Università degli studi di Padova
Dipartimento di Scienze Statistiche

Corso di Laurea Magistrale

in Scienze Statistiche

Financial ownership in European countries: a multilevel analysis

Relatore: Prof. Omar Paccagnella
Dipartimento di Scienze Statistiche
Università degli Studi di Padova

Co-relatore: Prof. Marcel A. Croon
Department of Methodology and Statistics
Tilburg University (The Netherlands)

Laureanda: Barbara Guarda
Matricola 1013768

Anno Accademico 2012/2013

Index

Introduction

Money is better than poverty, if only for financial reasons.

Woody Allen

Planning of his/her personal finances, a person may want to consider for own future needs a wide range of banking, investment and insurance products. Future planning and investment decisions play an important role to achieve financial stability, especially after retirement.

The topic of this thesis is about the holding of financial assets by European households in the years 2006/2007, whose head is at least 50 years old. It is to assess which might be the demographic or socio-economic factors that might push a households to decide to invest and how much. Much literature exists about this subject. The innovative aspect of this thesis is to consider all these factors regarding the European region of residence, offering consequently models with a structure at two levels.

Three research questions can be developed in this work. The first question is whether countries make a difference with regard to European households portfolio choices. Indeed, it could be argued that the current situation of economy and lifestyle in some countries may condition the behavior of their citizens on risky and not risky financial activities. The second research question will investigate which households' factors and countries characteristics have an impact on ownership of financial assets. Could lower educated people or poor people have a weaker portfolio of financial products compared with other groups? The third question concerns the financial assets, separately considered. Which are the factors affecting

the choice between a product or another one? All of these questions will be treated with the consideration of two level structured data: households and their countries of residence.

Some empirical results are confirmed, such as the strong relationship between real assets and financial assets or the significant effects of some demographic variables and socio-economic conditions of households and countries. What is new is the analysis by considering two levels of statistical units: households and European regions of residence. The data are structured in such a way to observe N units (households) grouped into J groups (regions). Being families naturally assigned to the region, our aim is to define the existence or not of some group effects. To capture some of the economic characteristics and well-being at the country level will be used two well-known and widely used indexes: the Gross Domestic Product and the Human Development Index. The GDP, macroeconomic indicator of development that represents the monetary value of the goods and services produced annually in a given country is accompanied by the HDI, which takes into account factors such as education and life expectancy to assess the quality of life in countries considered. The two indices may be potentially related, but using them as separate items in the analysis trying to explain two different aspects of a country.

Analysis instead will see second-level variables represented by regions, primarily for their relative abundance greater than countries and secondly, to capture any differences in social and demographic characteristics that can vary within the same nation. First, statistical techniques aimed at confirming the validity of the use of a two-tier structure will be used. We try to determine whether the internal homogeneity of the groups and the lack of homogeneity between the groups in terms of outcome variability is such as to suggest that a multilevel structure is suited to explaining the data. The multilevel analysis will be conducted initially for the total of the financial investment, including the following financial assets: bank accounts, bonds, stocks, retirement accounts, contractual savings for housing, mutual funds. Then, the multilevel analysis will be conducted considering each asset individually, in such a way to identify possible determinants for each asset. This will be compared with a multilevel logistic model, whose purpose is to explain the decision to participate in the financial market and no more the amount of the investment. Finally, a further analysis, with the same purpose

of the previous one, will be implemented: it is a *Heckman selection model*, whose advantage is to jointly consider, for units having a particular asset, also the invested amount.

Our detailed analysis of microeconomic sample data documents relatively low financial market participation and persistent differences across countries: many more Swedish, Danish and Swiss households participate on in the stock market than in those from Italy and, especially from Greece, Czech Republic and Poland. One consideration to keep in mind during the analysis will be the coverage of the sample: respondents aged 50 years and older. Conclusions will be carried out by detecting the most relevant results.

1. Why focusing on financial aspect?

'Finance' is a very wide term and it can be used to identify the study of the science of managing funds. It includes public, personal and business finance. It includes also things related to lending, spending and saving money. We are going to focus on personal finance. Personal finance refers to the financial decisions done by households. The aim is to invest or save money. When planning personal finances, a person would consider suitability to their needs of a range of banking, investment and insurance products (stocks, bonds, mutual funds, life insurance, retirement accounts and so on). Saving planning is an important part of long-term plans to gain financial stability, especially after retirement. Probably the financial aspects are among of the most important issues to be considered for the future and, thinking to a long-term plan, managing savings becomes an important component of personal finance. Savings would help people to make investments in the future so that they might have a secure life. But, we also need to keep in mind that "A business that makes nothing but money is a poor business." (Henry Ford)

As underlined in the introduction, the aim of this thesis is to investigate and understanding factors affecting the portfolio choices for European people aged 50 or more. Any evaluations on optimal allocation of financial portfolio will not be subject of this work. More precisely because of the characteristic of analyzed sample, attention will be given to the relationship between retirement and ownership of financial assets during the interpretation of the final results. Is it possible that, because of the ownership of more savings, financial investments

are higher after retirement? In fact, because of the population ageing, the lifestyle of older people is central to the social studying. Furthermore, which are the demographic and economic factors leading an individual to invest in finance? Are there any significant differences between European investors? Might country economy influence the decision? Is the well-being level of the country affecting the choice?

Financial wealth is an important variable to keep in mind when we want to investigate the quality of life of people. How people plan and manage their savings and their portfolio choices could be important to understand welfare level of Europeans. Referring in particular to the age of households, the value of financial assets can be interpreted as an indicator of future lifestyle and it can determine the future amount of liquidity. Consequently, the focus on financial assets aspects were decided because of our interesting in understanding the factors influencing lifestyle of European citizens. Moreover, portfolio choices affect the rate of wealth accumulation, due to differential returns on asset types. The asset specific risk also affects the distribution of wealth (Wolff, 2006).

A large number of studies about households portfolio was done since the 1950s. Some theoretical predictions, collected by Gollier (2002), show how wealthier people should invest more in risky assets than less wealthier households and invest a larger share of their wealth in risky assets (under the assumption of a decreasing absolute risk aversion). Besides, households that are more likely to be liquidity constrained in the future will invest less in risky assets (Gollier, 2002). Owning risky assets is compensated by higher expected returns on one's portfolio. An increase in risk aversion reduces the demand for risky assets. Our attempt is not to describe a dynamic scheme of portfolio choices during time, but the one of finding diversification factors in a static problem. Since of the general purpose of investing in financial assets to think the future after retirement, our interest is concentrated to households aged fifty or more. Some of them are already retired by their jobs, others are still working: is this variable affecting portfolio? Actually, the literature say us that households prefer to spread risk over their lifetime or before retirement: besides, people having pension funds over time will select riskier portfolio structures (Gollier, 2002).

Empirical findings about household finance cover three strongly correlated aspects: participation decision, households portfolio pattern and portfolio span, defined as the number of asset types held. Portfolio span is shown to be strongly associated with income, wealth and education variables (Wolff, 2006). The results of his another analysis indicate that life-cycle factors, such as income, wealth as well as race and education play determinant roles in determining portfolio pattern. Besides, a large number of households have no assets.

Table 1.1 depicts the distribution of portfolio span, by some household characteristics.

Portfolio span									
Age percentiles	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**Total**
<=61	5,9%	6,5%	4,4%	3,2%	1,3%	0,6%	0,1%	0,0%	22%
(61,66]	5,8%	5,2%	3,5%	2,1%	1,3%	0,5%	0,1%	0,0%	19%
(66,73]	8,6%	5,9%	3,5%	2,1%	1,1%	0,3%	0,1%	0,0%	21%
(73,81]	10,8%	5,1%	2,3%	1,0%	0,4%	0,1%	0,0%	0,0%	20%
>81	12,1%	4,0%	1,5%	0,6%	0,2%	0,0%	0,0%	0,0%	18%
Total	43,1%	26,7%	15,1%	9,0%	4,4%	1,4%	0,3%	0,0%	100%
Female									
Not married	21,2%	10,8%	5,3%	2,9%	1,3%	0,4%	0,1%	0,0%	42%
Married	4,6%	3,9%	2,2%	1,3%	0,6%	0,2%	0,0%	0,0%	13%
Male	**17,4%**	**12,0%**	**7,6%**	**4,8%**	**2,5%**	**0,8%**	**0,2%**	**0,0%**	**45%**
Not married	12,7%	8,0%	4,9%	3,1%	1,6%	0,5%	0,1%	0,0%	31%
Married	4,8%	4,0%	2,6%	1,7%	0,8%	0,3%	0,1%	0,0%	14%
Total	43,1%	26,7%	15,1%	9,0%	4,4%	1,4%	0,3%	0,0%	100%
Country									
Austria	1,4%	1,4%	0,7%	0,2%	0,1%	0,0%	0,0%	0,0%	4%
Belgium	3,7%	2,8%	1,6%	0,9%	0,5%	0,2%	0,0%	0,0%	10%
Czech Republic	3,9%	2,3%	1,5%	0,8%	0,2%	0,1%	0,0%	0,0%	9%
Denmark	1,7%	2,2%	1,7%	1,4%	0,7%	0,2%	0,0%	0,0%	8%
France	3,0%	2,6%	1,9%	1,1%	0,5%	0,2%	0,1%	0,0%	9%
Germany	2,6%	2,1%	1,3%	1,0%	0,4%	0,1%	0,0%	0,0%	8%
Greece	6,2%	0,9%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	7%
Italy	5,7%	2,3%	0,6%	0,2%	0,1%	0,0%	0,0%	0,0%	9%
Netherlands	3,8%	2,7%	1,3%	0,4%	0,1%	0,0%	0,0%	0,0%	8%
Poland	4,2%	2,9%	0,9%	0,1%	0,0%	0,0%	0,0%	0,0%	8%
Spain	4,5%	1,2%	0,4%	0,2%	0,1%	0,0%	0,0%	0,0%	6%
Sweden	1,0%	1,7%	2,0%	2,1%	1,4%	0,5%	0,1%	0,0%	9%
Switzerland	1,4%	1,5%	1,0%	0,6%	0,3%	0,1%	0,0%	0,0%	5%
Total	43,1%	26,7%	15,1%	9,0%	4,4%	1,4%	0,3%	0,0%	100%
Risk preferences									
Not willing to take any financial risks	38,9%	21,1%	9,4%	4,2%	1,6%	0,5%	0,1%	0,0%	76%
Above average financial risks expecting to earn above average returns	0,4%	0,8%	1,1%	1,4%	1,0%	0,4%	0,1%	0,0%	5%
Average financial risks expecting to earn average returns	3,6%	4,6%	4,4%	3,2%	1,7%	0,5%	0,1%	0,0%	18%
High financial risks expecting to earn substantial returns	0,2%	0,3%	0,3%	0,2%	0,1%	0,1%	0,0%	0,0%	1%
Total	43,1%	26,7%	15,1%	9,0%	4,4%	1,4%	0,3%	0,0%	100%
Employed	26,8%	14,2%	7,1%	3,3%	1,5%	0,3%	0,1%	0,0%	53%
Other	16,4%	12,5%	8,0%	5,7%	2,9%	1,1%	0,2%	0,0%	47%
Total	43,1%	26,7%	15,1%	9,0%	4,4%	1,4%	0,3%	0,0%	100%

Table 1.1 Portfolio span, by some characteristics.

Source: SHARE data.

Overall household portfolio span in our sample varies among different demographics groups. Table 1.1 shows portfolio span values by other variables, such as gender, country, age, risk preferences and job situation. According to the age distribution of the whole sample, the 20°-percentiles are at 61, 66, 73 and 81 years old respectively: portfolio span decreases with the increasing of ageing. For each different class of positive span, the number of households is decreasing with age. Anyway, the "null-portfolio" is more characteristic for older people. Other factors, such as gender, marital status (or having a partner), country of residence and preferences about financial risk seem to affect portfolio span. Female heads of household, not married or with low risk preferences, households living in Poland, Czech Republic or Greece have a less diversified portfolio. If we look at gender influence, it is notable that female heads tend to have a smaller span than the male. Females tend to prefer a reduced span and most of them prefer to invest nothing. In particular, the most of females answering the questionnaire is not married and it seems that they prefer to not invest. Risk preferences are strongly affecting portfolio choices. Households choosing no risks see a decreasing distribution of portfolio span. The same for people preferring average financial risks. Instead, for people who decide to take substantial risks, this distribution is hump shaped, because it has the maximum around 1 or 2 assets. We see two households having 7 assets and preferring average financial risk.

1.1 Financial and real assets

We are asking 'how households use financial instruments to attain their objectives'. Households must plan over long but finite horizons; they have important non-traded assets, notably their human capital; they hold illiquid assets, notably housing; they face constraints on their ability to borrow; and they are subject to complex taxation (Campbell, 2006).

