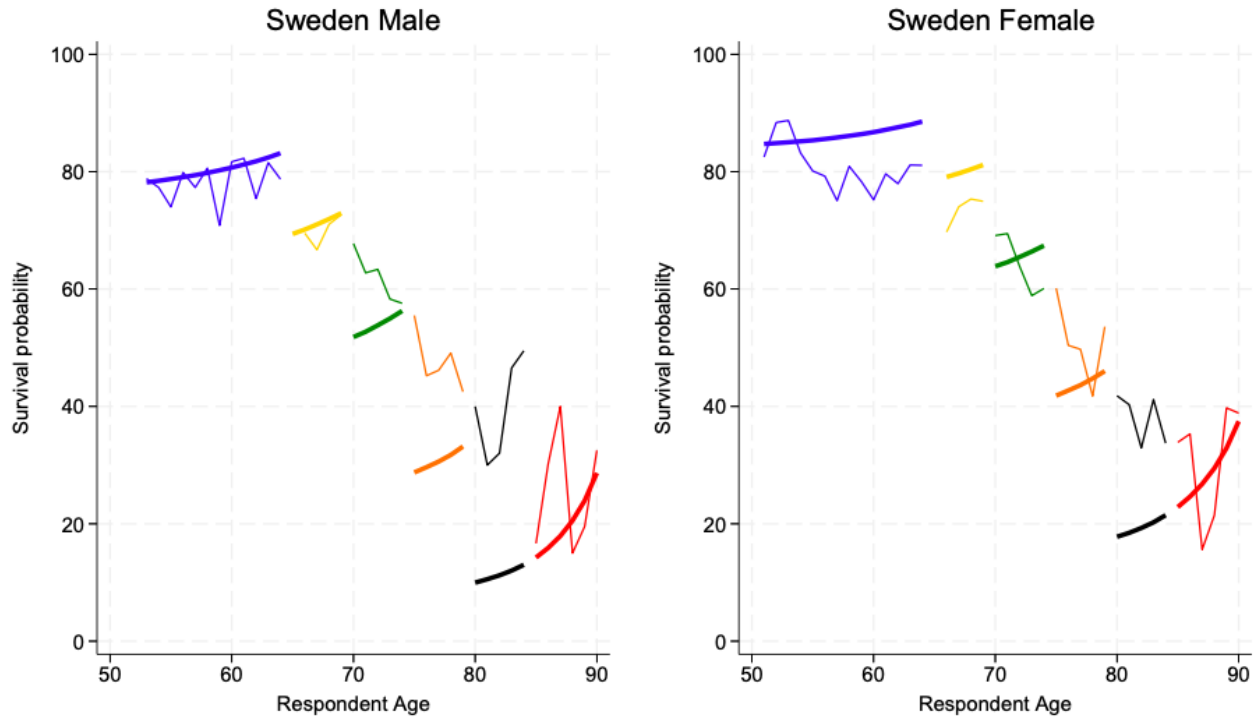


***Figure 15**: subjective and objective survival probabilities for Sweden by age and gender*



References

Angelini, V., Cavapozzi, D., Corazzini, L., and Paccagnella, O., Do Danes and Italians Rate Life Satisfaction in the Same Way? Using Vignettes to Correct for Individual-Specific Scale Biases. OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 76, 5 (2014) 0305–9049 DOI 10.1111/obes.12039

Angelini, V., Cavapozzi, D., Dispositional optimism and stock investments. Journal of economic psychology, 2017, Volume 59, Pages 113-128

Barberis, N., 2000. Investing for the long run when returns are predictable. J. Finance 55, 225–264. http://dx.doi.org/10.2139/ssrn.185376

Bec, F., Gollier, C., 2009. Assets Returns Volatility and Investment Horizon: The French Case. CESifo Working Paper No. 2622.

Benartzi, S., Thaler, R.H., Myopic Loss Aversion and the Equity Premium Puzzle. The Quarterly Journal of Economics, Vol. 110, No. 1 (Feb., 1995), pp. 73-92 (20 pages) https://doi.org/10.2307/2118511

Bergmann, M., T. Kneip, G. De Luca, and A. Scherpenzeel (2019). Survey participation in the Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 1-7. Based on Release 7.0.0. SHARE Working Paper Series 41-2019. Munich: MEA, Max Planck Institute for Social Law and Social Policy.

