STOCK MARKET PARTICIPATION IN THE EUROPEAN UNION

Niels Dillen

Stamnummer: 01506423

Olivier De Lille

Stamnummer: 01511855

Promotor: Prof. dr. Koen Inghelbrecht Commissaris: Nicolas Dierick

Masterproef voorgedragen tot het bekomen van de graad van:

Master of Science in de Handelswetenschappen

Academiejaar: 2017 - 2018



STOCK MARKET PARTICIPATION IN THE EUROPEAN UNION

Niels Dillen

Stamnummer: 01506423

Olivier De Lille

Stamnummer: 01511855

Promotor: Prof. dr. Koen Inghelbrecht Commissaris: Nicolas Dierick

Masterproef voorgedragen tot het bekomen van de graad van:

Master of Science in de Handelswetenschappen

Academiejaar: 2017 - 2018



Permission

We declare that the content of this Master's Dissertation may be consulted and/or reproduced, provided that the source is referenced.

Name students: Niels Dillen and Olivier De Lille

Signature Niels Dillen

Signature Olivier De Lille

Samenvatting

Deze masterproef is opgemaakt om een verklaring te geven voor de lage aandelenmarktparticipatie in de Europese Economische Gemeenschap. Dit gebeurt aan de hand van een onderzoek naar de meest gebruikte determinanten van aandelenparticipatie.

Het onderzoek is opgesteld met data uit de 'European Household Finance and Consumption Survey'. Deze data omvat antwoorden uit vijftien landen, die zich situeren in de Europese Economische Gemeenschap, met in totaal meer dan 62 000 huishoudens. De afhankelijke variabele openbaar verhandelde aandelen en de onafhankelijke variabelen educatie, geslacht, leeftijd, huishoudelijk bruto-inkomen, netto vermogen, aantal leden van het huishouden, arbeidsstatus en investeringshouding worden uit deze enquête gebruikt.

Van de vijftien beschikbare landen zijn er elf onderzocht. De onderzochte landen zijn: Oostenrijk, België, Cyprus, Duitsland, Spanje, Griekenland, Italië, Luxemburg, Nederland, Portugal en Slovenië. Dit geeft ons een totale database van ongeveer 31 400 huishoudens. Het eiland Malta is uit de database gehaald wegens het ontbreken van data in verband met de leeftijd voor de belangrijkste respondent. Finland en Frankrijk behoren niet tot de gebruikte data door het ontbreken van de variabele investeringshouding van de belangrijkste respondent. Slowakije is ook uit de dataset gehaald wegens problemen met convergentie. Hierdoor was het niet mogelijk om een model te schatten.

Van de negen onderzochte variabelen zijn er relatief weinig significante verbanden gevonden in de probit en logit regressies. De tekens van de verbanden liggen wel in de lijn met de literatuur. Vooral voor de variabelen geslacht, leeftijd, educatie en arbeidsstatus verwachtten we significante resultaten in de onderzochte landen. Uit onze resultaten hebben we de relatie tussen aandelenmarktparticipatie en de variabelen huishoudelijk bruto-inkomen, netto vermogen en risicoaversie wel kunnen bevestigen.

Acknowledgements

This master dissertation could not have been written without some external help, support and guidance of particular people. In this way we want to express our gratitude.

First and foremost, we want to thank our commissioner Mr. Nicolas Dierick. Throughout the dissertation he helped us with his expertise, suggestions and criticism. He was always eager to help us when we encountered difficulties or when we had any questions. Without him steering us in the right direction, it would have been impossible to complete this dissertation.

Next, we want to thank Prof. dr. Koen Inghelbrecht for making this dissertation possible as a promoter. His lectures were the foundation of our basic insights in econometrics.

Furthermore, without the help and support of our friends and family the process of writing this dissertation would have been much heavier. Thanks to the rereading by Emma Bolle, Ingrid Defever and Rudi Dillen our writing and grammar was improved.

Finally, if the European Central Bank would not have granted us permission to use the 'Household Finance and Consumption Survey', our research would not have been possible.

Table of contents

1	Introduction1
2	Stock Market Participation
2.1	Hard variables
2.1.1	Age3
2.1.2	Gender3
2.1.3	Education3
2.1.4	IQ3
2.1.5	Income4
2.1.6	Wealth4
2.2	Soft variables4
2.2.1	Participation barriers4
2.2.2	Risk-aversion5
2.2.3	Loss-aversion6
2.2.4	Trust6
2.2.5	Financial literacy6
2.2.6	Social interaction7
2.2.7	Personality traits8
2.2.8	Cognitive and non-cognitive skills8
2.2.9	Awareness9
2.2.10	Dispositional optimism9
2.3	Different perspective10
3	Data and methodology11
3.1	Data11
3.2	Methodology19
4	Results22
4.1	Binary response models22

4.1.1	Risk aversion	23
4.1.2	Gender	24
4.1.3	Age	24
4.1.4	Medium Education	25
4.1.5	High Education	25
4.1.6	Income	26
4.1.7	Wealth	26
4.1.8	Household members	27
4.1.9	Retiree	27
5	Conclusion	28
6	References	31
7	Appendix	34
7.1	Data structure	34
7.2	Frequency distribution of rescaled education levels	37
7.3	Frequency distribution of rescaled risk aversion	37
7.4	Binary response models with standard deviation	38

List of used abbreviations

АТ	Austria
BE	Belgium
CY	Cyprus
DE	Germany
ECB	European Central Bank
ES	Spain
GR	Greece
HFCS	Household Finance and Consumption Survey
IT	Italy
IQ	Intelligence Quotient
LU	Luxembourg
NL	The Netherlands
PISA	Programme for International Student Assessment
PT	Portugal
SI	Slovenia

List of tables

Table 1	Number of households
Table 2	Percentage of stock market participation
Table 3	Mean of main respondents' age
Table 4	Percentage of retirees
Table 5	Mean of household members
Table 6	Binary response models
Table 7	Summary of results from the probit models

List of graphs

Graph 1	Frequency distribution of levels of education
Graph 2	Average level of education
Graph 3	Frequency distribution of gender
Graph 4	Median of income
Graph 5	Median of net wealth
Graph 6	Frequency distribution of investment attitude

1 Introduction

Stock market participation describes the percentage of people who actively invest in the stock market. More specifically on markets and exchanges where the issuing and trading of equities takes place. Many households decide to not participate in the stock market (Haliassos & Bertaut, 1995). The reason behind this financial behaviour is to this day subject of many researches. Pioneers in this field of research concluded: "*Indirect stockholding, the degree of risk aversion alone, heterogeneity of beliefs, habit persistence or borrowing constraints in the form of a lower bound on wealth are found to not explain the phenomenon*" (Haliassos & Bertaut, 1995). There are still different opinions and results found in recent literature (Kaustia & Luotonen, 2016). The equity premium is an important factor to many households' wealth and stock market participation has a decisive effect on the economy of a country. Therefore, it is important to help diffuse the determinants of stock market participation.

In this dissertation the focus lies on participation in publicly traded shares. The stock market participation differs from country to country and between households. Many factors influence the decision to hold publicly traded shares or to stay away from these risky assets. Some of these demographic variables like education, age, gender, wealth and income have been frequently analysed before. Nevertheless, we did not find any existing literature that used the data from the Household Finance and Consumer Survey to relate these frequently analysed factors with stock market participation. The dataset gives the chance to execute this study by providing the necessary variables and making a cross-country comparison in the European Economic Area possible.

The HFCS contains data from fifteen countries with more than 62 000 households. The analyzed countries are Austria, Belgium, Cyprus, Germany, Spain, Greece, Italy, Luxembourg, The Netherlands, Portugal and Slovenia. Our research excluded four countries: Finland, France, Malta and Slovakia. Leaving us with data from eleven countries and around 31 400 households. The HFCS harmonized this microlevel data and is our preferred database to make a comparison between countries. The Survey of Health, Ageing and Retirement in Europe covers individuals aged fifty or older (SHARE, 2013), while HFCS contains micro-level data on households' finances and consumption varying from the age of 16 to 86. Since age is an important determinant, HFCS is chosen for our research.

This dissertation proceeds as follows. Chapter two breaks down the existing literature on explanatory variables of our dependent variable, stock market participation, and gives an overview of its analysed determinants. We divide the frequently used variables from the literature into hard and soft variables to give a clear overview.

Proceeding with chapter three including more information about the data and the structure of the database, followed with the descriptive statistics. Furthermore, we describe the methodology of our research in the second part of this chapter.

Chapter four starts with an overview of the models from the probit and logit regressions of every country complemented with the odds. The next chapter elaborates on the outcomes giving interpretation to the different variables for both regressions together with the interpretation of the odds. To conclude this dissertation, the last chapter offers general conclusions together with a reflection to the existing literature and some incentives for possible research in the future.

In the appendix we give an overview of the used variables from the HFCS data, explanatory graphs to describe the data distribution, the complete model with standard deviations and a do-file with some of our commands we used in STATA.