Households allocate their wealth into financial and real assets. Holding financial or real assets could help people to face retirement. Financial assets are economic resources. They are something that could be owned or controlled to produce a positive value. There are many kinds of financial assets for personal finance, such as bank accounts, bonds, stocks, mutual funds, individual retirement accounts, contractual savings and life insurance policies.

Household financial assets include cash, shares, funds, insurances and constitute an important part of the overall wealth and a source of revenue. In general, data on household participation in financial market play an important role in economic analyses. For instance they are used by governments in setting social protection policies, especially pension provisions, because they give an indication of how they are prepared for the future. Besides, considered with other factors like economic growth, they give an indication of how households would be able to face an economic downturn, if it happened.

There are different classifications for assets. We can discern risky financial assets, which may include indirect stockholdings such as mutual funds, retirement accounts and stocks, bonds or mortgages. However, there are also other kind of risky assets, including real estates and own businesses. Table 1.1 lists household assets considered in this thesis. We will include into the analysis also liabilities, such as mortgages and different types of loans (Haliassos at al, 2002).

Bank accounts	*A bank account is a financial arrangement between a depositor or debt holder and a bank. It considers the financial transactions between the customer and the bank. They can be of various type. In fact, one where a credit is maintained is called a deposit account, while an account to which a customer has a debt with the bank is called loan account. The most familiar kind of deposit account is savings' ones, while a loan account is best represented by a mortgage held by the bank.*
Bonds	*A bond is a tool of indebtedness of the bond issuer to the holders. It is a debt security, under which an individual has a debt with the holder and, depending on the terms of the bond, has to pay interest and/or to repay the principal at a later date.*
Stocks	*A stock is a portion of a corporation. It involves a claim on part of the corporation. Ownership is determined by the number of shares a person owns relative to the number of total shares. Stock prices are usually driven by expectations of corporate earnings. The stockholders have profit if the company pays a dividend. It is a way to reward stockholders, who are the actual owners of the company, for their investment.*
Mutual funds	*Mutual fund is a fund, managed by an investment company with the financial objective of generating high returns. It is a pool of funds collected from many investors with the aim to invest them in different stocks, bonds and similar assets. Collecting the money from investors, mutual funds invest in other assets constituting the patrimony of the undivided fund, of which each investor owns a certain number of shares. One of the main merits of mutual funds is that they give small investors access to professionally managed portfolios of stocks, bonds and other securities, which would be quite difficult to create with a small amount of capital. Each shareholder participates proportionally in the gain or loss of the fund.*

Individual retirement accounts	*An Individual Retirement Account is a form of retirement plan, provided by many financial institutions. It is an investing tool used by individuals to earn and earmark funds for retirement savings.*
Contractual savings for housing	*Contractual savings are defined as "housing solutions for those families whose needs cannot be met at market conditions and for which there are allocation rules. "* *From the perspective of households, Contractual Savings for Housing contracts facilitate the accumulation of equity and offer the prospect of a low-interest loan. The EU member states are characterized by the heterogeneity of housing situations and national of policies and the concept of social housing tends to vary from one country to another.*
Life insurance policies	*Life insurance is a contract between an individual and an insurer, where the insurer promises to pay a designated sum of money after the death of the insurance policy holder. Most people hold life insurance just for the ultimate payment in order to provide for their dependents. Life insurance policies can typically be of two major types: term insurance and permanent insurance. Term insurance will pay a death benefit only if the insured dies during the term of the policy. No benefits are paid if the insured lives beyond the term of the policy. For this reason, term insurance policies will carry the lowest premiums in the earlier years of the policy. However, as an individual gets older, term insurance gets more expensive A permanent life insurance policy, often referred to as whole life insurance, is intended to provide insurance coverage throughout the life of the insured. It is a life insurance policy that remains in force for the insured's whole life. Because of the inevitable death of the insurance holder, the insurance payout is made to the contract's beneficiaries. They also include a savings component, which accumulates a cash value in order to keep the premiums level regular (the cost of a whole life policy is somewhat more expensive than for a term policy in the earlier years because, as the insured gets older, their mortality rate increases).*

Mortgage	*The mortgage is a security interest on one thing others, set up to serve as security for a debt. However, it does not result in loss of possession by the debtor-owner of the asset subject of the asset. Mortgages are also known as "liens against property" or "claims on property."*
Housing	*The most important illiquid asset in household portfolios is housing. Houses are like long-term bonds and can be used to hedge changes in the relative price of housing and nonhousing consumption (Pelizzon and Weber (2005), Sinai and Souleles (2005)). Since houses are illiquid assets, homeowners find it costly to adjust their consumption of housing services in response to economic shocks. This illiquidity may discourage homeownership and financial risk taking by homeowners.*

Table 1.2 Financial assets.

2. The SHARE project

The question isn't at what age I want to retire, it's at what income.

George Foreman

In order to analyze financial participation of the people before or during retirement we choose to use data obtained from survey SHARE, the survey of Health, Ageing and Retirement in Europe (Borsch-Supan et al.).

Its aim is to collect multidisciplinary information about health, socio-economic status and social and family relationships of households aged 50 or more. It is coordinated centrally at the Munich Research Institute for the Economics of Aging. Three surveys took place until now. The first in 2004/2005, while the second one in 2006/2007. In our analysis we use data that relate only to second survey (Wave 2), available since 2006. Two new EU member states (the Czech Republic and Poland) have joined SHARE, in addition to Denmark, Sweden, Austria, France, Germany, Switzerland, Netherlands, Belgium, Italy, Spain and Greece.

The questionnaires in each wave have the following characteristics:

- Multidisciplinary
- Cross-nationality
- Lengthways

The collected data include variables about health (mental health, cognitive ability, risk behaviours, use of medical facilities), social networks (assistance within the family, transfer of goods and money, social relationships, volunteering), economic and financial status (employment, job characteristics, job opportunities after retirement age, sources and composition of the income, wealth and consumption, real estate, education, financial asset holdings).

The population of households is defined as "the set of families with at least one person who was born before 1957, who speaks the official language of the country and who does not live during the survey period in another country or an institution like a prison".

The main questionnaire consists of 20 modules, collecting information on almost all aspects of the life of European households. Since some applications relate to the individual and other the whole family, we can distinguish respondents as answering to him/herself, to the financial part and responding about housing and responding about the family.

Table 2.1 shows the sample composition by country.

Country	Total		M	F	Age under 50	Age between 50 and 64	Age between 65 and 74	Age 75+
Austria	845	4%	4%	4%	2%	3%	5%	4%
Belgium	2.096	9%	9%	9%	10%	9%	8%	10%
Czech Republic	1.906	1%	8%	9%	9%	9%	7%	7%
Denmark	1.704	8%	8%	1%	9%	8%	7%	7%
France	2.008	9%	8%	9%	13%	9%	8%	9%
Germany	1.642	7%	8%	7%	5%	7%	9%	7%
Greece	1.557	9%	9%	10%	18%	9%	9%	9%
Italy	1.894	9%	9%	9%	6%	8%	10%	9%
Netherlands	1.773	8%	8%	8%	5%	9%	7%	7%
Poland	1.739	7%	7%	7%	5%	8%	7%	7%
Spain	1.360	6%	7%	6%	5%	6%	7%	8%
Sweden	1.887	8%	8%	8%	4%	7%	9%	9%
Switzerland	1.024	4%	4%	4%	4%	4%	4%	4%
Total	21.435	100%	100%	100%	100%	100%	100%	100%

Table 2.1 Breakdown of wave 2 2006/2007 samples (release 2.5.0.) by country, sex and age.

In table 2.2 we report the set of variables used in the analysis.

Section	Type of respondent	Questions
Demographics	All respondents	Gender Marital status Year of birth Years of education General health
Children	Respondent for family	Number of children
Income	Respondent for house	Total income received by all households members in last month. Total income of other households members.
Consumption	Respondent for house	Is household able to make ends meet?
Job	All respondents	Current job situation. Received public benefits.
Financial asset	Financial respondent	Reason for not having a bank account. Amount in bank account, bonds, stocks, mutual funds, individual retirement account, contractual savings for housing, life policies. Own firm company business. Owe money and amount. Risk preferences.

Table 2.2 Sections of questionnaires.

3. A first analysis

3.1 Ownership and investment in financial activities

Table 3.1 shows the portfolio configuration in our sample. It regards "total wealth". We use it because of the quantitative importance of assets such as real estate and value of own share of businesses. Besides these components of wealth represent important factors affecting the composition of asset holdings. Our total wealth is formed by real estate and financial wealth with the addition of all debts.

The first one includes the 77% of the total worth, financial investments includes the 21% and the 2% of the total is due to debts. Households invest a lot in their main residence: the 50% of portfolio value is dedicated to it. Still considering the real estate, 8% consists in own share of business, 6% is given by mortgage on main residence and the 2% is car value. Regarding financial properties, the most of the total worth in portfolio is given by cash amount in bank accounts and individual retirement accounts (5% each). Another large part is given by life insurances (4%). The remainder includes stocks (3%), mutual funds (2%) and bonds (1%). Debts are usually due to friends or relatives (48% of the total debts), loans (24%), student loans (9%), overdue credit cards (6%), debts for cars or other vehicles (4%), overdue bills (2%) and other kind of debts (6%).

Real estate cover the largest slice of portfolio worth. They cover the 77% of the total of which the 65% is due to the main residence. Investing in own main house is the most widespread

real investment and it forms the great part of their future wealth, as we can observe from the value of loans (24%).

Portfolio	
	Percentage on total wealth
Household main residence	50%
Household other real estate	11%
Household value of own share of businesses	8%
Household cars	2%
Household mortgage on main residence	6%
Estates	**77%**
Bank accounts	5%
Bonds	1%
Stocks	3%
Mutual funds	2%
Individual retirement accounts	5%
Life insurances	4%
Countractual savings for housing	0,4%
Financial assets	21%
All debts	**2%**

Table 3.1 Total wealth composition.

Ownership of financial assets differs a lot from country to country. Just the holding of a bank accounts is pretty homogenous across them. Almost every household has a bank account where to save own cash. Lower values of asset holding are concentrated on South and East European countries, like Greece, Spain, Italy, Czech Republic and Poland. It is immediately visible as observations belonging to these countries have, on average, less financial assets than the North and West European countries (plus Germany in Central Europe). In general,

within the EU, in the Member States such as Czech Republic and Poland, the stock of financial assets are smaller than in the other countries.

	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
Austria	0,86%	0,79%	0,83%	0,83%	0,82%	4,14%
Belgium	2,07%	2,04%	2,06%	2,05%	2,05%	10,27%
Czech Republic	2,98%	0,75%	1,87%	1,87%	1,87%	9,34%
Denmark	1,67%	1,67%	1,67%	1,67%	1,67%	8,35%
France	1,97%	2,04%	1,89%	1,97%	1,96%	9,84%
Germany	1,37%	1,32%	1,59%	1,61%	1,61%	7,49%
Greece	0,00%	0,00%	0,11%	1,66%	1,39%	3,16%
Italy	2,29%	1,44%	2,03%	1,67%	1,85%	9,28%
Netherlands	1,74%	1,74%	1,74%	1,73%	1,73%	8,69%
Poland	0,00%	4,69%	0,42%	1,83%	1,58%	8,52%
Spain	1,44%	1,22%	1,36%	1,32%	1,33%	6,66%
Sweden	1,85%	1,91%	1,80%	1,84%	1,84%	9,25%
Switzerland	1,02%	0,98%	1,00%	1,00%	1,00%	5,02%

Table 3.2 Distribution of European households, by total investment amount percentile in financial activities.

In table 3.2 we have the percentage of investors grouped by country and total financial investment distribution. The 20° percentiles are calculated for each country to keep into account the differences among countries in terms of purchasing powers. We defined an investor as a person who has at least one financial asset in his/her portfolio (without take into account bank accounts). As denominator when calculating the percentage of portfolio composition, we employ a measure of "total financial wealth". Greece and Austria have the smallest proportion of total European investors, but it could be due by the small number of respondents in these countries. It seems that distribution across countries and financial investment percentiles is quite regular. This means that, conditionally to the different purchasing power, financial wealth is uniformly distributed across countries. Highest number of investors is in Belgium and France, followed by Czech Republic, Italy, Sweden, Netherlands, Poland, Denmark, Germany, Spain and Switzerland.

In table 3.3, the percentage of European citizens having the financial assets on the entire number of households of the same country is represented. Austria, Czech Republic and France tend to have a high percentage of own citizens having contractual savings for housing.

Denmark has a majority of households having stocks, bonds and retirement accounts and Sweden has a 32% of households holding mutual funds and the 24% prefer to invest in stocks. In Germany, households prefer to invest in bonds and contractual savings for housing, while in Italy we have 13% of households choosing bonds. Netherlands has investments particularly in life insurances, stocks and mutual funds. Belgium has a percentage closed to 10% for every asset. Poland and Greece have usually low number of investors, but the 12% of Polish people tend to invest in life insurances. Switzerland has values around 6-8% with a 13% of investors in bonds.