Billari, F.C, Favero, C.A., Saita, F., Online financial and demographic education for workers: Experimental evidence from an Italian Pension Fund. Journal of Banking and Finance, volume 151 (2023) 106849. https://doi.org/10.1016/j.jbankfin.2023.106849

Bloom, D.E., Canning, D., Moore, M., Song, Y., 2006. The Effect of Subjective Survival Probabilities on Retirement and Wealth in the United States. NBER Working Paper 12688. DOI 10.3386/w12688

Börsch-Supan, A. (2022). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 8.0.0. SHARE ERIC. Data set. DOI: 10.6103/SHARE.w6.800

Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B.Schaan, S. Stuck, S. Zuber (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). International Journal of Epidemiology. DOI: 10.1093/ije/dyt088.

Campbell, J.Y., Viceira, L., 2002. Strategic Asset Allocation: Portfolio Choice for Long-Term Investors. Oxford University Press.

Dimson, E., Marsh, P., Staunton, M., Equity Premia Around the World (October 7, 2011). http://dx.doi.org/10.2139/ssrn.1940165

Fagereng, A., Gottlieb, C., Guiso, L., 2013. Asset Market Participation and Portfolio Choice over the Life-Cycle. Netspar Discussion Paper. https://doi.org/10.1111/jofi.12484

Grevenbrock, N., Groneck, M., Ludwig, A., and Zimper, A. COGNITION, OPTIMISM, AND THE FORMATION OF AGE-DEPENDENT SURVIVAL BELIEFS. INTERNATIONAL ECONOMIC REVIEW Vol. 62, No. 2, May 2021 DOI 10.1111/iere.12497

Guiso, L., Michael, H., Jappelli, T., 2003. Household stockholding in Europe: where do we stand and where do we go? Econ. Policy 18, 123–170. https://www.jstor.org/stable/1344655

GUISO, L., SAPIENZA, P. and ZINGALES, L. (2008), Trusting the Stock Market. The Journal of Finance, 63: 2557-2600. https://doi.org/10.1111/j.1540-6261.2008.01408.x

Heckman, James J. “Sample Selection Bias as a Specification Error.” Econometrica 47, no. 1 (1979): 153–61. https://doi.org/10.2307/1912352.

Hossin, M.Z., The male disadvantage in life expectancy: can we close the gender gap? International Health, Volume 13, Issue 5, September 2021, Pages 482–484, https://doi.org/10.1093/inthealth/ihaa106

Hurd, M.D., 2009. Subjective probabilities in household surveys. Annu. Rev. Econ. 1, 543–564. https://doi.org/10.1146/annurev.economics.050708.142955

Kahneman, D., Tversky, A., Prospect Theory: An Analysis of Decision under Risk, Econometrica, Vol. 47, No. 2 (Mar., 1979), pp. 263-292 (29 pages). https://doi.org/10.2307/1914185

Kaustiaa, M., Conlinb, A., Luotonena, N. What drives stock market participation? The role of institutional, traditional, and behavioral factors. Journal of Banking & Finance, Volume 148, March 2023, 106743 https://doi.org/10.1016/j.jbankfin.2022.106743

Lusardi, A., Mitchell, O., Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education. January 2007 Business Economics 42(1):35-44 http://dx.doi.org/10.2145/20070104

Lusardi, A., Mitchell, O., The Economic Importance of Financial Literacy: Theory and Evidence. April 2013, Journal of Economic Literature 52(1). http://dx.doi.org/10.2139/ssrn.2243635

Mehra, R., Prescott, E.C., The equity premium: A puzzle. Journal of Monetary Economics, Volume 15, Issue 2, March 1985, Pages 145-161 https://doi.org/10.1016/0304-3932(85)90061-3

Perozek, M., 2008. Using subjective expectations to forecast longevity: do survey respondents know something we don't know? Demography 45, 95–113. https://doi.org/10.1353%2Fdem.2008.0010

Poterba, J.M., Samwick, A.A., 2001. Household portfolio allocation over the life-cycle. http://www.nber.org/books/ogur01-1

Puri, M., Robinson, D.T., Optimism and economic choice. Journal of Financial Economics 86 (2007) 71–99. DOI 10.1016/j.jfineco.2006.09.003

Samuelson, P.A., 1969. Portfolio selection by dynamic stochastic programming. Rev. Econ. Stat. 51, 239–246. https://doi.org/10.2307/1926559

Smith, V.K., Taylor Jr., D.H., Sloan, F.A., 2001. Longevity expectations and death: can people predict their own demise? Am. Econ. Rev. 91, 1126–1134. DOI 10.1257/aer.91.4.1126

Spaenjers, C., Spira, S.M., Subjective life horizon and portfolio choice. Journal of Economic Behavior & Organization 116 (2015) 94–106 https://doi.org/10.1016/j.jebo.2015.04.006

Van Rooij, Lusardi, Alessie, Financial Literacy, Retirement Planning and Household Wealth, The Economic Journal, Volume 122, Issue 560, May 2012, Pages 449–478, https://doi.org/10.1111/j.1468-0297.2012.02501.x