2 Stock Market Participation

2.1 Hard variables

2.1.1 Age

Age is related to stock market participation as it is a critical component in how individuals adjust their exposure to the stock market. When aging, people adjust their attitude towards stockholding and tend to lower their participation in the stock market. Young investors have a different investment horizon than older investors which plays an important role in the investment strategy (Ameriks & Zeldes, 2004).

2.1.2 Gender

Gender is another important determinant of stock market participation. Men and women have a different investment and handling approach towards the stock market (Almenberg & Dreber, 2015). Women are more reluctant to participate in the stock market than men. Overall, men show a higher stock market participation rate (Jianakoplos & Bernasek, 1998).

2.1.3 Education

Education has a significant impact on stock market participation (Campbell, 2006). Respondents with a high level of education show greater interest and are more likely to invest in the stock market. The type of education also plays an important role in relation to stock market participation. Among economists, investors with a longer education have a higher probability of participating in the stock market than investors with a shorter economic education. Moreover, among highly educated investors, economists show a significantly higher stock market probability than investors with a different educational background (Joensen, Christiansen, & Rangvid, 2007).

2.1.4 IQ

The IQ is further related to education. IQ is found to have a remarkably, significant impact on participation in the stock market and diversification of risk in a later stadium of life. Respondents with a higher IQ are more likely to participate in the stock market (Grinblatt & Linnainmaa, 2011).

2.1.5 Income

Income plays a significant role in stock market participation and its returns. When relating stock market returns to labour income, a positive correlation is found. Meaning the higher the income from labour, the higher the returns from stock markets. Entering the stock market is a lot easier for someone with a high income than someone with a lower income. Furthermore, the marginal benefit of holding money is much lower for a person with a high income (Davidoff, 2006).

2.1.6 Wealth

Labour income is considered as the most important source of wealth for households while the most important asset is housing. It is therefore expected that decisions about housing incorporates the desire to hedge against income risk. Empirical findings show that homeowners are wealthier and have more financial assets than renters. Because homeowners' income covary positively with housing prices, the purchase of stock depends on that factor. Thus, the stock market returns are not only positively correlated with labour income, but also with housing prices (Davidoff, 2006). Furthermore, a significant share of household wealth and income is mostly owned by entrepreneurial households (Gentry & Hubbard, 2004).

2.2 Soft variables

2.2.1 Participation barriers

Participation barriers are holding the potential investors back to invest in the stock market. Examples of barriers are the amount of time someone needs to spend on learning the market's mechanics or the entry costs of the stock market itself. The participation barriers are lowered by a higher level of education. It reduces the cost to enter the stock market by making it easier to understand the different financial assets and the market's mechanics. On the other hand, respondents with a low level of education are reluctant to participate in the stock market (Campbell, 2006). Entry costs to the stock market are often participation barriers for potential investors. Stockholding is mostly concentrated at the upper end of the wealth distribution due to these entry costs. Which implies that households at the lower part of the wealth distribution are discouraged by these costs to enter the stock market (Sule, 2006).

2.2.2 Risk-aversion

The level of risk aversion is related to many hard variables mentioned a priori. Risk aversion has a significant negative relation with education. When more educated, the level of risk aversion decreases (Sung & Hanna, 1996). Risk aversion related to gender shows women are more risk-averse compared to men. This indicates women are reluctant to the participate in the stock market as stocks are considered risky assets (Almenberg & Dreber, 2015). Similar research confirms these findings (Halko, Kaustia, & Alanko, 2012). Particularly interesting is the fact that the research is done in Finland, considered a gender equal country, giving some clear insights. As gender is considered a strong predictor of risk, the returns of households headed by men are influenced by their higher level of risk tolerance. This implies that a household guided by a man is more likely to have riskier assets. Thus, risk aversion is strongly related to these variables. Therefore the level of risk aversion is related to the investment decision of a household (Guiso & Paiella, 2004).

Risk aversion is also related to age. The older we get, the more risk averse we become (Sung & Hanna, 1996). When comparing a salary earner with someone who is retired, a clear difference is seen. The salary earner is more risk tolerant than the retired person (Yao, Wang, & Sharpe, 2011). When getting older and coming closer to retirement, the intention to take a lot of risk reduces significantly. There is simply no time left anymore to recover from losses, if any. This stands opposite against a younger person who still has a lot of time left to make up for losses and act accordingly. These conclusions are supported by the 'life-cycle risk aversion hypothesis' which proves that how older a population gets, the higher the average risk aversion will be (Bakshi & Chen, 1994). Therefore, age is related to risk aversion and to stock market participation as it is a critical component in how individuals adjust their exposure to the stock market. As of today, we see an increase in the importance of individually controlled and managed pension accounts. Meaning that economists and policymakers are getting more and more interested in how individuals allocate their financial portfolios. However, there is no uniform prediction if the share of wealth held in stocks should increase or decrease with age. Professionals give the advice to reduce the part of financial wealth held in equities when aging. In practice, individuals do not gradually decrease equity shares as they age (Ameriks & Zeldes, 2004).

2.2.3 Loss-aversion

Just as risk aversion, loss aversion is noted to play a significant role on stock market participation (Dimmock & Kouwenberg, 2010). Loss-aversion is defined by Kouwenberg as follows: "*Loss-aversion implies that households frame events as either gains or losses relative to a reference point, and weight losses more heavily than gains.*" Regarding to households, those with higher loss-aversion show significantly less effort to participate in equity markets. Furthermore, the probability of direct stockholding reduces significantly more than the probability of owning mutual funds as higher loss-aversion occurs. However, a significant relation between loss-aversion and household portfolio allocations to equity is not found after controlling for sample selection. Loss aversion is thus described as an important feature of households' investment decision making process. It partially helps explain the puzzling phenomenon of household financial behaviour (Dimmock & Kouwenberg, 2010).

2.2.4 Trust

A similar outcome is found when relating education to trust in the stock market. If education level increases, the lack of trust in the stock market decreases (Guiso, Sapienza & Zingales, 2008). Thus, a lack of trust explains why individuals do not participate in the stock market, even if there aren't any other frictions. A better education in terms of the stock market reduces the negative effect of lack of trust. Also, the understanding of the determinants of investors' (which are possibly biased) perception of the trustworthiness of the stock market becomes crucial. This effect of trust is confirmed in more recent literature (Hagman, 2015). There is a clear difference across countries in stock market participation rate, even controlling for wealth levels. Furthermore, the level of trust varies between these countries and is related to stock market participation in his model. It shows an increase of approximately 30% stock market participation per marginal increase of level of trust. Trust in the government institutions clearly plays a severe role in stock market participation.

2.2.5 Financial literacy

Van Rooij, Lusardi and Alessie analysed the stockholding puzzle by relating financial literacy to stock market participation. First, they set up two types of indices to measure the level of financial sophistication. Secondly, they help shaping the

methodology of measuring financial sophistication by creating questions with critically chosen words. Lastly, they contribute to the stockholding puzzle by setting up a model to measure the relationship between this level of financial literacy and stock market participation. Respondents with low levels of financial literacy tend to shy away from the stock market. Richer and more educated households are less likely to make financial mistakes than less educated households (Lusardi, Alessie, & Van Rooij, 2011). This field of research is further expanded by relating overconfidence of financial knowledge to stock market participation (Xia, Wang, & Li, 2014). The financial literacy overconfidence from respondents is positively correlated with stock market participation. If a respondent tends to be overconfident about their financial literacy they are more likely to participate on the stock market but this does not necessarily mean they will have a higher performance. Under-confident respondents on the other hand, show a negative correlation to stock market participation (Xia, Wang, & Li, 2014).

2.2.6 Social interaction

There is an important relationship between social interaction and stock market participation (Hong, Kubik, & Stein, 2004). Households that either know their neighbours or attend church are more likely to invest in the stock market. More precisely they found that social households have a four percent higher probability of participating in the stock market. This effect of sociability is even stronger among white, higher educated households with an above-average wealth, increasing the probability of stock market participation with eight percent. Finally, those states where stock-market participation rates are higher show a much stronger impact of sociability (Hong, Kubik, & Stein, 2004). Further research in this field reveals a great impact of peer performance on future stock market participation of nonparticipants (Kaustia & Knüpfer, 2012). This outcome-based social influence has a significant impact on the stock market entries of new investors. Investors tend to communicate selectively about their returns and refrain from discussing bad outcomes. This social pattern in communication is persuasive for new investors and motivates them to participate in the stock market. More precisely, given a positive return the existing investors gained in a specific month in a neighbourhood, encourage new investors to participate the following month. Only positive returns are related to entry. Stock market participation is not only influenced by peers from the same neighbourhood. There is other evidence of certain family effects on stock

market participation. The within-family and community effects have a significant impact on the subsequent participation of an individual (Hellström, Zetterdahl, & Hanes, 2013). Following positive parental and partner stock market participation, the probability of participating in the stock market increases significantly. This effect is stronger for individuals with relatively low financial literacy and a greater level of interpersonal trust on average. These findings extend to both male and female for within-family effects, while community effects mainly pertain to males. The role of trust plays an important role and is positively related to stock market participation as mentioned before (Hellström, Zetterdahl, & Hanes, 2013).