	Bank accounts	Bonds	C. savings housing	Retirement accounts	Life insurances	Mutual funds	Stocks
Austria	4%	2%	19%	2%	4%	3%	2%
Belgium	12%	13%	8%	13%	10%	12%	13%
Czechia	7%	1%	27%	16%	6%	2%	2%
Denmark	9%	17%	1%	15%	10%	10%	21%
France	11%	3%	24%	15%	8%	10%	8%
Germany	9%	16%	19%	5%	10%	9%	7%
Greece	4%	0%	0%	0%	1%	0%	1%
Italy	9%	13%	0%	1%	3%	3%	3%
Netherlands	10%	4%	0%	3%	11%	8%	9%
Poland	3%	1%	1%	1%	12%	0%	1%
Spain	6%	2%	0%	4%	3%	2%	2%
Sweden	10%	16%	1%	20%	17%	32%	24%
Switzerland	6%	13%	1%	6%	5%	8%	8%
Total	100%	100%	100%	100%	100%	100%	100%

Table 3.3 Investors, by country and financial asset on the total of own citizens

Table 3.4 depicts percentage of investors on number of citizens of the same country. Just the 26% of Polish and the 45% of Greek individuals has a bank account and for Czech and Polish households this percentage is about 65%. From table 3.5 we can note that this 65% of Czech people has, on average, a quite high value of bank accounts, if compared with other

countries in central Europe. Therefore, there is a suspicion that a small portion of Czech people represents a big part of total amount.

	Bank accounts	Bonds	C. savings housing	Retirement accounts	Life insurances	Mutual funds	Stocks
Austria	90%	4%	49%	9%	23%	8%	8%
Belgium	97%	12%	8%	27%	24%	16%	22%
Czechia	65%	1%	31%	36%	16%	3%	4%
Denmark	96%	19%	1%	38%	30%	16%	43%
France	97%	3%	26%	32%	19%	13%	15%
Germany	95%	19%	25%	14%	32%	15%	15%
Greece	43%	0%	0%	0%	2%	0%	2%
Italy	80%	14%	0%	2%	9%	5%	6%
Netherlands	96%	4%	0%	9%	31%	12%	17%
Poland	26%	1%	1%	2%	35%	1%	1%
Spain	83%	2%	1%	12%	10%	4%	5%
Sweden	95%	17%	1%	48%	44%	45%	44%
Switzerland	95%	25%	1%	28%	23%	22%	28%
Total	81%	9%	10%	21%	23%	13%	16%

Table 3.4 Average values of investment, by country and financial asset

Denmark, Sweden and Switzerland have very wide slice of citizens holding different kinds of financial assets. Almost half of Swedish people hold mutual funds, individual retirement accounts, whole life insurances and stocks; the percentage decreases for bonds (17%) and contractual savings for housing (1%). Particular attention has to be given to contractual savings for housing. European countries are characterized by the heterogeneity of housing situations and national of policies. The focus for financial asset is based on certain characteristics of social housing because they can lead to different interpretations from country to country:

- their specific role can be generally expressed as answering to the housing needs of families in terms of access and permanence in decent housing at affordable prices;

- the definition of criteria for allocation and target groups;
- average rents differ from each other, reflecting the economic conditions in different countries;
- sector size, measured as a percentage of the stock of housing rented out of the total housing stock.

  It can be seen that the countries of Northern and Western Europe tend to have an attention for social sector larger than the Mediterranean countries.

  Denmark and Sweden have really low percentages of citizens who benefit these social housing. In these countries, social housing model defines the good of living a public responsibility towards total population. These low values could be explained because of the high average rent for social housing, reflecting economic situations of the countries. The same model is given for Netherlands. Austria has a great percentage of the social rented housing on total number of new constructions. Based on data available in 2005, this percentage is about 30% in Austria. Instead, in our sample, households living in a social house is about 50% of the total. In this way, we note as Austrian social policies are addressed in particular to older people. Same available data for Denmark show an interesting difference: they tell that the 20,7% of total number of constructions are social houses, but our sample underlines that just the 0.82% of population with more than 50 years old benefits. This may be explained again with the social model of the country: it does not focus on a narrower category of beneficiaries, typically the poorest families - elderly, disabled, unemployed, single parents, but the assignment is done through waiting lists or not priority criteria or a number of vacant apartments may be reserved for those families who have an urgent housing needs.
  Czech Republic and Poland are exceptions in the landscape of Eastern Europe. They were not reached by the mass privatization of 90s, so the public sector of social housing is still present, especially in the Czech Republic.

  Investments in bonds vary greatly across European countries to another. The highest figure is Switzerland, with the 25% of its households participating in bonds market. In Germany, Italy

and Belgium, individuals prefer to invest directly in bonds, while in other European countries such investments are mainly through other kind of assets. Italian households participating in bonds market are 14%. Denmark, Sweden and Germany people are around 19%. In Belgium, the percentage for the same financial asset is 12% and other countries is usually less than 5%.

It is possible to distinguish two different ways to ensure the livelihood of retirees in Europe. The first one is based on creating a type of insurance that provides services related proportionally to income received during employment and financed with contributions made mandatory by the State. The second instead contemplates the adoption of a universal type providing a minimum pension equal for all, funded largely through general taxation. In the latter system, in order to maintain the same level of income had during working life, it is necessary to ensure individual supplementary retirement schemes. The universalistic system is adopted by Sweden and Denmark. In Denmark, in particular, there is a pure universalistic system, under which all citizens, regardless of the pursuit of employment, are entitled to receive a minimum pension (flat-rate) upon completion of a certain age. On the other hand, a pure employment-type system is adopted by France, Belgium, Germany and Austria. A mixed employment-type model is adopted by Italy, Netherlands and Switzerland (Soede et al.).
Recovering a cluster analysis from the Social Cultural Planning office about positioning of European countries on the basis of some welfare characteristics, we find out Denmark and Sweden in the same cluster where the general social security scores really high and conversely, the pension system, scores very low. The variable representing pension system covers the main benefits to supplement their income - including pensions, maternity leave and custody arrangements - taxes and social security contributions, and the rules of the labor market. In another cluster we discover Germany, Belgium, France and Austria: for these countries, social security and pension policy are located at a medium level compared to others. Netherlands has almost the same scores, but its social security level is higher. The third cluster contains Italy and Spain: Italian pension system has high scoring, but the worst degree of social security, like for Spain. Anyway, Spain has the highest standard for pension

system. The last one includes Poland and Czech Republic, where both variables are quite failing.

Supplementary pensions in our sample are really widespread in Denmark and Sweden (38% and 48% respectively) and also the investment is very high. Countries having a pure employment model observe a participation between 10% for Austria and 32% for France of households who address to additional investments to supplement their retirement. Around 10% of households hold an individual retirement account in Spain and Netherlands. Italy, Greece and Poland have lowest percentages, closed to 2. A particular case is Czech Republic, whose households holding retirement accounts are the 36%!

Wealth

Household wealth is defined as the total market value of dwellings, consumer durable goods and financial assets, net of debts (Wolff, 2006).

Generally, there is a positive correlation between financial investments and the others household resources, measured either by income or real estate.

Having higher incomes leads, generally, to bigger investments in finance.

Fin. assets / Incomes	**<20°**	**20°-40°**	**40°-60°**	**60°-80°**	**>80°**	Total
<20°	5,43%	6,32%	2,66%	2,91%	1,89%	20%
>80°	2,47%	2,09%	3,43%	4,88%	7,60%	20%
20°-40°	4,95%	4,96%	3,89%	3,80%	2,31%	20%
40°-60°	3,58%	4,11%	4,32%	4,46%	3,64%	20%
60°-80°	2,84%	3,14%	4,07%	5,00%	5,28%	20%
Total	20%	20%	20%	20%	20%	**100%**

Table 3.5 Distribution of incomes by financial assets percentiles.

Table 3.5 shows the sum of all households incomes 20°-percentiles by financial assets amount 20°-percentiles. It presents real values, taking into account differences among countries in terms of purchasing power. We see that distribution is quite regular, saying that households tend to invest in financial market a portion quite stable conditionally to their incomes. People staying in the first income-classes tend to invest small amounts and people staying in the highest incomes-classes tend to invest more elevated amounts.

Education

"Wealth is not the only household characteristic that may predict its willingness to take financial risk. Income, age, race, education, and self-reported attitudes to risk may also be important" (J. Campbell, 2006).

The literature documents that a higher education involves some returns, such as a higher wage, and accordingly a higher expected return on savings through increased access to the financial market. Our variable represents the number of household years of education, that sometimes is not explanatory of the real level of education, because of the existence of people who employ a longer time than expected to complete a course of study.

We can confirm that participation in financial market is related with the years of education. Households having at least one financial asset, still without considering bank accounts, show, on average, higher levels of education (with a mean of more than ten years). Instead, households not participating in the financial market (having less than 8 years of education).

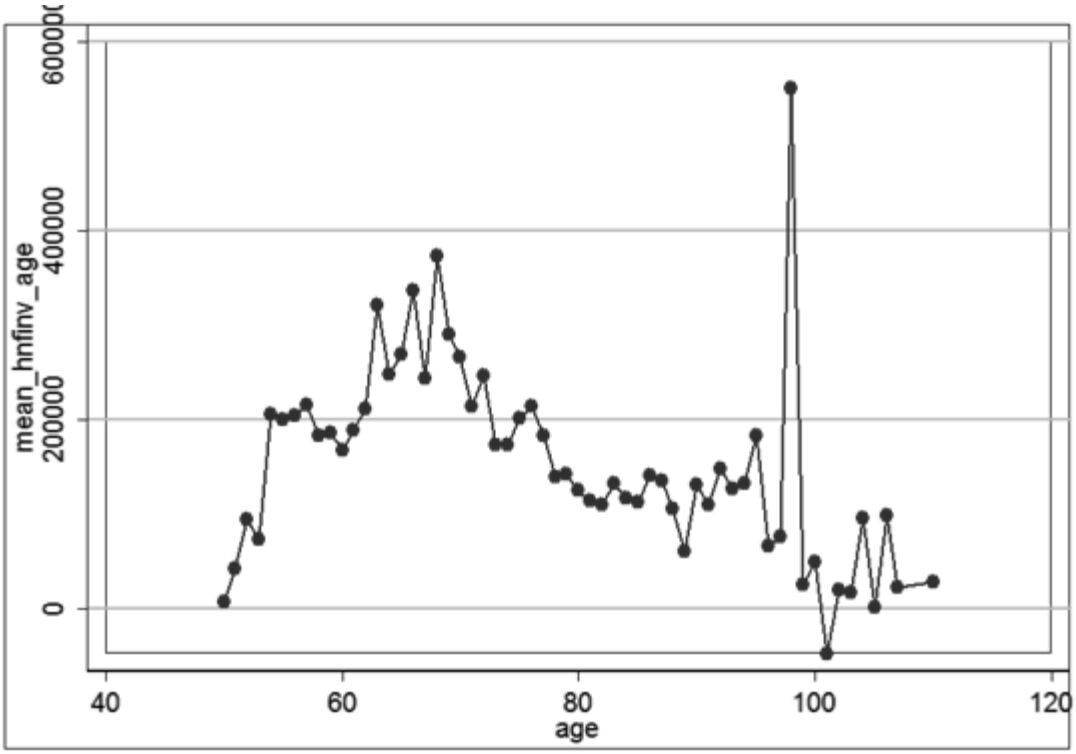
		<20°	20°-40°	40°-60°	60°-80°	>80°	Total
high	Austria	13%	17%	17%	29%	24%	100%
	Belgium	3%	10%	17%	24%	47%	100%
	Czechia	15%	16%	36%	27%	6%	100%
	Denmark	5%	12%	13%	24%	46%	100%
	France	4%	16%	21%	25%	33%	100%
	Germany	6%	12%	17%	27%	38%	100%
	Greece	30%	14%	27%	25%	4%	100%
	Italy	9%	16%	25%	26%	24%	100%
	Netherlands	3%	11%	15%	28%	43%	100%
	Poland	32%	30%	20%	12%	5%	100%
	Spain	9%	15%	19%	25%	32%	100%
	Sweden	3%	5%	16%	32%	44%	100%
	Switzerland	4%	4%	11%	26%	56%	100%
medium	Austria	12%	25%	28%	21%	15%	100%
	Belgium	7%	19%	21%	23%	30%	100%
	Czechia	32%	27%	26%	13%	2%	100%
	Denmark	8%	19%	13%	22%	38%	100%
	France	9%	22%	24%	26%	20%	100%
	Germany	9%	23%	22%	28%	19%	100%
	Greece	50%	12%	19%	15%	4%	100%
	Italy	12%	24%	27%	27%	10%	100%
	Netherlands	6%	18%	20%	29%	28%	100%
	Poland	45%	31%	15%	7%	2%	100%
	Spain	20%	28%	18%	19%	15%	100%
	Sweden	5%	9%	18%	33%	35%	100%
	Switzerland	7%	8%	14%	27%	45%	100%
low	Austria	16%	28%	25%	20%	10%	100%
	Belgium	16%	28%	22%	18%	16%	100%
	Czechia	47%	28%	20%	4%	0%	100%
	Denmark	17%	27%	19%	22%	17%	100%
	France	17%	36%	22%	17%	8%	100%
	Germany	24%	30%	21%	16%	9%	100%
	Greece	66%	12%	16%	5%	1%	100%
	Italy	30%	28%	26%	12%	4%	100%
	Netherlands	9%	32%	27%	19%	14%	100%
	Poland	71%	21%	6%	2%	0%	100%
	Spain	29%	31%	21%	11%	8%	100%
	Sweden	9%	16%	25%	28%	22%	100%
	Switzerland	5%	11%	14%	27%	42%	100%
	Total	20%	20%	20%	20%	20%	100%

Table 3.6 Total investments 20°-percentiles, by educational levels

Education influence is confirmed again from table 3.6. Percentages in the highest percentile (>80°) increase everywhere with the increasing of education level. In general, in the first classes the opposite happens: number of households having low investments tend to have medium or low education.

Age

Graph 3.1 explores the age-participation relation. The profile has a hump shape. The investment tends to grow up till 65 years old, and then it starts decreasing. In correspondence of 97 years old, we find out a peak into the investments, probably due to an outlier. Period between 60 and 80 has the highest value in investment.



Graph 3.1 Average investments amount by age.

% investors/tot Europeans	Percentile 20° (<=61)	Percentile 40° (61,66]	Percentile 60° (66,73]	Percentile 80° (73,81]	Tot (>81)	Total
Austria	1,0%	1,0%	1,0%	1,0%	2,0%	6,0%
Belgium	2,0%	2,0%	2,0%	1,0%	2,0%	9,0%
Czech Republic	2,0%	2,0%	2,0%	2,0%	1,0%	9,0%
Denmark	3,0%	2,0%	2,0%	2,0%	1,0%	10,0%
France	2,0%	2,0%	2,0%	2,0%	3,0%	11,0%
Germany	2,0%	2,0%	2,0%	1,0%	1,0%	8,0%
Greece	0,5%	1,0%	0,0%	0,0%	1,0%	2,5%
Italy	1,0%	1,0%	1,0%	1,0%	0,0%	4,0%
Netherlands	1,5%	2,0%	2,0%	2,0%	1,0%	8,5%
Poland	1,0%	1,0%	1,0%	2,0%	2,0%	7,0%
Spain	1,0%	1,0%	0,0%	0,0%	0,0%	2,0%
Sweden	2,0%	2,0%	4,0%	4,0%	3,0%	15,0%
Switzerland	1,0%	1,0%	1,0%	2,0%	3,0%	8,0%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 3.7_Proportion of households investing in risky assets, by age percentiles

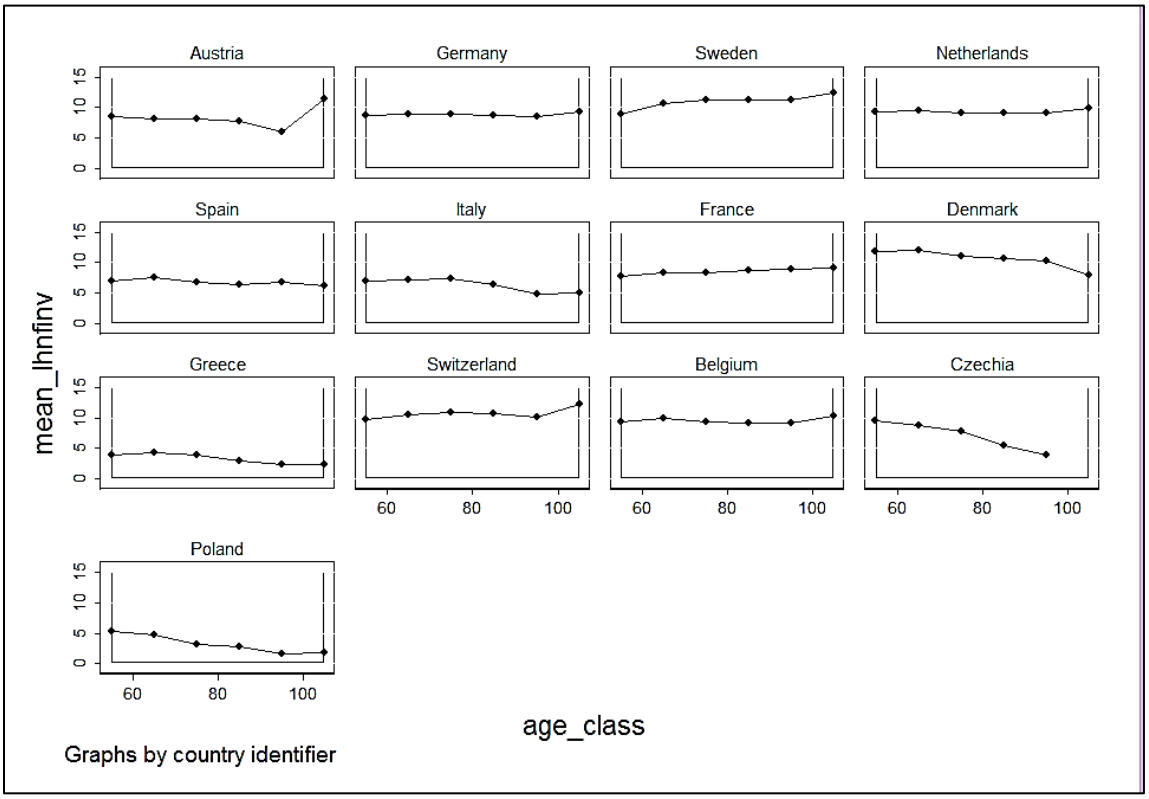
% investors/total citizens	Percentile 20° <=61	Percentile 40° (61,66]	Percentile 60° (66,73]	Percentile 80° (73,81]	Tot >81	Total
Austria	14%	**35%**	28%	20%	3%	100%
Belgium	7%	22%	26%	**28%**	17%	100%
Czech Republic	**32%**	1%	9%	27%	30%	100%
Denmark	10%	4%	8%	14%	**65%**	100%
France	12%	**30%**	26%	25%	7%	100%
Germany	10%	24%	28%	**30%**	8%	100%
Greece	**59%**	18%	18%	5%	0%	100%
Italy	25%	28%	**31%**	13%	3%	100%
Netherlands	7%	22%	28%	**30%**	13%	100%
Poland	**55%**	20%	16%	7%	1%	100%
Spain	22%	**37%**	23%	15%	3%	100%
Sweden	16%	2%	4%	14%	**63%**	100%
Switzerland	10%	5%	14%	30%	**41%**	100%
Total	20%	20%	20%	20%	20%	100%

Table 3.8_Proportion of households investing in risky assets, by age percentiles

Tables 3.7 and 3.8 list the distribution of investors in European countries, by age percentiles. In the first one, we see that the majority of investors, defined as those with at least one financial asset (with the exception of bank account), lives in Sweden and France (15% and

11% respectively), followed by Belgium and the Czech Republic (9%). The countries with the lowest number of investors are Greece and Spain. In Czech Republic, Poland and Greece the majority of investors in risky assets is concentrated in the first age group (less than 61 years old), while in Austria, France and Spain the largest number is between 61 and 66 years. In Italy is between 66 and 73 years, in Belgium, the Netherlands and Germany the population of investors in risky assets is in the age group between 73 and 81 years. For Denmark, Sweden and Switzerland the most important class is that one of over 81 years old.

❖



Graph 3.2 Average net investments amount (logarithm) by age and country

The investment level is rather stable. Greece and Poland present less elevated average investments (around 150 euro). Czech Republic has great investments at the beginning but

decreasing during the period. Austria, Spain, Italy and France has a mean of 1000 invested euro: in correspondence of 95 years old, for Italians there is a flexion leading investment to 150 euro. French people instead see an increase of their investments going between 1000 to more than 2900 euro during their life. An interesting shape is found out for Austria at 90 years old, while Germany, Sweden, Netherlands, Denmark and Belgium have the highest values of total investments.

Risk preferences

The table shows financial ownership by risk attitude. The variable is coded as:

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

The most of the household is concentrated in the middle of the distribution. They are not totally risk averse and they do not take substantial risks. They prefer to invest with an average level of risk. Considering not risk-averse households, they are usually concentrated on high levels of financial participation, in particular for those ones who are not risk averse. Instead, investments considerably go down when households are average or totally risk averse. By looking at Germany, Denmark, Sweden, Netherlands and Belgium we can observe an high portion of investors preferring elevated risks.

	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
Not willing to take any financial risks	18,42%	17,55%	16,14%	13,88%	9,63%	75,61%
Take above average financial risks expecting to earn above average returns	0,21%	0,29%	0,59%	1,30%	2,75%	5,15%
Take average financial risks expecting to earn average returns	1,15%	2,09%	3,14%	4,65%	7,11%	18,15%
Take substantial financial risks expecting to earn substantial returns	0,17%	0,11%	0,13%	0,17%	0,51%	1,09%
Total	20%	20%	20%	20%	20%	100%

Table 3.9_Proportion of households, by asset investment percentiles and risk preferences

As we can see from table 3.9, the great part of the people prefers not to take any financial risk (76%), but only the 17% of them has a negative or equal to zero investment. Just the 1% of the sample usually takes high average financial risks.

Citizens from Spain, Poland and Austria prefer to take no financial risks, while Denmark, Sweden, Switzerland and Belgium see own citizens investing in risky financial assets.

	Not willing to take any financial risks	Take above average financial risks expecting to earn above average returns	Take average financial risks expecting to earn average returns	Take substantial financial risks expecting to earn substantial returns	Total
Austria	82%	2%	16%	0%	100%
Belgium	72%	3%	24%	1%	100%
Czechia	70%	2%	27%	0%	100%
Denmark	56%	10%	31%	2%	100%
France	78%	2%	18%	1%	100%
Germany	73%	3%	24%	1%	100%
Greece	82%	3%	14%	1%	100%
Italy	85%	2%	12%	1%	100%
Netherlands	74%	2%	22%	2%	100%
Poland	91%	0%	8%	1%	100%
Spain	92%	1%	7%	0%	100%
Sweden	64%	29%	5%	2%	100%
Switzerland	67%	2%	29%	2%	100%
Total	76%	5%	18%	1%	100%

Table 3.10 Households by country and risk preferences

Gender and marital status

	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
Female	**23%**	**22%**	**21%**	**18%**	**16%**	**100%**
Not married	24%	23%	21%	18%	15%	100%
Married	21%	21%	21%	20%	18%	100%
Male	**16%**	**17%**	**19%**	**22%**	**25%**	**100%**
Not married	16%	17%	19%	23%	26%	100%
Married	17%	18%	19%	22%	24%	100%
Total	20%	20%	20%	20%	20%	100%

Table 3.11 Households by sex and marital status

Table 3.11 presents investment amounts by sex and marital status. Married heads are more concentrated in classes of higher investments, especially for men who generally have the highest percentage just after the last percentile. The majority of women however is found to have lower capital investment.

4. Analysis

4.1 Why using a multilevel structure?

Grouped data arise in almost all areas of statistical application. Sometimes the grouping structure is simple, where each case belongs to single group and there is only one grouping factor. More complex dataset have a hierarchical or nested structure or include longitudinal or spatial elements. All such data share the common feature of correlation of observation within the same group and so analyses that assume independence of observations will inappropriate. The use of random effects is one common and convenient way to model such grouping structure.

Julian Faraway

When the data structure is hierarchical from cluster sampling, the assumption of independent observations is not tenable, and an analysis dealing with the correlations among observations (i.e., multilevel modeling) is required (Raudenbush & Bryk, 2002). A single-level approach ignoring the data dependency in nested data can lead to biased parameter estimates and degraded standard errors.

Our dataset includes 21435 households grouped in 130 European regions (13 countries), resulting from variable *nuts2* from the original dataset. Since countries and regions vary with respect to measured private consumption, gross investments, government spending and import/export spending. Households are not isolated units in Europe, but they are socially,

culturally and economically influenced by the country where they live. Regardless of the origin of the hierarchy, each group will tend to differentiate themselves from others because:
- the behavior of each individual is influenced by the group of belonging;
- the characteristics and peculiarities of the group are influenced by the individuals who compose it. The hierarchical structure in our case is due to the individual's residence in one of the considered European countries.

One of the merits of the multilevel modelling is its ability to treat some variables, like Gross Domestic Product or Human Development Index (country-level indexes), as both characteristic of units and contextual variables in the same analysis.