2.2.7 Personality traits

Other factors determining the level of stock market participation are a person's personality traits (Conlin, Kyröläinen, Kaakinen, Järvelin, Perttunen, & Svento, 2015). More precisely, traits like exploratory excitability, extravagance, sentimentality and dependence have large effects on stock market participation. The first trait, exploratory excitability, forms a subscale of novelty seeking and describes a person's need to actively seek new things. Extravagance is a second subscale of novelty seeking and measures the preference for spending money over saving money. Sentimentality, a subscale of reward dependence, shows how someone is affected by emotional stimuli. The last most important personal trait, dependence, resembles the needs of a person's approval of others. These four traits have the most impact on a person's probability of participating in the stock market (Conlin, Kyröläinen, Kaakinen, Järvelin, Perttunen, & Svento, 2015). This research implies that policies have limited impact, such as improved education, when personality traits have a strong effect on behaviour. Furthermore, these traits help explain the heterogeneity of individual's behaviour towards stock market participation, even amongst the wealthy and well-educated respondents.

2.2.8 Cognitive and non-cognitive skills

A further evolution in the literature relates the cognitive and non-cognitive skills to stock market participation of immigrants and natives in the United States (Luik & Steinhardt, 2016). In line with previous studies of financial behaviour they conclude a substantial gap in stock market participation between these two groups. The gap is partly driven by lower numerical skills, memory skills and less intact mental status of immigrants in terms of cognitive skills. For the non-cognitive skills, the difference in investment horizon and time preference have the biggest impact on explaining the financial gap. Natives tend to have a longer investment horizon than immigrants which explains the differences in time preferences (Luik & Steinhardt, 2016). Wealth, income and education remain the most influential factors in explaining the gap nonetheless.

2.2.9 Awareness

The extent to which consumers are aware of available financial assets depends on the incentives of asset suppliers to spread information about the instruments they issue. A significant fraction of consumers is unaware of the existence of available financial instruments such as stocks and mutual funds (Guiso & Jappelli, 2005). When the cost of spreading information is low and when aware households are more likely to adopt, distributors and producers of financial assets have a stronger incentive to distribute information. The lack of awareness helps to resolve part of the stockholding puzzle. A specific obstacle to stockholding, that goes beyond the generic reference to fixed adoption costs as a cause of nonparticipation typically found in the literature, is the ignorance of investment opportunities. If all investors were aware of risky securities, the level of participation in risky financial markets increases considerably. Entry costs are still an important reason for nonparticipation, even when awareness is present. If the entry costs are low, awareness is increased. The explanation for this interaction can be found in the fact that financial intermediaries' incentive to invest resources to inform potential investors is greater. This further increases stock market participation. Awareness is thus positively affected by demographic variables like education, wealth and income that increase the probability of being active in the stock market. If investors have awareness about the existence of risky securities, there is more participation in risky financial markets. Lower entry costs increase awareness, which in turn increases further stock market participation (Guiso & Jappelli, 2005).

2.2.10 Dispositional optimism

Portfolio decisions of older Europeans are affected by dispositional optimism. This is a potential source of overconfidence and mostly lead to inefficient financial decisions. Analysing the relationship between dispositional optimism and stock market participation in twelve European countries (Angelini & Cavapozzi, 2017), an indicator of dispositional optimism, based on the miscalibration between subjective and objective survival probabilities, is developed. The relationship is estimated for cognitive skills, personality traits and for a set of demographic and socio-economic characteristics. Personality traits are seen as risk aversion, trust and social interactions. Focused on the ownership of stocks, risk tolerant agents are more vulnerable for dispositional optimism. Someone who is rather extremely optimistic than pessimistic, has a significant increase in the probability of holding stock of about thirteen and fourteen percentage points, regarding direct and total ownership. On the other side, for risk averse agents, dispositional optimism is a negligible role player in financial behaviour. An increase of eleven percentage points in the total share of wealth invest in stocks can be attributed to risk tolerant investors whom are extremely positive and have no trust in other people. Pessimistic individuals might develop unattractive beliefs of the actual costs of trading and managing stocks. Even if they are not risk averse, they can decide to not participate in the market. This gives an opportunity to lower barriers to financial market participation. Investing in education and numeracy might offset the negative impact of a negative attitude towards life (Angelini & Cavapozzi, 2017).

2.3 Different perspective

As this field of research is continuously expanded, the variables explaining market participation and their force of explanation is changing as well. Different insights and approaches are being researched. By using a measure of risk aversion, which is shown to be very powerful in explaining actual risk-taking behaviour, and using data from an extended time period (2004-2013), Kaustia and Luotonen provide a new insight in stock market participation. Their research brings a whole new perspective using variables like sociability, cognitive skill and religiosity. Their results are not consistent anymore with earlier findings. One of the reasons for this is the lack of controlling for risk aversion in existing studies they state. The regression model of Kaustia and Luotonen explains thirty percent of the variation, while less than a fifth of the variation is explained, using the frequently used variables. This paper is a precursor to identify more measures relevant for investing in stocks (Kaustia & Luotonen, 2016).

3 Data and methodology

3.1 Data

The Household Finance and Consumption Survey (HFCS) is the dataset we use for our research. The HFCS is a joint project of all the central banks of the Euro system and the National Statistical Institutes of France, Finland and Portugal. It covers all countries in the Euro area excluding Ireland and Estonia (in the first wave). This dataset is conducted at individual country level where each participating institution finances and conducts its own wealth survey. It is an effort of the Governing Council of the European Central Bank to cover the whole euro area with samples that provide representativeness both at the euro aggregate level and the individual country level. The micro-level data has been collected in a harmonised way in fifteen euro area members for a sample of more than 62 000 households. It allows to compare the economic structure and the impact of different institutional features across individual countries, which is a key distinguishing feature of the HFCS (European Central Bank , 2012).

Several Euro system countries already have a similar survey constructed (e.g. DNB Household Survey), while others set up a completely new survey. These existing surveys are harmonised with the new surveys by adding the missing HFCS-specific questions. There are two waves of the HFCS, the first wave started in 2008 and the second wave started in 2015. Our research uses data from the first wave. The survey is conducted over several years and takes time for all individual countries to complete. Therefore, the survey fieldwork is not carried out in the same period of time in all countries and should be carefully used for cross-country comparisons. In particular, wealth and income could sometimes refer to different years. Furthermore, the questionnaires are not identical in all fifteen countries. Taking the cultural and institutional differences between euro area countries in consideration, the formulation of questions varies per country to obtain comparable data. However, the responsible institutions produce harmonised output for their respective country with a common template questionnaire as a benchmark.

The dataset is constructed in three main parts. The first part of the questionnaire, containing most of the questions, refers to the household as a whole. This part is completed by a single person from the respective household, the main respondent. This reference person is chosen according to international standards of Canberra

Group (UNECE, 2011) using these sequential steps; household type, the person with the highest income and the eldest person (European Central Bank, 2013). This is usually the household member taking the lead in financial household decisions. The first part covers questions from following areas: real assets and their financing, liabilities and credit constraints, private businesses and financial assets, intergenerational transfers and gifts and consumption/savings. The other two parts are completed on an individual-household-member-level. These parts question areas as: employment, future pension entitlements and labour-related income (other income sources being covered at the household level). In addition to this dissertation there is a relevant data appendix available for those who take interest in the data structure (European Central Bank, 2012).

In this research models are constructed per country for eleven countries in total. Malta did not provide data of the main respondent's age. Therefore, Malta is excluded from the analysis to avoid omitted variable bias. Finland and France are excluded from the dataset due to the lack of data on the investment attitude of the main respondent. This variable is interpreted as a risk aversion measurement in the model. Due to convergence problems Slovakia is excluded from our research as well. This gives a total population of around 31 400 households over eleven different countries. Table 1 shows an overview of the total amount of households per country.

Country	Households
Slovenia	343
Luxembourg	950
Cyprus	1237
Netherlands	1301
Belgium	2327
Austria	2380
Greece	2971
Germany	3565
Portugal	4404
Spain	6197
Italy	7951

Table 1: Number of households

The models contain nine different variables based on existing literature. We disclose which factors determine the decision to participate in the stock market. The first variable is the dependent variable; stock market participation. This variable is questioned as follows:

"(Do you/does anyone in your household) own stock shares in any publicly traded companies?"

$$0 = No$$

1 = Yes

These results are displayed in percentages in table 2 per country. There is a clear difference in stock market participation between all countries. Cyprus has the highest participation rate compared to other countries in the dataset. Overall, there is a relatively low percentage of respondents participating in the stock market by holding publicly traded shares as confirmed in many previous studies (Haliassos & Bertaut, 1995).

Country	Stock market participation
Greece	2,70%
Italy	4,58%
Austria	5,31%
Portugal	5,40%
Luxembourg	9,97%
Slovenia	9,99%
Spain	10,39%
Germany	10,60%
Netherlands	10,43%
Belgium	14,75%
Cyprus	34,56%

Table 2: Percentage of stock market participation

Based on the previous studies, stated in the second section of this dissertation, the explanatory variables we added to the models are age, gender, income, wealth, education, investment attitude, number of household members and labour status. First, the level of education is added to the model. More precisely, it shows the highest level of education completed by the main respondent. The level of education is divided in four categories as follows:

Level 1 = Primary or below

Level 2 = Lower secondary

Level 3 = Upper secondary + Post-secondary





Graph 1: Frequency distribution of levels of education

Graph 1 shows the frequency distribution of education per country. Level one states the lowest level of education, level four shows the highest level of education. There is a clear difference in level of education between the countries. Portugal stands out from the rest with a very high representation of the lowest education level. Cyprus, Belgium and The Netherlands are most represented in the highest completed level of education. Overall, we can see an average education level of two to three in the following graph. Germany has the highest average of education.



Graph 2: Average level of education

Next, we examine the control variable age. The average age of the main respondents is around fifty years for all countries in the population of the survey. Table 3 shows the mean of the main respondents' age per country.

Country	Age
Luxembourg	49,87
Greece	49,94
Austria	50,98
Cyprus	51,09
Slovenia	51,23
Netherlands	51,67
Germany	51,91
Belgium	52,15
Spain	52,66
Portugal	54,32
Italy	55,85

Table 3: Mean of main respondents' age

The next explanatory variable used in the model is gender. This variable consists of two codes and is introduced as a dummy variable:

0 = Female

1 = Male

Portugal has a very high representation of males in the dataset. Other countries show a more balanced representation of genders in the dataset.



Graph 3: Frequency distribution of gender

The next independent variable included, is income. The variable "total gross household income" is used to interpret income. It is measured as gross income and is defined as the sum of labour and non-labour income for all household members. It is calculated by taking the sum of many underlying types of income. For those who take interest in the underlying structure of income, take a closer look at how variable DI200 is constructed in the data appendix. The following graph shows the median income of the main respondent per year per country. The median is chosen over mean to exclude extreme values. Outliers have an impact on the calculation of mean and could give a misperception of income.



Graph 4: Median of income

There is a clear difference between several countries of the survey for yearly income. Luxembourg is the country with the highest income, having a median of

about 65 000 Euros per year. Portugal has the lowest income of the countries with a median of around 15 000 Euros per year. Overall, most countries have a median income of roughly 30 000 Euros per year.

Furthermore, there is the independent variable wealth. This is calculated by taking the total household assets, excluding public and occupational pension wealth, minus the total outstanding household's liabilities. The median net wealth per country is shown in graph 5. Again, Luxembourg stands out with the highest median of net wealth per main respondent. Slovakia and Germany stand out at the lower bound of the graph with a median net wealth of about 50 000 per respondent.



Graph 5: Median of net wealth

The next variable in our model is the labour status of our respondents. The retirement status is introduced as a dummy variable. There are nine answers in the survey to labour status. The answers are rescaled to a dummy variable with respondents being either retiree or not retired. Table 4 gives an overview of the percentage of retirees per country.

- 0 = Not retired
- 1 = Retired

Country	Retirees
Spain	20,69%
Netherlands	20,72%
Cyprus	23,83%
Luxembourg	24,34%
Greece	27,95%

Germany	29,73%
Belgium	32,47%
Portugal	34,88%
Austria	35,48%
Italy	38,57%
Slovenia	40,58%

Table 4: Percentage of retirees

Furthermore, the amount of household members is another explanatory variable in our model, which can be found in the survey as a derived variable. We use this variable as a measurement of household size. This numerical variable ranges from one to sixteen. Overall, households have around two to three household members on an average base. This is derived from table 5, giving an overview of the average amount of household members per country.

Country	Household members
Germany	2,04
Austria	2,13
Netherlands	2,22
Belgium	2,31
Luxembourg	2,48
Italy	2,53
Slovenia	2,57
Greece	2,64
Spain	2,68
Portugal	2,71
Cyprus	2,76

Table 5: Mean of household members

The last variable used in the models, is the investment attitude of the main respondent. This variable is used as an interpretation of risk aversion. The investment attitude is measured in four different levels of risk.

- 1 = Take substantial financial risks expecting to earn substantial returns
- 2 = Take above average financial risks expecting to earn above average returns
- 3 = Take average financial risks expecting to earn average returns
- 4 = Not willing to take any financial risk

Overall, most respondents are reluctant to taking any financial risk expecting a corresponding return. In all countries minimum fifty percent of the respondents show a zero risk tolerance and choose to not take any financial risk. However, Italy shows a relatively high representation of above average risk with eighteen percent of the respondents. Slovenia, Portugal and Spain are the most conservative countries in the dataset with highest representation of risk aversion.



Graph 6: Frequency distribution of investment attitude

3.2 Methodology

The main goal of this dissertation is to confirm the existing literature on most frequently used determinants influencing stock market participation. In other words, we test if these variables remain their significant impact on stock market participation on a household level per country. To answer this, each country is analysed on their level of stock market participation. We set up regression models per country and make a cross-country analysis. Most frequently used variables include age, gender, wealth, income and education. These variables are complemented in our models with the amount of household members per household, investment attitude and retirement status.

Before using the dataset there are a lot of implications to take into consideration. The dataset consists of three main parts as mentioned before and needs to be merged into one large dataset in order to perform any analyses on the data. Furthermore, this complex survey data is provided with replicate weights by the European Central Bank in order to calculate the correct parameters and their respective significance level. For our models we use an amount of 1000 replicate weights to receive best results, this comes at the expense of computation time (Ertz & Münnich, 2015). This complex survey data is multiple imputed and requires specific measurements to calculate correct estimators. This caused convergence problems with several commands.

After implementing the dataset we set up regression models per country. To measure the maximum likelihood of stock market participation we start with probit models, following with logit models and its odds and odds ratios. Two different binary choice models are set up per country to give an elaborated interpretation. Due to problems with convergence we cannot analyse the margins from our probit models, therefore we choose to also analyse the logit models and use its odds and odds ratios as an interpretation of the impact of the variables. Logit and probit models estimate the maximum likelihood of a certain categorical dependent variable. These type of regression models are most fitted to measure the probability of stock market participation, which is a dichotomous dependent variable.

The main differences with a linear regression comes in three ways. First, the linear regression model captions outcomes of the dependent variable outside the zero and one range. The probit and logit models are S-shaped curves which vary between zero and one. Second, this functional form allows the variables to have different marginal effects while the linear regression assumes the same marginal effect per variable. Third main difference is problems with heteroskedasticity. Using a dichotomous dependent variable will result in error terms taking one of the two possible values, which results in heteroskedasticity. Probit and logit models solve these issues with abnormal distributed error terms. The estimation results from probit and logit models are very similar. The models differ slightly from each other in their approach to measure the effects. The logit model uses the logistic function to transform the linear model. In certain fields of research the one is preferred above the other, but it's a matter of personal preference (Inghelbrecht, 2016).

The odds ratio is a practical measurement to interpret the effect certain explanatory variables have, keeping other variables constant. It is fairly easy to calculate by taking the exponentiation of the coefficients from our logit models, but it has to be interpreted cautiously.

Taking a closer look at all explanatory variables separately, shows some precautions to be made in order to use them in our models. Age is not modified and implemented as a numerical variable ranging from 16 to 85. Gender is introduced as a dummy variable, as stated a priori. Risk aversion is implemented as a dummy variable to simplify our models. We rescale the four levels of investment attitude to zero and one values as follows:

0 = No risk aversion (willing to take financial risk)

1 = Risk averse (not willing to take any financial risk)

Furthermore, we rescaled the levels of education to fit our models. Due to low representation of the lowest education level in several countries convergence is not possible for their models. Therefore, the primary or below education level is merged with lower secondary education level. This gives following three education levels:

Level 1 = Low education (Primary or below + Lower secondary)

Level 2 = Medium education (Upper secondary + Post-secondary)

Level 3 = High education (Tertiary)

This gives a new perspective of the frequency distributions per country, which is found in the appendix. We expect a positive correlation between the level of education and the probability to enter the stock market. A higher level of education should indicate a higher probability of stock market participation.

Wealth and income are introduced as the natural logarithm of their original values to simplify our model. We transform these variables to implement the percentage of change rather than the absolute change in values. To measure the impact of retirement status we use labour status from the dataset. This is introduced as a dummy variable in our model as described a priori. The amount of household members is implemented as a numerical variable. We expect all explanatory variables to have a significant effect on stock market participation on household level per country.