Since we are interested in the effect of the macro-level variable z on the micro-level variable y, controlling for the micro-level variables x_p, a multilevel structure is adequate.

We find households nested within countries. If we calculate the Intraclass Correlation, that is a measure of how strongly units in the same group resemble each other defined as the proportion of the total variance explained by the variability between groups, we see how the observations are not independent.

About estimates, one of the most widely used alternatives to OLS is ML (Kleinbaum, 1996) and another one, really closed to ML, is REML. The first one doesn't take into account the number of parameters used in model estimation, so regression coefficients estimated with small samples may be biased. Instead, REML takes into consideration the number of parameters used in model estimation and, in this way, REML estimators for small samples are less biased. In general, as the number of parameters estimated increases, the difference between ML and REML estimates becomes larger.

4.2 Does group membership matter?

Significant relationships can exist among group means even if individuals are randomly assigned to groups (Bliese, 2000; Hammond, 1973). Hence, it is important to understand if group membership matters.

One of the most important aspects in group-level analysis is the necessity to determine how much of a variable total variance is due to the groups. This is so important because the variance due to the group-level properties has theoretical implications about underlying group processes. Demonstrating the existence of group agreement is necessary to justify the using of multilevel modeling. To establish agreement, one merely needs to demonstrate that responses from group-members are more similar each other than would be expected by chance (Klein et al., 1994). The term *within group agreement* refers to the degree to which ratings from individuals are interchangeable; that is, agreement reflects the degree to which raters provide essentially the same rating (Kozlowsky & Hattrup, 1992, Tilsley & Weiss, 1075). ICC1 and ICC2, the two major forms of Intraclass Correlation Coefficients, are calculated from a one-way random effect ANOVA model. Here, the variable of interest is predicted from a group membership factor.

$$Y_{ij} = \mu + U_j + R_{ij}$$

where Y_{ij} is the outcome value observed for micro-unit i within macro-unit j, μ is the population grand mean, U_j is the effect of group j^{th} and R_{ij} is the residual effect for micro-unit i within the group j. Hence, group j has true mean $\mu + U_j$, and each value of the dependent variable that deviates from the group-mean is given by some values of R_{ij}. This one is a random variable and it means that Y_{ij} values differ each other randomly. It is assumed that all variables are independent, U_j has mean 0 and population variance τ^2 (the

population between group variance) and the R_{ij} terms have mean 0 and variance σ^2(the population within group variance). So, total variance of the dependent variable is $\mathrm{Var}(Y_{ij}) = \tau^2 + \sigma^2$. The number of macro-units is N and the number of micro-units within j^{th} group is denoted by n_j.

In general, the Intraclass Correlation Coefficient is the proportion of variance that is accounted for by the group level (Snijders, Bosker, 1999).

The first form of ICC is ICC(1) (Bartko, 1976, James, 1982, McGrow & Wong, 1996). It is defined as the ratio of between group variance and total variance. When group sizes differ each other, it can be used the Blalock's formula (1972):

$$\mathrm{ICC}(1) = \frac{\text{Between Group Mean Squared} - \text{Within Group Mean Squared}}{\text{Between Group Mean Squared} + \{(n_j - 1) * \text{Within Group Mean Squared}\}}$$

The second form of Intraclass Correlation Coefficient is ICC(2). It refers to reliability of the group means.

$$\mathrm{ICC}(2) = \frac{\text{Between Group Mean Squared} - \text{Within Group Mean Squared}}{\text{Between Group Mean Squared}}$$

In the next chapter, we try to establish if the two-level structure of data is necessary when applied into the model and if the within agreement of units may be considered non-random.

4.3 Random Group Resampling

There are some methods to test if membership group matters. One of the most popular is Random Group Resampling (RGR), from Bliese and Halverson (2002). In few words, it compares multilevel results based on actual group membership to multilevel results based on pseudo-group membership. It tests whether results using aggregate-level variables from actual groups differ from results using aggregate-level variables created from randomly formed groups. In regression analyses, it is fairly easy to determine whether or not group membership matters by contrasting within-group and between-group correlations and regression coefficients.

RGR is a versatile resampling procedure similar to the permutation test proposed by Fisher in 1930 and related to the bootstrap and jackknife approaches (Efron & Tibshirani, 1993). The permutation test evaluates whether two samples come from the same population. If two samples belong to the same population, it has to be that the group-mean differences between the two pseudo-groups should equal the group-mean differences between the two real groups. Pseudo-groups are created by sorting the observations from the lowest to the highest and then randomly selecting and assigning n1 observations to the first group and n2 observations to the second one (where n_j equals to the number of observations in the j^{th} group). The comparison between the group-mean differences from the actual groups and the group-mean differences from the pseudo-groups allows to estimate the probability that the observed group means from the actual groups come from a single population.

RGR creates as many pseudo-groups as there are actual groups, to test if agreement in pseudo groups is different from the one in real groups. When group-level results from actual groups significantly differ from pseudo-group results, group effects are identified as being present and group membership is considered important to the interpretation of the data (Bliese and Halverson, 2002).

RGR is a procedure to assess within-group agreement in multilevel data. The aim of the RGR procedure is to determine a null distribution of random response when the data comprise a collection of groups, with raters being nested within groups. The distribution of these pseudo groups' variances is then used to determine the expected random variance (Lüdtke and Robitzsch, 2008). A powerful R function is the rgr.agree. It uses RGR to create pseudo groups and calculate pseudo group variances. The rgr.agree algorithm creates pseudo groups that are identical in size characteristics to the actual groups. Table 4.3 and 4.4 provide the standard deviations, variance estimates and the estimate of the z-value if group are regions (table 4.1) and countries (table 4.2).

Summary Statistics for Random and Real Groups (if groups = regions)				
# Random Groups	Avg Random Group Variance	SD Random Group Variance	Avg Real Group Variance	Z-value
960	48.8	14.3	39.0	-8.6

Table 4.1 Summary of RGR function, with groups are regions.

Summary Statistics for Random and Real Groups (if groups = countries)				
# Random Groups	Avg Random Group Variance	SD Random Group Variance	Avg Real Group Variance	Z-value
1001	48.3	2.3	42.5	-9.2

Table 4.2 Summary of RGR function, with groups are countries.

From the first section of the outcome, we can see statistics that allow us to contrast within-group variances from actual group with within-group variances from pseudo groups. The z-value for both the outputs tells us that the within-group variances from actual groups were significantly smaller than the within-group variances from the random groups. For regions, the average within-group variance for random groups is 48.8, with standard deviations of 14.3. For countries the average within-group variance for random groups is similar to that obtained for regions (48.3), but standard deviations are much smaller (2.3). The average within-group variance in actual groups are smaller, 39 for regions and 42.5 for countries.

By using the *quantile* statistics on pseudo groups variance, we can understand how many groups display agreement or not. For regions, the test establishes that variance has to be within the confidence interval of (39.8, 61.6). It must be smaller than 39.8, that it is in the lowest 10% of all pseudo groups variances. In this way, we can say that groups with variance lower than 35.8 have, according to this criteria, a strong agreement, because they are different from pseudo groups. One hundred and five of the total number of groups (131) meet this criteria.

In the same way, we can estimate the variance in order that it should considered 'large' compared to pseudo groups.

A variance equal or larger than 61.6 is too big compared to pseudo groups standards. By using this criteria, we find that 20 regions with an apparent lack of agreement. They are:

DK009 Sonderjyllands amt 62.4409404

CZ08 Moravskoslezsko/Moravian-Silesian 62.5104440

DK005 Vestsjaellands amt 63.9667206

CZ01 Praha/Prague 64.3836864

DK004 Roskilde amt 66.4807670

CZ04 Severozapad/Nothwest 64717231

ITF5 Basilicata 68.4032141

CH07 Ticino 70.0466240

DK00B Vejle amt 72.2765977

SE02 Oestra Mellansverige 76.1237312

DK00F Nordjyllands amt 80.4047472

SE09 Smaland med oearna 86.8310331

CH01 Iemanique 88.7259496

SE0A Vaestsverige 89.3599440

SE04 Sydsverige 92.0256950

SE06 Norra Mellansverige 102.1601403

DK006 Storstroms amt 110.1516969

SE07 Mellersta Norrland 111.1899427

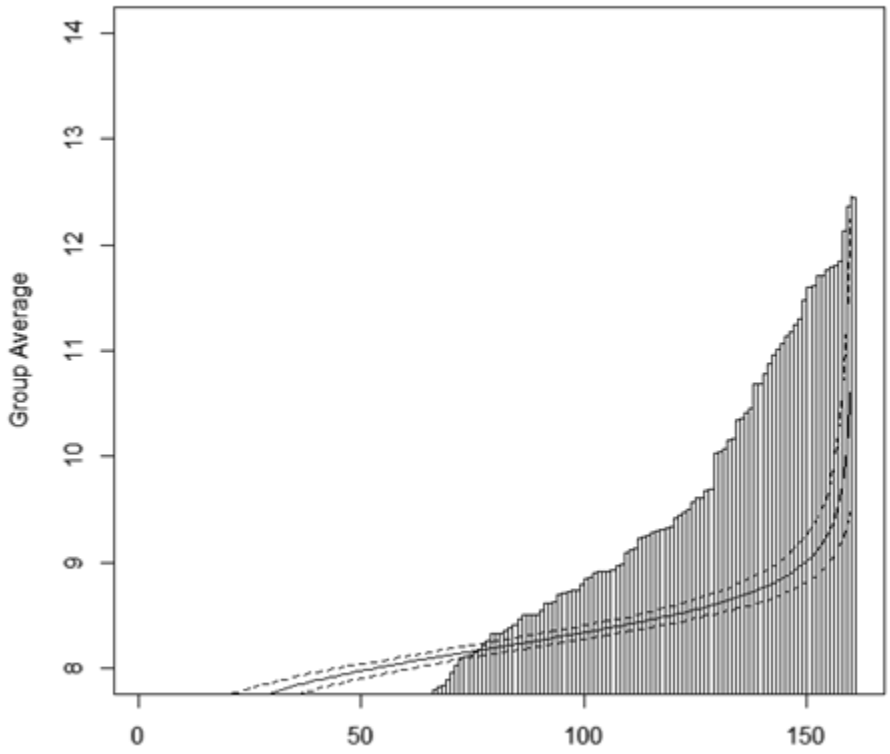
ES63 Ciudad Autónoma de Ceuta135.2090227

SE08 Oevre Norrland 137.4701411

The first expression indicates the code used in the SHARE dataset to refer a determined region, the second reports the name of the same and into parenthesis we can find the corresponding within-group variance.

Three are regions staying in Czech Republic. There are seven regions in Sweden, two in Switzerland, six in Denmark, one in Spain and one in Italy. This kind of analysis suggests that using a two-level structure for our data is a correct way to proceed. Similar results are obtained for countries.

Anyway, we must consider that a high value of ICC(1) can be due to one or two highly anomalous groups rather than indicating generally shared group properties among the entire sample (Bliese, 2012). To see if this happens in our sample, we use an approach which compares observed group means with group means resulting from a randomly assigning of individuals to pseudo groups.



Graph 4.1

The bar chart in graph 4.1 represents each groups' average rating of logarithm of net amount of financial assets value sorted from highest to lowest. The line represents a random distribution of random groups, that is the expected distribution if there were no group-level effects associated with these data. The dotted lines represent the upper and lower 95% confidence bootstrap interval estimates.

Because there are no one or two groups which are clearly different from the pseudo-group distribution, we can say that ICC(1) value does not seem to be caused by one or two aberrant groups.

4.4 Within-And-Between Analysis

The Covariance Theorem, discussed by Robinson (1950) and Dansereau et al. (1984), decomposes a raw correlation from a two-level nested design into 6 components. In particular, the covariance theorem allows to split a correlation into two separate parts – the portion of the correlation attributable to within-group processes, and the portion of the correlation attributable to between-group processes. For that purpose, a Within-And-Between-Analysis (WABA) is done. WABA technique, developed by Dansereau et al. (1982), uses weighted unit averages in calculating between-units correlation as well as correlations based on within-unit variation. Data are transformed prior to entry into a correlation program by partitioning each respondent's raw score on each measure into two components. These two components are a weighted between-unit score, which is shared by all members of a second-level unit and a within-unit component, which is that individual's unique, relative position above or below the units average (McNemar, 1955 and Dansereau et al., 1982).

When the sample is formed by N individuals, grouped in J groups, the covariance theorem can be stated as follows:

$$r_{xyN}\sigma_{xN}\sigma_{yN} = r_{xyB}\sigma_{xB}\sigma_{yB} + r_{xyW}\sigma_{xW}\sigma_{yW}$$

where x and y are two variables, W is the within-group effect and B is the between-group effect.