4 Results

4.1 Binary response models

		Probit	Logit	Odds
Austria	Risk aversion	7956909***	-1.7092***	0,1810105
	Gender	.295191**	.6275875**	1,8730863
	Age	0068212	0093725	0,9906713
	Medium Education	.3099766	.6769264	1,9678201
	High Education	.3257826	.73029	2,0756825
	Income	.2016568*	.349997	1,4190633
	Wealth	.2143598***	.4130748***	1,5114581
	Household members	0978005	1869105	0.829518
	Retiree	.3007481	.5011409	1,6506034
elgium	Risk aversion	9323058***	-1.60655***	0,2005784
	Gender	.3051009**	.5201631**	1,682302
	Age	.0137265*	.0246667*	1,0249734
	Medium Education	. 3331043*	.6218964*	1,8624567
	High Education	. 6594521***	1.217461***	3,3785986
	Income	. 1018864	.1842656	1,2023351
	Wealth	.2819771***	.5291073***	1.6974163
	Household members	0609645	1096435	1 1158802
	Retiree	- 0927359	- 1688066	0 8446722
	Retiree	.0527555	.1000000	0,0440722
/prus	Risk aversion	2066439	3443351	0,7086914
	Gender	. 1329607	.197855	1,2187857
	Age	0299156***	0487407***	1 0499481
	Medium Education	1175796	1870934	1 2057399
	High Education	1120848	17/0768	1 1012186
		2627206***	6050107***	1 0212002
		1140000**	10050197****	1,0312003
	wealth	. 1149889***	.1985251*	1,2196026
	Household members	00/2/22	0123531	0,9877229
	Retiree	5593179**	9134022**	0,4011571
ermanv	Risk aversion	687047***	-1.298264***	0,2730053
	Gender	. 1652777	.3203398	1.3775958
	Age	0009987	.0004189	1.000419
	Medium Education	1432018	3545068	1 4254774
	High Education	3138828	6756191	1 9652/93
	Income	30/7/60***	551278**	1 735/605
	Wealth	1026520***	2470775***	1,7554095
	Wediti	120222**	.34/0//3****	1,4149204
	Housenoid members	128323**	22/0/9/*	0,7968573
-	Retiree	0801184	-0.748133	0,4/32493
pain	Risk aversion	6066171***	-1.098***	0,3335375
	Gender	. 1103658	.2184063	1,2440924
	Age	. 0095202*	.0175515**	1,0177064
	Medium Education	. 452435***	.8402552***	2,3169582
	High Education	. 5684529***	1.050289***	2,8584771
	Income	. 2625817**	.4734107**	1,6054606
	Wealth	2815876***	.5519535***	1,7366422
	Household members	0504192	0953879	0.9090203
	Retiree	.1148866	.1823078	1,1999835
reece	Risk aversion	3266868	6852587	0,5039598
	Gender	. 1422875	.3127592	1,3671923
	Age	. 013944*	.0303967*	1,0308634
	Medium Education	. 3984585	.8702045	2,387399
	High Education	. 3290501	.7304726	2,0760615
	Income	. 2379467	.4957693	1,6417608
	Wealth	. 3512857**	.7933679**	2,2108298
	Household members	0016203	- 0001571	0 9998429
	Retiree	0807613	1620578	0,8503921
		. 5557,515	.1020070	0,0000021
aly	Risk aversion	3785665***	8065873***	0,4463788
	Gender	. 3330477***	682491***	1 9788008

	Age	. 0053128	.0097687	1,0098166		
	Medium Education	. 4359322***	.9547389***	2,5979922		
	High Education	. 383775**	.775869**	2,1724792		
	Income	. 4118366***	.7785191**	2,1782441		
	Wealth	. 2661059***	.5489925***	1,7315076		
	Household members	1284115**	2289598**	0,7953605		
	Retiree	1619808	2989556	0,7415923		
				-,		
Luxemboura	Risk aversion	8645893***	-1.573553***	0,2073073		
	Gender	. 133593	.3278729	1,3880125		
	Age	0056179	0148877	0,9852226		
	Medium Education	1061206	0003665	0,9996336		
	High Education	. 4776415	1.046068	2,8464369		
	Income	. 3231237**	.5683255*	1,7653086		
	Wealth	. 1341358**	.227987	1,256069		
	Household members	1453613**	2623359*	0,7692526		
	Retiree	. 0218333	.2808553	1,324262		
Netherlands	Risk aversion	-1.126191***	-2.123099***	0,1196602		
	Gender	.1349787	.2188685	1,2446676		
	Age	.0160303	.0303778	1,0308439		
	Medium Education	.2740914	.5277996	1,6951981		
	High Education	.1899702	.3974187	1,4879788		
	Income	.1285262	.2622133	1,2998038		
	Wealth	.2171491**	.4187373**	1,520041		
	Household members	.0316701	.0329622	1,0335115		
	Retiree	0417783	1204254	0,8865432		
Portugal	Risk aversion	-0.7166433***	-1.24617***	0,2876042		
	Gender	-0.0120997	0411492	0,9596859		
	Age	-0.0004077	0044518	0,9955581		
	Medium Education	0.3199284*	.6672688*	1,9489072		
	High Education	0.3386832*	.6376	1,8919348		
	Income	0.5031443***	.9754794***	2,6524385		
	Wealth	0.2526277***	.5098817***	1,6650942		
	Household members	-0.0807618	1486393	0,8618799		
	Retiree	0.0922852	.2535937	1,2886481		
Slovenia	Risk aversion	-0.6654204***	-1.202041**	0,3005801		
	Gender	0.1033998	.1426079	1,1532775		
	Age	0.0049196	.0120737	1,0121469		
	Medium Education	0.232416	.4971515	1,6440316		
	High Education	0.5726002*	1.051091*	2,8607705		
	Income	0.0218402	.0781557	1,081291		
	Wealth	0.1290745	.2394944	1,2706066		
	Household members	0.0755845	.14525	1,1563286		
	Retiree	0.2635131	.5086381	1,663025		
	* p < 0.05; ** p < 0.01; *** p < 0.001					

Table 6: Binary response models

Table 6 shows the regressions results and the odds for every country analysed. We only report the coefficients in this table to limit the space taken. The complete model with standard deviations is implemented in the appendix. It is important to state that for Slovenia, the sample size is only representative at the euro area level. For all other countries the sample size is chosen to be representative also at the country level (European Central Bank, 2013).

4.1.1 Risk aversion

In our probit models, risk aversion is highly significant in every country analysed, except for Cyprus and Greece. Risk aversion is negatively significant related to stock market participation. This implicates that the more risk averse an individual is, the less likely the individual is to participate in the stock market.

Risk aversion remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models. Only for Slovenia the variable risk aversion is not as highly significant as in the probit model.

The odds for risk aversion with the highest parameter are found in Italy and is interpreted as follows. If all the other variables stay at a constant value, the odds of participating into the stock market for a risk averse individual over the odds of participating into the stock market for a risk-taking individual is 0,44. In percentage, the odds for a risk averse individual are 56% lower than the odds for a risk-taking individual.

4.1.2 Gender

Looking at our probit models, gender is highly significant in Italy while being less significant in Belgium and Austria. Gender is positively significant related to stock market participation. Implicating that if the individual is a man, the more likely the individual is to participate in the stock market.

Gender remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models.

The odds for gender with the lowest parameter are found in Belgium. If all the other variables stay at a constant value, the odds of participating into the stock market for a male over the odds of participating into the stock market for a female is 1,68. In percentage, the odds for males are 68% higher than the odds for females.

4.1.3 Age

In our probit models, age is highly significant in Cyprus, while being less significant in Belgium, Spain and Greece. Age is positively significant related to stock market participation. This implicates that the older an individual is, the more likely the individual is to participate in the stock market.

Age remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models. Only for Spain, a higher level of significance is shown for age. The odds with the highest significance level for age are found in Cyprus. If all the other variables stay at a constant value, the odds of participating into the stock market for an older individual over the odds of participating into the stock market for a younger individual is 1,05. In percentage, the odds for older individuals are 5% higher than the odds for younger individuals.

4.1.4 Medium Education

In our probit models, medium education is highly significant in Spain and Italy while being less significant in Belgium and Portugal. Medium education is positively significant related to stock market participation. This implicates that an individual with a medium education is more likely to participate in the stock market than an individual with a lower education.

Medium education remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models.

The odds where medium education is highly significant are found in Spain. If all the other variables stay at a constant value, the odds of participating into the stock market for a medium educated individual over the odds of participating into the stock market for a low educated individual is 2,32. In percentage, the odds for medium educated individuals are 132% higher than the odds for low educated individuals.

4.1.5 High Education

Looking at our probit models, high education is highly significant in Belgium and Spain while being less significant in Italy, Portugal and Slovenia. High education is positively significant related to stock market participation. Implicating that if an individual has a high education, the more likely the individual is to participate in the stock market.

High education remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models, except for one country. A big difference is seen in Portugal where high education is not significant anymore.

The odds for high education discussed are found in Belgium. If all the other variables stay at a constant value, the odds of participating into the stock market for a high educated individual over the odds of participating into the stock market

for a low educated individual are 3,38. In percentage, a high educated individual is 238% higher than the odds for a low educated individual.

4.1.6 Income

In our probit models, income is highly significant in Cyprus, Germany, Italy and Portugal while being less significant in Austria, Spain and Luxembourg. Income is positively significant related to stock market participation. This implicates that the higher the income of an individual, the more likely the individual is to participate on the stock market.

Income remains significant in most of our logit models and gives approximately the same estimates as the ones from our probit models. Only in Austria the variable income is not significant anymore while for Germany, Italy and Luxembourg it has a lower significance level.

The odds for income with the highest parameter are found in Portugal. If all the other variables stay at a constant value, the odds of participating into the stock market for an individual with a higher income over the odds of participating into the stock market for an individual with a lower income is 2,65. In percentage, the odds for an individual with a higher income is 165% higher than the odds for an individual with a lower income.

4.1.7 Wealth

Looking at our probit models, wealth is highly significant in Austria, Belgium, Germany, Spain, Italy and Portugal while being less significant in Cyprus, Greece, Luxembourg and The Netherlands. Wealth is positively significant related to stock market participation. This implicates that the wealthier an individual is, the more likely the individual is to participate in the stock market.

Wealth remains significant in most of our logit models and gives approximately the same estimates as the ones from our probit models. Only for Luxembourg the variable wealth is not significant anymore

The odds for wealth with the highest parameter are found in Greece. If all the other variables stay at a constant value, the odds of participating into the stock market for a wealthier individual over the odds of participating into the stock market for a less wealthier individual is 2,21. In percentage, the odds for a wealthier individual are 121% higher than the odds for a less wealthier individual.

4.1.8 Household members

In our probit models, household members is significant in Germany, Italy and Luxembourg. Household members is negatively significant related to stock market participation. Implicating that a household with more members is less likely to participate in the stock market than a household with less members.

Household members remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models. Only for Germany and Luxembourg the variable household members is not as highly significant as in the probit model.

The odds for household members with the highest significance are found in Italy. If all the other variables stay at a constant value, the odds of participating into the stock market for a household with more members over the odds of participating into the stock market for a household with less members are 0,80. In percentage, the odds for a household with more members is 20% lower than the odds for a household with less members.

4.1.9 Retiree

In our probit models, retiree is only negatively significant related to stock market participation in Cyprus. This implicates that a retired individual is less likely to participate in the stock market than a non-retired individual.

Retiree remains significant in our logit models, it gives approximately the same estimates as the ones from our probit models.

The odds for retiree in Cyprus is interpreted as follows. If all the other variables stay at a constant value, the odds of participating into the stock market for a retired individual over the odds of participating into the stock market for a non-retired individual is 0,40. In percentage, the odds for a retired individual is 60% lower than the odds for a non-retired individual.

5 Conclusion

Table 7 summarises the non-significant and significant variables in a clear overview. The X's state a significant effect of the explanatory variable in our estimated probit models per country.

	Risk aversion	Age	Gender	Education	Income	Wealth	Household members	Retiree
Austria	Х		Х		Х	Х		
Belgium	Х	Х	Х	х		Х		
Cyprus		Х			Х	Х		Х
Germany	Х				Х	Х	х	
Spain	Х	Х		Х	Х	Х		
Greece		х				Х		
Italy	Х		Х	Х	Х	Х	Х	
Luxembourg	Х				Х	Х	х	
Netherlands	Х					Х		
Portugal	Х			х	Х	Х		
Slovenia	Х			Х				

Table 7: Summary of results from the probit models

Risk aversion, income and wealth are confirmed to have a significant effect in at least seven out of eleven countries. Therefore, our research can confirm the relationship towards stock market participation for these variables. These results correspond with the literature (Davidoff, 2006). The other variables gender, age, education, retiree and household members are only significant in maximum five out of eleven countries. Consequently, our research cannot confirm the relationship towards stock market participation for these variables.

Gender was reported significant for only three out of eleven countries. We cannot confirm the relation towards stock market participation for all countries. The existing literature states women participate less in the stock market (Jianakoplos & Bernasek, 1998). Women are more risk-averse compared to men and consequently are less likely to participate in the stock market (Halko, Kaustia, & Alanko, 2012). Our research could not confirm the significance of gender.

Age is found significant in only four out of eleven countries and does not correspond with the findings in the existing literature (Bakshi & Chen, 1994)(Sung & Hanna, 1996). They state that the older an individual gets, the more risk averse the individual will be. A more recent work is in line with our results, individuals do not gradually decrease equity shares as they age (Ameriks & Zeldes, 2004).

Retirement status is only confirmed for one out of eleven countries. Most estimated models did not report a significant effect of retirement status. Therefore, we cannot confirm the significantly negative effect of retirement concluded by the study of Yao, Wang and Sharpe. The literature indicates that a salary earner will be more risk tolerant than a retired person and therefore have a higher probability to enter the stock market (Yao, Wang, & Sharpe, 2011).

The significant effect of education on stock market participation is found for five out of eleven countries. This positive effect of education is supported in previous literature (Campbell, 2006)(Sung & Hanna, 1996). They indicate that when more educated, the level of risk aversion will decrease and participation in the stock market increases. We cannot confirm this relation on a significant level. The fact that education is not significant in every country could indicate that institutional factors play a more important role than these demographic variables.

The PISA score helps understanding this by giving global rankings on student performance in mathematics, reading and science. The PISA score could be interpreted as a measurement of quality of education. Six out of eleven countries from our dataset scored a higher than average PISA score. However, only two out of these six countries, Belgium and Portugal, showed a significantly positive effect of education on stock market participation. Portugal has the lowest average education level in our database, while Belgium has the third highest educated respondents on average base. We cannot conclude any consistent results out of this. It could be interesting to take a further investigation in this direction for future research. It is necessary to mention that some consideration about the PISA score is needed. While we work with data varying from the age of 16 to 86, the PISA score only works with data from fifteen-year-olds (OECD, 2016).

Overall, we conclude a lack of consistency in significant effect from our explanatory variables. One of the reasons why our models did not report age as a significant determinant of stock market participation could be due to high correlation between risk aversion and age. The same accounts for gender. The implementation of risk aversion together with gender and age could lead to multicollinearity and affect our results, although our risk aversion measurement is quite simple.

6 References

Almenberg, J., & Dreber, A. (2015). Gender, Stock Market Participation and Financial Literacy. *Economics Letters*, *137*, 140-142.

Ameriks, J., & Zeldes, S. P. (2004). How do household portfolio shares vary with age? *working paper, Columbia University*.

Angelini, V., & Cavapozzi, D. (2017). Dispositional optimism and stock investments. *Journal of Economic Psychology*, *59*, 113-128.

Bakshi, G., & Chen, Z. (1994). Baby boom, population aging and capital markets. *Journal of Business*, 165-202.

Calvet, L. E., Campbell, J. Y., & Sodini, P. (2009). Measuring the financial sophistication of households (No. w14699). *National Bureau of Economic Research*.

Campbell, J. Y. (2006). Household finance. *The journal of finance , 61* (4), 1553-1604.

Conlin, A., Kyröläinen, P., Kaakinen, M., Järvelin, M.-R., Perttunen, J., & Svento, R. (2015). Personality traits and stock market participation. *ournal of Empirical Finance*, *33*, pp. 34-50.

Davidoff, T. (2006). Labor income, housing prices and homeownership. *Journal of Urban Economics*, *59* (2), 209-235.

Dimmock, S. G., & Kouwenberg, R. (2010). Loss-aversion and household portfolio choice. *Journal of Empirical Finance*, *17* (3), pp. 441-459.

Ertz, F., & Münnich, R. T. (2015). Pitfalls of regression modellig using complex survey data.

European Central Bank . (2012). HFCS USER DATABASE DESCRIPTION DOC. UDB1 WAVE 1.

European Central Bank. (2012). HFCS CORE VARIABLES CATALOGUE DOC. UDB 2 WAVE 1.

31

European Central Bank. (2013). *The Eurosystem Household Finance and Consumption Survey METHODOLOGICAL REPORT FOR THE FIRST WAVE.*

European Central Bank. (2013). The Eurosystem Household Finance and Consumption Survey Results from the first wave.

Gentry, W. M., & Hubbard, G. R. (2004). Entrepreneurship and Household Saving. *Advances in economic analysis & policy , 4* (1).

Grinblatt, M., & Linnainmaa, J. (2011). IQ and Stock Market Participation. *The Journal of Finance*, 66 (6), pp. 2121-2164.

Guiso, L., & Jappelli, T. (2005). Awareness and stock market participation. *Review* of *Finance*, 9 (4), 537-567.

Guiso, L., & Paiella, M. (2004). The role of risk aversion in predicting individual behaviors.

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *the Journal of Finance*, *63* (6), 2557-2600.

Hagman, M. (2015). The Effects of Trust on Stock Market Participation: A Cross-Sectional Study Based on 15 Countries.

Haliassos, M., & Bertaut, C. C. (1995). Why do so Few Hold Stocks? *The economic journal*, pp. 1110-1129.

Halko, M.-L., Kaustia, M., & Alanko, E. (2012). The gender effect in risky asset holdings. *Journal of Economic Behavior & Organization , 83* (1), 66-81.

Hellström, J., Zetterdahl, E., & Hanes, N. (2013). Loved Ones Matter: Family Effects and Stock Market Participation.

Hong, H., Kubik, J., & Stein, J. C. (2004). Social Interaction and Stock-Market participation. *The journal of finance*, *59* (1), pp. 137-163.

Inghelbrecht, K. (2016). Qualitative Choice Models [Powerpoint slides].

Jianakoplos, N., & Bernasek, A. (1998). Are women more risk averse? *Economic inquiry*, *36* (4), 620-630.