In particular, we have

$$r_{xy,T} = r_{xy,B}\left(\frac{\sigma_{xB}}{\sigma_{xT}}\right)\left(\frac{\sigma_{yB}}{\sigma_{yT}}\right) + r_{xy,W}\left(\frac{\sigma_{xW}}{\sigma_{xT}}\right)\left(\frac{\sigma_{yW}}{\sigma_{yT}}\right) = r_{xy,B}\,\eta_{xB}\eta_{yB} + r_{xy,W}\,\eta_{xW}\eta_{yW}$$

where $\eta_{xB} = \frac{\sigma_{xB}}{\sigma_{xT}} = \frac{\sigma_{xB}}{\sqrt{\sigma_{xB}^2 + \sigma_{xW}^2}}$ and $\eta_{xW} = \frac{\sigma_{xW}}{\sigma_{xT}} = \frac{\sigma_{xW}}{\sqrt{\sigma_{xB}^2 + \sigma_{xW}^2}}$. Similarly for y.

This theorem states that the total covariance between x and y variables, where the units are grouped in J groups, is equal to a sum of within and between group covariance. From this formula, we can derive three kinds of correlation, such as the raw correlation, the between-units correlation and within-units correlation. The first one is the raw unadjusted correlation of N units. The second one is the between-group correlation, based on the J units, where each of them is represented by the average score and the third one is the within-group correlation, based on residual correlation, after the calculation of between group correlation.
The *waba* function in R provides the covariance theorem components for the relationship between the two variables, x and y.
In the analysis below, we are considering two different group-variables: regions and countries. This is because in multilevel model we will use regions as group level, but the underlying idea is that citizens are more differentiate each other by living in different countries rather than in different regions. There are two principal reasons for this choice. The first one is that we are using country-level variables to assess unobservable differences between units. However, by using regions as group variable, we can also catch institutional differences. The second one is that using 13 groups (number of countries considered in the sample) in a multilevel modeling could be not enough for the accuracy of estimates (Maas & Hoox, 2005 and Paccagnella, 2011), while regions included in the sample are 131! We chose x and y as the net amount in financial investments and sum of all households incomes respectively, to underline the strong relation between the two variables in a groups-context.
The tables below include outputs of the analysis.

Covariance Theorem (groups = region)						
$raw\ corr.$	$eta\ Bx$	$eta\ By$	$corr_B$	$eta\ Wx$	$eta\ Wy$	$corr_W$
0.4296537	0.6138491	0.6775216	0.7078402	0.7894234	0.7355028	0.2329666

Table 4.2

Covariance Theorem (groups = countries)						
$raw\ corr.$	$eta\ Bx$	$eta\ By$	$corr_B$	$eta\ Wx$	$eta\ Wy$	$corr_W$
0.4296537	0.5777236	0.6673772	0.7295704	0.8162325	0.7447198	0.244069

Table 4.4

In the first case, for example, the formula of covariance theorem shows that the raw correlation of $0.43 = \mathrm{r_{xyN}}\sigma_{\mathrm{xN}}\sigma_{\mathrm{yN}}$ is due to:

$$(.613 * .68 * +.7) + (.79 * .74 * +.23)$$

In the second case, instead:

$$(.58 * .67 * +.73) + (.82 * .74 * +.24)$$

The first part of the general formula represents the between-group component of the correlation, and second one represents the within-group component of the correlation. We can see that the group-mean correlations of .70 and .73, respectively for regions and countries, are definitely larger than the within-group correlations of .23 and .24.

Since these two correlations are independent, we can contrast them using a particular function. It is needed because of the WABA's simultaneous consideration of two scores. This function performs a transformation of the two correlations and then tests for the differences between these two component correlations. The transformation is:

$$\mathrm{Z} = \frac{z'_b - z'_W}{\sqrt{\frac{1}{J-3} + \frac{1}{N-J-3}}} \sim N(0,1)$$

The z-value for WABA considering regions as group variable is -5.29, so we conclude that the two correlations are significantly different for each other. In other words, the between-group correlation is significantly larger than the within-group correlation. On the other hand, the z-

value for WABA considering countries as group variable is -0.63, that is the correlations are not statistically significant. It is evident that correlation between financial investment and household incomes is statistically significant with respect to region belonging. We will proceed just for regions, because they will define the group-belonging of 1-level units in the analysis of the next chapters.

Another routine performs the covariance theorem decomposition, but builds upon this work by incorporating Random Group Resampling. RGR is used to randomly assign individuals to pseudo groups. This creates sampling distributions of the covariance theorem components, and allows us to compare actual group covariance components to pseudo group covariance components.

	RawCorr	EtaBx	EtaBy	CorrB	EtaWx	EtaWy	CorrW
N. Rep.	1000	1000	1000	1000	1000	1000	1000
Mean	0.234	0.077	0.077	0.231	0.996	0.997	0.234
SD	0	4.7e-03	4.9e-03	0.008	3.67e-04	3.8e-04	5.05e-04

Table 4.5 Groups = regions

The table 4.5 gives the number of random repetitions, the means and the standard deviations from analysis. Raw correlation has a standard deviation of zero because it does not change, compared to the previous table (4.2).

The between-group correlation has a standard deviation of 0.008. It is evident that all of covariance theorem components in the actual groups significantly vary from their counterparts in the pseudo group analysis. This is obvious because most actual group components are not closed to two standard deviations of the pseudo group means. To test for significance, however, we are going two calculate a confidence intervals.

Quantile	EtaBx	EtaBy	CorrB	EtaWx	EtaWy	CorrW
2.5%	0.07	0.07	0.069	0.996	0.996	0.233
97.5%	0.09	0.09	0.39	0.997	0.997	0.235

Table 4.6 Groups = regions

Because all of the covariance theorem values based on the actual groups are outside of the 95% confidence interval estimates (table 4.6), we may consider significantly different the differences in this resampling design. That is all of the actual group results are significantly different than would be expected by chance.
So, we can conclude that there is a strong effect due to belonging to different regions between investment in financial activities and income variables.

In the end, to have another kind of checking about this structure, we can perform also a simple Log-likelihood ratio test. To verify the existence of random effects in the model initially and the presence of effects due to second-level variables then, we apply the Log-likelihood ratio test.
Comparing model without random effects and model with random effects, we find that the log-likelihood results indicate that the model with the random effect for the net amount of financial assets provides a significantly better fit than the model without these random effects. Table 4.7 shows the output of a X^2 with one degree of freedom. We reject the null hypothesis of equality of the two models and we can say that random effects model better fits data.

	Df	AIC	logLik	L.Ratio	P-Value
Model with random effects	17	139944.7	-69955.34		
Model without random effects	16	140669.6	-70318.81	726.9387	0

Table 4.7

Comparing model without second-level variables and model with these variables (GDP and HDI indicators), we find that the results indicate that the model with the presence of GDP and HDI for the net amount of financial assets provides a significantly better fit than the model without them (table 4.8).

	Df	AIC	logLik	L.Ratio	P-Value
Model with second-level variables	15	140052.3	-70011.13		
Model without second-level variables	17	139944.7	-69955.34	111.5782	0

Table 4.8

5. Modeling

5.1 Multilevel Model

A multilevel problem concerns a population with a hierarchical structure (i.e. level-1 units nested in level-2 units). Multilevel models are helpful because, using grouped data, observations in the same group are generally more similar than the observations in different groups (Hox, 1995). Multilevel models provide a more accurate and comprehensive description of the relationships in clustered data than conventional models, by correcting underestimated standard errors, estimating components of variance at several levels. One can think of Multilevel Random Coefficient (MRC) models as ordinary regression models that have additional variance terms for handling non-independence due to group membership (Bliese, 2012). How does group membership produce additional sources of variance in data? Moreover, multilevel models may be used to examine whether the relationship between variables at the lowest level depends on a variable at another level. The first difference from a regression model is the presence of an error term for each level considered in the analysis. Dependent variable has to be a first-level variable because our aim is to investigate the lowest level. The basic idea is that a result has an individual as well as a group effect. In this sense, the individual effect is referred to the lowest level of sample structure. The underlying idea in our analysis is that the amount of financial assets may be influenced by households decisions as well as region ownership characteristics. To do this, we will let the intercept vary between groups. This reflects the tendency for some regions to have high average of dependent variable values, while other groups have low average dependent variable values. In order to capture economic and wellbeing effects we will add some group-level variables into the model. They differ across groups, but are fixed within-groups.

Now we are going to present the structure of the models for multilevel data on two levels.

Let be J ($j = 1, \dots, J$) the number of groups with a different number n_j ($i = 1, \dots, n_j$) of individuals within each group. Y_{ij} is the response variable measured at the end of the hierarchy (detected, then, for each individual), X_{ijp} are the explanatory variables ($p = 1, \dots, P$) measured at the individual level of individuals and W_j is an explanatory variable measured at the group level. Let be ε_{ij} a not observable effect at the first level and u_{0j} a not observable effect at second level.

$$Y_{ij} = \alpha_j + \beta_p X_{ijp} + \varepsilon_{ij}$$

The intercept, α_j, can be broken down into three parts: an overall or average value of the intercept γ_{00}, a group dependent part of the intercept u_{0j} and the final part, $\gamma_{01} W_j$, which includes the characteristics of the jth group:

$$\alpha_j = \gamma_{00} + \gamma_{01} W_j + u_{0j}$$

Groups are regarding as a random sample from a population of groups. In the above expression, $\gamma_{00} + \gamma_{01} W_j$ represents the expected value of the intercept for groups with W_j characteristics. Instead, u_{0j} is the deviation of the jth group from the expected value.

A multilevel model specifies also the error term distribution:

$$\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$$

$$u_{0j} \sim N(0, \sigma_u^2)$$

$$\mathrm{cov}\left(\varepsilon_{ij}, u_{0j}\right) = 0$$

The fixed portion of the model is $\gamma_{00} + \gamma_{01} W_j + \beta_p X_{ijp}$ and the random component is $u_{0j} + \varepsilon_{ij}$. The final model is:

$$Y_{ij} = \gamma_{00} + \gamma_{01} W_j + \beta_p X_{ijp} + u_{0j} + \varepsilon_{ij}$$

This model, called *random intercept model*, states that the dependent variable is a function of a common intercept γ_{00}, and two error terms: the between-group error term, u_{0j}, and the within-group error term, ε_{ij}. The model essentially states that any dependent variable value can be described in terms of an overall mean plus some error associated with group membership and some individual errors.

The Intraclass Correlation Coefficient is defined as:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2}$$

Under the normality distribution assumption of ε_{ij} and u_{0j}, there are two main methods to estimate the regression coefficients: ML (Maximum Likelihood) and REML (Restricted Maximum Likelihood). They basically differ each other for the estimation of the variance components. REML estimates the variance components taking into account the loss of degrees of freedom resulting from the estimation of the regression parameters, while ML does not. Consequently, REML estimators haven't a downward bias for the variance components, as occurred for ML. Software R uses REML procedure as default.

5.2 Estimating of multilevel models for financial ownership

Heterogeneity in portfolio allocations could be explained by keeping into account heterogeneity in preferences and heterogeneity in circumstances, including a wide range of potential explanatory factors like for instance risk preferences, background and demographic factors, information asymmetries and transaction costs. Transaction costs include taxes, fixed and variable costs of trading in securities markets, and also the time or psychic costs of learning about asset markets. Because we don't have information about transaction costs, we will progress our empirical models without keeping into account them. We will consider background factors, such as labour income, incomes from pension (and every kind of income), and household's real estate. As demographic factors we include gender, age, education, job situation.

Financial resources, income and education are correlated: education is positively correlated with income and negatively correlated with age. To account for this correlation and to isolate the contribution of each factor while holding others variables, we report multilevel regressions for the invested amount in financial assets. The regressions control not only for income, real estate, age and education, but also for number of children, risk attitude, current job situation and two country-level variables to keep into account country-effects, representing population well-being and economics. Response variable is the inverse hyperbolic sine of household's net invested amount in financial assets (*net_fin_asset*). It is given by the following formula:

$$\log(y + \sqrt{y^2 + 1})$$

where y is the sum of the amounts of money into bank accounts, investment in bonds, stocks, mutual funds, individual retirement accounts, contractual savings for housing and life insurances. This choice is done because of the presence of units having no financial assets and, differently from the standard logarithmic transformation, the inverse hyperbolic sine function returns 0 if y is 0.

In the table below, different kinds of financial asset are proposed.

y	*Household net financial assets*
z	$\log(y + \sqrt{y^2 + 1})$
hbaccv	*Household bank accounts*
hbondv	*Household government and corporate bonds*
hstocv	*Household stocks*
hmutfv	*Household mutual funds*
hirav	*Household individual retirement accounts*
hcontv	*Household contractual savings for housing*
hlinsv	*Household whole life insurance*
hliabv	*Household debts (no mortgage)*

Table 5.1 *Financial assets regarded in the analysis*

Following the results given by Campbell, we are going to use some explanatory variables that include these concepts. We argue that factors related with the total amount of financial assets could be the income, the real estates, the number of educational years, the age and, how Campbell said, the self-reported attitude to risk. We added also the age squared because we noted that the dependent variable increases till a certain age and then it starts decreasing.