32

Joensen, J. S., Christiansen, C., & Rangvid, J. (2007). Are Economists More Likely to Hold Stocks? *Review of Finance*, *12* (3), pp. 465-496.

Kaustia, M., & Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics*, *104* (2), pp. 321-338.

Kaustia, M., & Luotonen, N. (2016). What drives the heterogeneity in portfolio choice? The role of institutional, traditional, and behavioral factors.

Luik, M.-A., & Steinhardt, M. F. (2016). Immigrant-native differences in stockholding–The role of cognitive and non-cognitive skills. *Journal of Empirical Finance*, *38*, pp. 103-119.

Lusardi, A., Alessie, R., & Van Rooij, M. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101* (2), pp. 449-472.

OECD. (2016). PISA 2015 Results in Focus.

SHARE. (2013). Survey of Health, Ageing and Retirement in Europe.

Sule, A. (2006). Entry costs and stock market participation over the life cycle. *Review of Economic Dynamics*, 9 (4), pp. 588-611.

Sung, J., & Hanna, S. (1996). Factors Related To Risk Tolerance. *Journal of Financial Counseling and Planning*, 7 (11).

UNECE. (2011). Canberra Group handbook on household income statistics.

Xia, T., Wang, Z., & Li, K. (2014). Financial literacy overconfidence and stock market participation. *Social indicators research*, *119* (3), pp. 1233-1245.

Yao, R., Wang, F., & Sharpe, D. L. (2011). Decomposing the Age Effect on Risk Tolerance. *The Journal of Socio-Economics , 40* (6), 879-887.

7 Appendix

7.1 Data structure

ID Variable	Description	Reference unit	Coding
SA0010	Household identification number	Household and individual	Numerical value, 6 digits
SA0100	Country	Household and individual	Two letters country code
IM0100	Implicate ID	Household and individual	Numerical value, 1 digit
RA0010	Personal ID	Household and individual	Numerical value, 2 digits
Household level variable	Description	Question	Coding
HD1500	Household owns publicly traded shares	(Do you/does anyone in your household) own stock shares in any publicly traded companies?	0 - No 1 - Yes
HD1800	Investment attitudes	Which of the following statements comes closest to describing the amount of financial risk that you (and your husband/wife/partner) are willing to take when you save or make investments?	 1- Take substantial financial risks expecting to earn substantial returns 2 - Take above average financial risks expecting to earn above average returns 3 - Take average financial risks expecting to earn average returns 4 - Not willing to take any financial risk
HG0210	Income from regular private transfers	What was the total gross amount over (the last 12 months / the last calendar year)?	Numerical value in EUR, 9 digits
HG0310	Gross rental income from real estate property	What was the total gross amount over (the last 12 months / the last calendar year)?	Numerical value in EUR, 9 digits
HG0410	Gross income from financial investments	What was the total gross amount over (the last 12 months / the last calendar year)?	Numerical value in EUR, 9 digits
HG0510	Gross income from private business other than self- employment	What was the total gross amount over (the last 12 months / the last calendar year)?	Numerical value in EUR, 9 digits
HG0610	Residual income variable	If it's not possible to add an income under its corresponding category or the income does not belong to any of these income components. The residual income variable HG0610 is to be used, with the text description of the kind of	/

	income provided in HG0620.			
Variables for all	Description	Question	Coding	
RA0200	Gender	What is X's(your) gender?	0 – Female 1 – Male	
RA0300	Age	What is X's(your) age?	Numerical value, 3 digits	
Personal questionnaire variables (persons 16 and older)	Description	Question	Coding	
PA0200	Highest level of education completed	What is the highest level of education (you/he/she) (has/have) completed?	Categories based on ISCED- 97 classification: 1 – Primary or below (No formal education or below ISCED 1 + ISCED 1: Primary education) 2 – Lower secondary (ISCED 2: Lower secondary or second stage of basic education) 3 – Upper secondary (ISCED 3: Upper secondary + ISCED 4: Post-secondary) 5 – Tertiary (ISCED 5: First stage tertiary + ISCED 6: Second stage tertiary)	
PE0100\$x	Labour status	What is (your/X's) current employment status. Which categories best describe (your/his/her) situation? Please start with the most important employment status.	 Doing regular work for pay / self-employed/working in family business On sick/maternity/other leave (except holidays), planning to return to work Unemployed Student/pupil/unpaid intern Retiree or early retiree Permanently disabled Compulsory military service or equivalent social service Fulfilling domestic tasks Other not working for pay (specify) 	
PG0110	Gross cash employee income	What was the total gross amount over (the last 12 months / last calendar year)? Please include income from regular wages or salaries, as well as any overtime pay, tips, bonuses, profit sharing benefits	Numerical value in EUR, 9 digits	

		(unless part of the pension arrangements).	
PG0210	Gross self employment income (profit/losses of unincorporated enterprises)	What was the total gross amount (over the whole 12 months / over the whole last calendar year)?	Numerical value in EUR, 9 digits
Derived variables (D-files)	Description	Survey definition	Coding
DH0001	Number of household members	/	Numerical value
DI1100	Employee income	Sum of PG0110 for household members	Numerical value
DI1200	Self-employment income	Sum of PG0210 for household members	Numerical value
DI1500	Income from pensions	Income from public, occupational and private pension plans.	Numerical value
DI1600	Regular socials transfers (except pensions)	Unemployment benefits + gross income from regular social transfers	Numerical value
DI2000	Total household gross income	DI1100+DI1200+DI1500+D I1600+HG0210+HG0310+H G0410+HG0510+HG0610	Numerical value
DN3001	Net wealth	Total household assets excluding public and occupational pension wealth minus total outstanding household's liabilities	Numerical value



7.2 Frequency distribution of rescaled education levels

7.3 Frequency distribution of rescaled risk aversion



7.4 Binary response models with standard deviation

		Probit	Logit	Odds
Austria	Risk aversion	7956909*** (.1272527)	-1.7092*** (.2814156)	0,1810105
	Gender	.295191**	.6275875**	1,8730863
	Age	0068212	0093725	0.9906713
	Madium Education	(.0056216) .3099766	(.0110348) .6769264	1.0070201
	Medium Education	(.2989255)	(.6981411)	1,9678201
	High Education	(.320758)	(.7317467)	2,0756825
	Income	.2016568* (.100738)	.349997 (.1932845)	1,4190633
	Wealth	.2143598***	.4130748***	1,5114581
	Household members	0978005	1869105	0 829518
	Household members	(.0540366)	(.1150424)	0,029510
	Retiree	(.1785902)	(.3491768)	1,6506034
Delaisse		- 9323028***	-1 60655***	
Beigium	Risk aversion	(.1060858)	(.1892024)	0,2005784
	Gender	.3051009** (.0999111)	.5201631** (.1854198)	1,682302
	Age	.0137265*	.0246667*	1.0249734
		(.0057125) . 3331043*	(.0104614) .6218964*	_,
	Medium Education	(.161648)	(.3122201)	1,8624567
	High Education	. 6594521*** (.1600352)	1.21/461*** (.3074949)	3,3785986
	Income	. 1018864	.1842656	1,2023351
	Wealth	(.0638242) .2819771***	.5291073***	1 (074162
	wealth	(.0588149)	(.109216)	1,6974163
	Household members	(.0392588)	(.0722883)	1,1158802
	Retiree	0927359 (.1570091)	1688066 (.2862491)	0,8446722
Cyprus	Risk aversion	2066439 (1234405)	3443351 (2074598)	0,7086914
	Gender	. 1329607	.197855	1,2187857
		(.1134446) . 0299156***	(.1919592) .0487407***	1,220,007
	Age	(.0056363)	(.0094804)	1,0499481
	Medium Education	. 11/5/96 (.1607203)	.1870934 (.2704245)	1,2057399
	High Education	. 1120848	.1749768	1,1912186
	Incomo	. 3637396***	.6050197***	1 0212002
	Income	(.0948842) 1149889**	(.1601348) 1985251*	1,0512005
	Wealth	(.0432007)	(.0778022)	1,2196026
	Household members	0072722 (.0461636)	0123531 (.0777348)	0,9877229
	Retiree	5593179**	9134022**	0,4011571
		(.2010294)	(.559755)	
Germany	Risk aversion	687047***	-1.298264***	0,2730053
	Condor	. 1652777	.3203398	1 3775058
	Gender	(.1054992)	(.1974645) 0004189	1,3773330
	Age	(.0061521)	(.0112992)	1,000419
	Medium Education	. 1432018 (.279578)	.3545068 (.6490432)	1,4254774
	High Education	. 3138828	.6756191	1,9652493
	Income	. 3047469***	.551278**	1.7354695
		(.0842378) . 1836538***	(.1620063) .3470775***	_,, 00 1000
	Wealth	(.0337321)	(.0639049)	1,4149264

	Household members	128323** (.0486818)	2270797* (.0915209)	0,7968573
	Retiree	0801184	-0.748133	0,4732493
		Probit	Logit	Odds
Spain	Risk aversion	6066171*** (.0948354)	-1.098*** (.1719702)	0,3335375
	Gender	. 1103658	.2184063	1,2440924
	Age	. 0095202* (.0038655)	.0175515** (.0076902)	1,0177064
	Medium Education	. 452435*** (.1127724)	.8402552*** (.2176615)	2,3169582
	High Education	. 5684529***	1.050289***	2,8584771
	Income	. 2625817**	.4734107**	1,6054606
	Woolth	(.0789653) . 2815876***	(.1549314) .5519535***	1 7366422
		(.0469705) 0504192	(.0888848) 0953879	1,7500422
	Household members	(.044878)	(.0855472)	0,9090203
	Retiree	(.1120948)	(.2169442)	1,1999835
Greece	Risk aversion	3266868	6852587 (.4422082)	0,5039598
	Gender	. 1422875	.3127592	1,3671923
	Ane	. 013944*	.0303967*	1.0308634
		(.0065978) . 3984585	(.013914) .8702045	2,0000001
	Medium Education	(.2066929)	(.4921722)	2,387399
	High Education	(.2112414)	(.5073082)	2,0760615
	Income	. 2379467 (.1596778)	.4957693 (.328596)	1,6417608
	Wealth	. 3512857**	.7933679**	2,2108298
	Household members	. 0016203	0001571	0,9998429
	Potiroo	(.0749567) 0807613	(.1615677) 1620578	0.8503021
	Retiree	(.1973944)	(.4372434)	0,0505921
Thele		- 3785665***	- 8065873***	
Italy	Risk aversion	(.1058446)	(.2303614)	0,4463788
	Gender	. 3330477*** (.0755813)	.682491*** (.1620507)	1,9788008
	Age	. 0053128	.0097687	1,0098166
	Medium Education	. 4359322***	.9547389***	2,5979922
	High Education	. 383775**	.775869**	2 1724702
		(.1361914) . 4118366***	(.2865964) .7785191**	2,1724732
	Income	(.1340585)	(.2550119)	2,1782441
	Wealth	(.0550228)	(.1132705)	1,7315076
	Household members	1284115** (.038019) 1619808	2289598** (.0747074) 2989556	0,7953605
	Retiree	(.1110817)	(.2275226)	0,7415923
Luxembourg	Risk aversion	8645893*** (.1606397)	-1.573553*** (.3113335)	0,2073073
	Gender	. 133593	.3278729 (3511448)	1,3880125
	Age	0056179	0148877	0,9852226
	Medium Education	(.0119607) 1061206	(.0223181) 0003665	0 0006336
		(.2910653) . 4776415	(.6330847) 1.046068	0,3330000
	High Education	(.2890248)	(.6271948)	2,8464369
	Income	(.1426064)	(.2704708)	1,7653086
	Wealth	. 1341358** (.0660449)	.227987 (.1264394)	1,256069

	Household members	1453613** (0625625)	2623359* (1183036)	0,7692526
	Potiroo	. 0218333	.2808553	1 324262
	Retilee	(.3518548)	(.6368864)	1,524202
		Probit	Logit	Odds
Netherlands	Risk aversion	-1.126191*** (.1680573)	-2.123099*** (.3130759)	0,1196602
	Gender	.1349787 (.1704464)	.2188685 (.3216859)	1,2446676
	Age	.0160303 (.0092295)	.0303778 (.0168298)	1,0308439
	Medium Education	.2740914 (.1970467)	.5277996 (.3930944)	1,6951981
	High Education	.1899702 (.1813193)	.3974187 (.3690552)	1,4879788
	Income	.1285262 (.110862)	.2622133 (.2146895)	1,2998038
	Wealth	.2171491** (.0628331)	.4187373** (.1244524)	1,520041
	Household members	.0316701 (.065844)	.0329622 (.1206395)	1,0335115
	Retiree	0417783 (.2091651)	1204254 (.4032449)	0,8865432
Portugal	Risk aversion	-0.7166433*** (.152723)	-1.24617*** (.2984772)	0,2876042
	Gender	-0.0120997 (0.146277)	0411492 (.3082239)	0,9596859
	Age	-0.0004077 (0.0050874)	0044518 (.010244)	0,9955581
	Medium Education	0.3199284* (0.1538778)	.6672688* (.3201805)	1,9489072
	High Education	0.3386832* (0.1676514)	.6376 (.3351591)	1,8919348
	Income	0.5031443*** (0.1050067)	.9754794*** (.2120371)	2,6524385
	Wealth	0.2526277*** (0.0545938)	.5098817*** (.1116144)	1,6650942
	Household members	-0.0807618 (0.459518)	1486393 (.0913195)	0,8618799
	Retiree	0.0922852 (0.1475802)	.2535937 (.3061608)	1,2886481
Slovenia	Risk aversion	-0.6654204*** (0.1860994)	-1.202041** (.3573598)	0,3005801
	Gender	0.1033998 (0.1535207)	.1426079 (.3015169)	1,1532775
	Age	0.0049196 (0.0078712)	.0120737 (.0161709)	1,0121469
	Medium Education	0.232416 (0.1932593)	.4971515 (.4063941)	1,6440316
	High Education	0.5726002* (0.2232513)	1.051091* (.4667328)	2,8607705
	Income	0.0218402 (0.082454)	.0781557 (.1796628)	1,081291
	Wealth	0.1290745 (0.0661061)	.2394944 (.1422449)	1,2706066
	Household members	0.0755845 (0.0696507)	.14525 (.137741)	1,1563286
	Retiree	0.2635131 (0.2785331)	.5086381 (.5761785)	1,663025

7.5 DO-File from Stata

```
*Descriptive statistics*
*Mean*
use "data", clear
keep if sa0100 == 1
replace pa0200_1 = 0 if pa0200_1 == 1
replace pa0200_1 = 1 if pa0200_1 == 2
replace pa0200_1 = 2 if pa0200_1 == 3
replace pa0200_1 = 3 if pa0200_1 == 5
drop if ra0300_1 == .
drop if dh0001 == .
mi estimate, esampvaryok vceok: svy: mean ra0300_1 dh0001 pa0200_1
```

```
*Median*
use "data", clear
keep if sa0100 == 1
drop if dn3001 == .
drop if di2000 == .
mi estimate, esampvaryok vceok: svy: medianize di2000 dn3001
```

```
*Frequency distribution*
use "data", clear
keep if sa0100 = = 1
replace pa0200 \ 1 = 0 if pa0200 \ 1 = 1
replace pa0200_1 = 1 if pa0200_1 == 2
replace pa0200_1 = 2 if pa0200_1 == 3
replace pa0200_1 = 3 if pa0200_1 = 5
replace hd1500 = 0 if hd1500 = 2
replace ra0200 1 = 0 if ra0200 1 = 2
drop if hd1500 ==.
drop if hd1800 ==.
drop if ra0200_1 = =.
drop if pa0200_1 == .
drop if pe0100a_1 ==.
mi estimate, esampvaryok vceok: svy: proportion hd1500 hd1800 ra0200_1
pa0200_1 pe0100a_1
```

```
*Binary response models*
*Logit*
use "data", clear
keep if sa0100 = = 1
replace pa0200_1 = 0 if pa0200_1 = 1
replace pa0200 \ 1 = 0 if pa0200 \ 1 = 2
replace pa0200 \ 1 = 1 if pa0200 \ 1 = 3
replace pa0200_1 = 2 if pa0200_1 == 5
replace hd1500 = 0 if hd1500 = 2
replace ra0200_1 = 0 if ra0200_1 == 2
drop if hd1500 ==.
drop if hd1800 ==.
drop if ra0200_1 = = .
drop if ra0300_1 = = .
drop if pa0200_1 == .
drop if dn3001 ==.
drop if di2000 ==.
drop if dh0001 ==.
drop if pe0100a \ 1 == .
gen \ln dn3001 = \ln(dn3001)
gen \ln_{di2000} = \ln(di2000)
```

mi estimate, esampvaryok vceok: svy: logit hd1500 4.hd1800 i.ra0200_1 ra0300_1 i.pa0200_1 ln_di2000 ln_dn3001 dh0001 5.pe0100a_1, difficult

```
*probit*
use "data", clear
keep if sa0100 == 1
replace pa0200_1 = 0 if pa0200_1 == 1
replace pa0200_1 = 0 if pa0200_1 == 2
replace pa0200_1 = 1 if pa0200_1 == 3
replace pa0200_1 = 2 if pa0200_1 == 5
replace hd1500 = 0 if hd1500 == 2
replace ra0200_1 = 0 if ra0200_1 == 2
drop if hd1500 == .
drop if hd1800 == .
```

drop if ra0200_1 == . drop if ra0300_1 == . drop if pa0200_1 == . drop if dn3001 == . drop if di2000 == . drop if dh0001 == . drop if pe0100a_1 == . gen ln_dn3001 = ln(dn3001) gen ln_di2000 = ln(di2000)

mi estimate, esampvaryok vceok: svy: probit hd1500 4.hd1800 i.ra0200_1 ra0300_1 i.pa0200_1 ln_di2000 ln_dn3001 dh0001 5.pe0100a_1, difficult