The variable including information about real estate (hrav, household real assets net of any debts on them) is the sum of other variables, that is: homev, horesv, hownbv, hcarv and hmortv. The variable explaining household's income (lhgtincv, household total gross income)

is, in turn, the sum of great number of other variables, that is: hdipv, yindv, annpen1v,…, annpen5v, annpen7v, … , annpen16v, annpultv, annprltv, annreg1v, … , annreg5v and yrentv.

In order to understand future findings, the context of each country should be considered. The impact of countries on response variable is controlled for variables which indicate the country's index Gross Domestic Product (GDP) and the Human Development Index (HDI). This choice is due to the idea that part of the investments could be influenced by the wealth and the wellbeing of countries. We can claim that individual's wealth and welfare in a particular country are also determined by the social wellbeing and economic levels of the same country. In this sense, also individual's investments should be influenced by indicators like country's market situation, average income, life expectancy, education. From Eurostat, we can know that stock of financial assets held by the households as a percentage of pro-capite GDP in 2006 ranged from 62% for Poland to 355% in Switzerland. Within Europe, the Netherlands, Belgium and Italy rank high while, in general, it can be said that in the new Member States (Czech Republic and Poland,), the stocks of financial assets as a percentage of GDP are smaller than in the other countries. So, stock dimension seems to be correlated to GDP, although other variables, such as development of financial markets can be important.

Because of the presence of many variables linked to household`s wealth and their relationship with the financial assets ownership, it was our decision using a Principal Component Analysis. In fact, they are highly correlated and they measure different aspects of the same general concept of the household`s estate value.

Principal components are independent linear combination of the original variables, whose total variance is equal to that of the observed ones.

Principal components analysis is a technique to manage high-dimensional data and, by using the dependencies between the variables, to represent them in a more tractable, lower-dimensional form. Principal component analysis (PCA) can be used to determine a better set of weights for defining a composite variable that summarizes the data without losing too

much of the information in the data (Croon). It is also one of the simplest and most robust ways of doing such dimensionality reduction.

Such analysis is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA determines the optimal values of the weights in such a way that the linear combination Y preserves a maximal amount of the information available in the data matrix. The information given by a single observed variable is often measured by its variance. So, the purpose of a principal component analysis is to determine that linear combination of the observed variables which attains the largest variance (Croon).

Variables	
homev	Household main residence.
horesv	Household other real estate.
hownbv	Household value of own share businesses.
hmortv	Household mortgage on main residence.
hcarv	Household cars.

Table 5.2 Principal components (eigenvectors)

After standardizing the variables, the outputs are the following. There are a lot of criteria to decide how many principal components are to consider. One is that to keep those components which eigenvalues are bigger than 1 (table 5.3). Another one is to look at the proportion of explained variance (table 5.4).

Eigenvalues of correlation matrix				
Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
1.89017	.984859	.83703	.716626	.571315

Table 5.3 Principal components (eigenvalues)

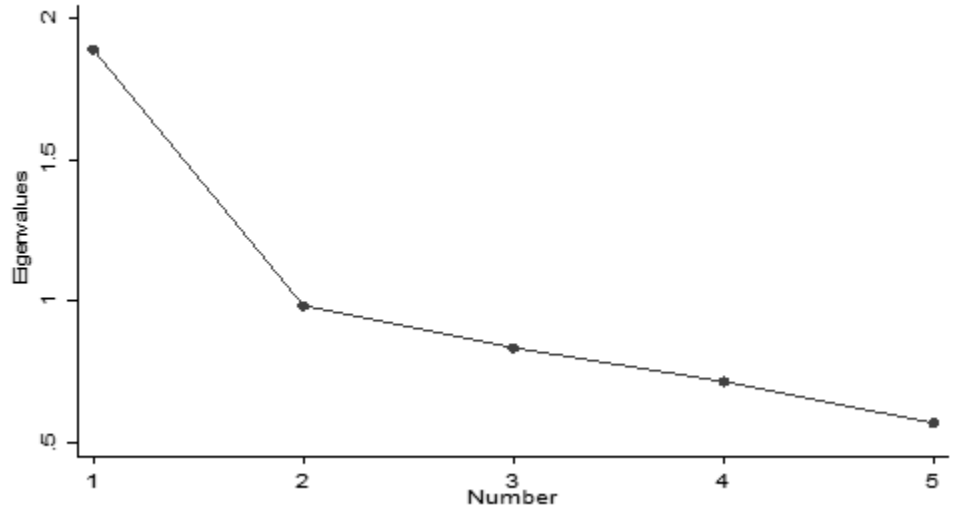
Explained cumulative variance				
Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
0.3780	0.5750	0.7424	0.8857	1

Table 5.4 Principal components (explained variance)

Variables	Comp.1	Comp. 2	Comp. 3	Comp. 4	Comp. 5	Unexplained
homev	**0.5165**	-0.4066	0.0355	0.2373	-0.7144	0
horesv	0.3838	**0.5824**	-0.4217	0.5681	0.1138	0
hownbv	0.3509	**0.5058**	0.7662	-0.1759	-0.0545	0
hmortv	**0.4945**	-0.4795	0.1731	0.1482	0.6882	0
hcarv	**0.4673**	0.0986	-0.4515	-0.7537	0.0089	0

Table 5.5 Principal components (eigenvectors)

The following picture shows the screeplot of the PCA analysis. Screeplot is a plot, in descending order of magnitude, of the eigenvalues of a correlation matrix. It helps to visualize the relative importance of the factors — a sharp drop in the plot signals that subsequent factors are ignorable.



Graph 5.1 Scree plot of eigenvalues

We decide to choose two principal components, that are able to explain about the 58% of the total variable. First variable could be interpreted as something referring to main house and cars while the second one represents the other real estate and households share of businesses.

Meanings of all explanatory variables are explained in the table 5.6.

gender	*Sex (female/male) of the respondent*
partner	*Marital status of the respondent (if married or with a partner)*
health	*General health of the respondent (1-excellent, 2-very good, 3-good, 4-fair, 5-poor)*
homev	*Household main residence (value)*
horesv	*Household other real estate*
hownbv	*Household value of own share of businesses*
hcarv	*Household cars' value*
hmortv	*Household mortgage on main residence*
end_meet	*Is household able to make ends meet?* *1=with great difficulty* *2=fairly difficult* *3=fairly easily* *4=easily*
riskpref	*Risk aversion* *1=I take high risks expecting high returns* *2=I take above average financial risks expecting above average returns* *3=I take average financial risks expecting average returns* *4=I don`t take any risk*
edu_years	*Number of educational years*
employed	*Current employment situation:* *0. Other 1. Employed or self-employed (including working for family business)*

GDP	*Gross Domestic Product: private consumption + gross investment + government spending + (exports – imports)*
HDI	*Human Development Index: is a composite statistic of life expectancy, education, and income indices to rank countries about human development*

Table 5.6

5.3 Main results

According to the list of variables explained in the table 5.2, the estimated model is given below.

$$\log\left(y_{ij} + \sqrt{y_{ij}^2 + 1}\right) = \alpha_j + \beta_1 \text{age}_{ij} + \beta_2 \text{age2}_{ij} + \beta_3 \text{pca1}_{ij} + \beta_4 \text{pca2}_{ij} + \beta_5 \text{edu\_years}_{ij} + \beta_6 \text{nchild}_{ij} + \beta_7 \text{lhgtincv}_{ij} + \beta_8 \text{employed}_{ij} + \beta_9 \text{gender}_{ij} + \beta_{10} \text{partner}_{ij} + \beta_{11} \text{health}_{ij} + \beta_{11} \text{riskpreferences}_{ij} + \varepsilon_{ij}$$

$$\alpha_j = \gamma_{00} + \gamma_{01} \text{GDP}_j + \gamma_{02} \text{HDI}_j + u_{0j}$$

Where y_{ij} is the net investment in financial assets, j = 1, …, 131 identifies region belonging and i = 1, ... , 21435 identifies the households.

Table 5.7 summarizes the main results of the estimation of this model. Standard errors are reported in parenthesis below each estimated coefficient, and coefficients significant at the 10% level or better are indicated with stars. This analysis is performed to measure the relationship of different demographics, economic and social characteristics of the citizens on the choices about financial activities coming from different regions. The ICC in empty model is 29%. It can be concluded that regions do make a difference: 29% (rounding) of the variance in response variable is due to the regions. In Model 1 we first control for individual-level

characteristics. Adding these variables immediately reduces ICC to 24%. And finally, by adding also country-level variables, the ICC will reduce to 17%. The variance component corresponding to the random intercept has decreased to 1.34 (in the empty model was 2.64). Since the ICC can be interpreted as the proportion of the variance not explained by covariates that is due to variation between classes, (in other words, that is the proportion of total variance accounted for by the higher level), it means that covariates inserted in the formula explained part of variance due to differences between groups.

Looking at the control variables, in Model 1, we find that old people tend to have more invested money in financial assets, that is to say, age is a significant variable. The current job situation of the household head does not affect the response. Furthermore, the higher their educational level (expressed in years), the higher is the financial investment. At the same time, having high level of wealth (represented by the variables representing real estate and incomes) has a appositive relationship with the invested amount. Both the variables are significant, the former indicating the other real estate and the latter one indicating the household income. A factor which make go down the financial investment is number of children: it takes a negative sign.

In Model 2, estimates are similar to the ones in Model 1, but countries characteristics are added. They are Gross Domestic Product and Human Development Index (standardized values). The first variable is strongly significant: the higher is the value of this index, the larger is the response. In other words, it means that a person living in a country with a high value in GDP invests more money in financial activities than one living in a country with lower GDP index ceteris paribus. On the other hand, the HDI index is not significant and it doesn't discriminate among countries.

	Empty Model	Model 1	Model 2
Individual-level variables			
Sex	-	0.13 (0.09)	0.12 (0.09)
Health	-	-0.48*** (0.04)	-0.45*** (0.04)
Married	-	0.68*** (0.11)	0.76*** (0.11)
Employed	-	-0.06 (0.12)	-0.03 (0.12)
age	-	0.39*** (0.06)	0.38*** (0.06)
age 2	-	-0.002*** (0.0004)	-0.002*** (0.0004)
Pca1	-	0.04 (0.04)	0.03 (0.04)
Pca2	-	0.34*** (0.04)	0.35*** (0.04)
lhgtincv	-	0.53*** (0.03)	0.51*** (0.03)
riskpref	-	-0.92*** (0.08)	-0.9*** (0.08)
edu_years	-	0.13*** (0.01)	0.13*** (0.01)
nchild	-	-0.32*** (0.03)	-0.32*** (0.03)
Country-level variables			
GDP	-	-	1.88*** (0.35)
HDI	-	-	-0.43 (0.36)
Intercept	7.65*** (0.22)	-8.42*** (2.27)	-7.97*** (2.26)
Intra-class correlation	29%	24%	17.6%
AIC criterion	141184	140049	139941

Table 5.7 Total investment in financial asset

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Estimated coefficients are highly significant, except for dummy indicating the job situation, the first principal component resuming household main residence value and the gender of respondent. The negative coefficient for risk preferences variable indicates that the individual approach to risk negatively influences financial assets holding. The coefficient for the variable counting number of children per households is negative as well. In this way, we can say that a family with a large number of children has an expected amount of financial assets smaller than another one with less children. Health is an ordinal variable and a poor health status is related negatively with the amount in assets. The other variables, the second principal component, number of education, GDP and dummy representing the marital status have instead positive coefficients. So, the increasing one of these characteristics has a positive impact on the dependent variable. The signs of coefficients corresponding to variables age

and *age2* tell us as the value of financial activities holding by households first increases till a certain age and then decreases. Besides, participation level is higher for households with highest education. This can be explained on the returns to education, because a higher education involves a wage premium and accordingly a higher expected return on saving through increased access to the financial market.

A risk-averse person has a smaller financial assets worth than one who takes some risks. In particular, moving from a person who takes some risks to a totally risk-averse person bumps down the expected score by 90%. The same happens if we move from a person who usually takes high risks to another who prefers above average risks.

Having one year of education more, increases the financial assets ownership of 13%, *ceteris paribus*. Having a child more, instead, makes the expected value of response variable lower to 32%, ceteris paribus. It makes sense even that the coefficient corresponding to the second principal component (*pca2*), which includes the information about owned real estate, is positive. The relation between real estate (without considering the main house) and the financial investment is positive.

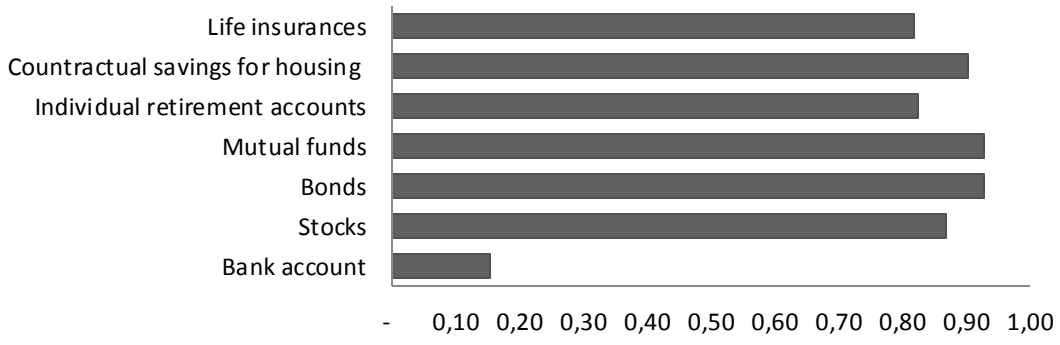
6. Different modeling for each financial asset

The majority of the households hold neither common stock nor other risky financial securities, except for bank accounts.

In the previous chapter, the analysis focused on the total investment in financial activities held by people in the sample. However, as we already observed, the amount in bank accounts plays a decisive role compared to the other assets, because the most of people holds just a bank account in their portfolio and no other kinds of financial activities. That is why we now are going to analyze separately each of them.

Because of the strong presence of units having no assets, we have decided to carry on two different regressions: a multilevel logistic regression first and a standard multilevel regression then. In this way, we can see the relationship between different types of individual and country characteristics and the financial market ownership decision and the invested amount level, across different European countries. For this second analysis, only positive amounts are considered. We can analyze first the factors conditioning the choice to invest in some financial activities or not and then the level of such investments.

As shown in graph 6.1, the proportion of people having no bank accounts is about 15% overall, but for all the other financial assets it grows up at more than 80%! We can deduce that households generally prefer not to bear a financial risk. We can also see how less than the 10% of the sample has bonds, stocks or contractual savings for housing, while the proportion of households holding an individual retirement account is about 18%.



Graph 6.1 Proportion of zeros.

In the implemented logistic models, the dependent variable is binary, taking value of 1 if the unit has a certain asset and 0 otherwise. In the standard multilevel regression, the response variable is the logarithm of the amount invested in that product (all values equal to zero were not accounted). We can compare factors influencing the participation decision and factors influencing the investment after positive decision.

Models will be implemented for each asset: bank account, stock, bond, mutual fund, individual retirement account, contractual saving for housing and whole life insurance.

6.1 The logistic multilevel model

We report estimated from logistic multilevel regressions for the ownership decision and from standard multilevel regressions for the invested amount in financial assets.

We are able to take into account both the two-level structure of data and the dichotomous form of dependent variable to explain financial participation. This variable takes value 1 if the unit has that particular financial asset, otherwise it takes value 0. Similarly to standard multilevel regression, the group structure is defined by the presence of micro observations nested within macro observations. There is a collection of N units grouped in *J* ($j = 1, \dots, J$) groups. Within each group there are n_j observations. The total sample size is then $N = \sum_j n_j$. We specify the same logistic regression to be fitted within each group and the coefficients of the logistic regressions are thought to vary across them. The success probability is denoted by $P_{ij} = \Pr(y_{ij} = 1)$, modelled by means of a *logit* link function, and we assume that y_{ij} has a Bernoulli distribution. The errors ε_{ij} are assumed to follow a standard logistic distribution with mean 0 and variance $\frac{\pi^2}{3}$.

The level-1 model can be written as:

$$\begin{aligned}\text{logit}(P_{ij}) = {} & \alpha_j + \beta_1 \text{age}_{ij} + \beta_2 \text{age2}_{ij} + \beta_3 \text{pca1}_{ij} + \beta_4 \text{pca2}_{ij} + \beta_5 \text{edu\_years}_{ij} \\ & + \beta_6 \text{nchild}_{ij} + \beta_7 \text{lhgtincv}_{ij} + \beta_8 \text{employed}_{ij} + \beta_9 \text{gender}_{ij} \\ & + \beta_{10} \text{partner}_{ij} + \beta_{11} \text{health}_{ij} + \beta_{11} \text{riskpreferences}_{ij} + \varepsilon_{ij}\end{aligned}$$

The level-2 model is given by:

$$\alpha_j = \gamma_{00} + \gamma_{01} \mathrm{GDP}_j + \gamma_{02} \mathrm{HDI}_j + u_{0j}$$

Here, $\mathrm{logit}\left(P_{ij}\right) = \ln\left(\frac{P_{ij}}{1-P_{ij}}\right)$ is called *log-odds*. Logit function is an increasing function defined for numbers between 0 to 1 and its range is from minus infinity to plus infinity.

Therefore, the ICC formulation is now equal to:

$$\mathrm{ICC} = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \frac{\pi^2}{3}}$$

Each table in the next pages includes two different models: one is a logistic multilevel regression, done for the dummy dependent variable *having/not having* that specific financial product and the second one is a traditional multilevel regression, done for the positive invested amount in the same financial asset, introduced through a logarithmic transformation.

The probability of bank accounts for a unit is modeled as a function of respondent gender, marital status (*partner*), current job situation (*employed*), general health (*health*), age and its squared, household's income (*lhgtincv*), a principal component resuming household's main house value (*main house*), another principal component representing the other real estate value (*other real estate*), household preferences about risk in finance (*riskpref*), years of education (*edu_years*) and number of children (*nchild*). Level-2 models take into account the differences between regions. This modeling means that within-region intercepts of each region vary systematically with region belonging and some country characteristics.

The tables summarize demographic and countries effects on the decisions about each financial asset. The third column of the table reports the logistic multilevel results, while the second column shows multilevel regression estimates on the same explanatory variables.

For illustrating the quantitative importance of each effect in the logistic regressions, each table also reports the participation probability for a reference household, and the change in this probability caused by a unit change in each variable, *ceteris paribus*. The reference household is defined as a not-employed male householder aged 65 years, with a partner, with a good health, 10 years of education and one child; he also takes no risks, with average principal components values and the sample average value for logarithm of all incomes and living in Poland (the country whereby GDP and HDI values are the lowest ones). In the last column, the probability estimates are calculated after a unit change in age *ceteris paribus*, a unit change in the number of children *ceteris paribus*, a change in general health level (from good health to poor health), a change in real estate values (from average values to maximum values), an increase of 1 year in education *ceteris paribus* and a variation in the job situation (employed) *ceteris paribus*, 10000 euro more for logarithm of all incomes *ceteris paribus*, a change in risk attitude (he/she takes high risks) *ceteris paribus* and the country-level indexes are increased for the next country with the bigger values of them, the Czech Republic *ceteris paribus*.

Bank accounts

Almost everyone holds a bank account. Table 6.1 summarizes the reasons for not having a bank account in Europe. The most frequent reason is the paucity of money, so we expect to find a strong relationship between variables related to wealth, such as income and real estate, and our dependent variable.

Reason for not having a bank account	*%*	*Cumulative*
do not like dealing with banks	8%	8%
minimum balance/service charges are too	2%	10%
no bank has convenient hours or location	0%	10%
do not need/want a bank account	12%	22%
do not have enough money	60%	82%
savings are managed by children or others	3%	85%
actually I/we do have an account	5%	90%
some other reason	10%	100%

Table 6.1 Reason for not having a bank account

Table 6.2 shows the distribution of the households having or not having a bank account, by financial asset amount percentiles and country. Among all those who have a bank account (over 80%), a very few (5.4%) has a total amount of zero or negative financial investment. Households having bank accounts are distributed uniformly across the percentiles, while familes not having a bank accounts are concentrated in the first class and many of them live in Poland and Greece. Denmark and Sweden hold the highest percentages among those who invest a lot. People who do not have a bank account, generally have no other kinds of financial assets.

Bank account	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**16,0%**	**1,7%**	**0,7%**	**0,3%**	**0,1%**	**18,8%**
Austria	0,3%	0,1%	0,0%	0,0%	0,0%	0,4%
Belgium	0,3%	0,0%	0,0%	0,0%	0,0%	0,3%
Czechia	2,6%	0,3%	0,2%	0,0%	0,0%	3,1%
Denmark	0,3%	0,0%	0,0%	0,0%	0,0%	0,3%
France	0,2%	0,0%	0,0%	0,0%	0,0%	0,3%
Germany	0,3%	0,0%	0,0%	0,0%	0,0%	0,4%
Greece	4,1%	0,0%	0,0%	0,0%	0,0%	4,1%
Italy	1,7%	0,0%	0,0%	0,0%	0,0%	1,8%
Netherlands	0,3%	0,0%	0,0%	0,0%	0,0%	0,3%
Poland	4,3%	1,3%	0,4%	0,1%	0,0%	6,0%
Spain	1,1%	0,0%	0,0%	0,0%	0,0%	1,1%
Sweden	0,3%	0,1%	0,0%	0,0%	0,0%	0,5%
Switzerland	0,2%	0,0%	0,0%	0,0%	0,0%	0,2%
Yes	**4,0%**	**18,3%**	**19,3%**	**19,7%**	**19,9%**	**81,2%**
Austria	0,3%	1,0%	1,0%	0,8%	0,5%	3,5%
Belgium	0,4%	1,8%	2,0%	2,1%	3,2%	9,4%
Czechia	0,1%	2,0%	2,2%	1,2%	0,2%	5,8%
Denmark	0,3%	1,2%	1,1%	1,8%	3,1%	7,6%
France	0,7%	2,3%	2,1%	2,2%	1,9%	9,1%
Germany	0,4%	1,5%	1,5%	2,0%	1,8%	7,3%
Greece	0,0%	0,9%	1,3%	0,7%	0,2%	3,2%
Italy	0,4%	2,3%	2,3%	1,5%	0,6%	7,1%
Netherlands	0,2%	1,6%	1,7%	2,1%	2,3%	8,0%
Poland	0,3%	0,8%	0,5%	0,3%	0,1%	2,1%
Spain	0,5%	1,8%	1,3%	0,9%	0,8%	5,2%
Sweden	0,1%	0,8%	1,7%	2,7%	3,0%	8,4%
Switzerland	0,1%	0,4%	0,6%	1,2%	2,2%	4,5%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.2 Distribution of bank account holding, by country and total financial investment percentiles

Bank accounts	Amount	Participation	Probability estimates
Individual-level variables			
Sex	0.13*** (0.03)	0.09* (0.05)	100%
Health	-0.1*** (0.01)	-0.07 ** (0.02)	100%
Partner	0.12*** (0.03)	0.53*** (0.06)	100%
Employed	0.09* (0.03)	0.07 (0.06)	100%
age	0.1*** (0.02)	0.39 (0.03)	100%
age 2	-0.001*** (0.0001)	-0.0003 (0.0002)	100%
Main house	0.1*** (0.01)	0.23*** (0.05)	100%
Other real estate	0.03* (0.01)	0.05 (0.07)	100%
lhgtincv	0.25*** (0.01)	0.18*** (0.01)	100%
riskpref	-0.24*** (0.02)	-0.5*** (0.05)	100%
edu_years	0.04*** (0.003)	0.11*** (0.01)	100%
nchild	-0.09***(0.09)	-0.04* (0.01)	100%
Country-level variables			
GDP	0.49*** (0.14)	0.74*** (0.09)	100%
HDI	-0.32 . (0.14)	0.57*** (0.09)	
Intercept	3.49*** (0.65)	0.28 (1.13)	
Intra-class correlation	0.32	0.04	
Probability estimate for reference household			100%

Table 6.3 – Model estimates for bank accounts

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 6.3 reports the findings of the multilevel logistic and standard multilevel analysis for the dummy *have/not have a bank account* and the amount held respectively. The ICC in the empty logistic model is 0.64 and so it can be concluded that regions do make a difference: about 64% of the total variance is due to the regions. Then we control for individual-level and group-level characteristics. Adding these variables the ICC immediately reduces to 4%. This means that factors included in our analysis can explain much of the variability between regions in terms of financial participation and invested amount.

In the logistic model we find the respondent's age not significant; however, if we look at the standard multilevel model results, the same variable is strongly significant and it is positively

related to the amount. Furthermore, the higher the educational level, the more people have a bank account and the higher is the amount, *ceteris paribus*. The number of children affect negatively both the probability to have an account and its amount, *ceteris paribus.* The age effect is hump-shaped, reflecting the tendency to invest more until a certain age and less then, as suggested by the life-cycle theory. The strong correlation between wealth and financial investment is showed in the output. As we pointed out before, the main reason to do not have a bank account is to do not have enough money. Our results confirm the strong tendency for the richest and highest-income households to participate in financial markets. Respondent gender is another significant factor in explaining bank accounts holding. If the head of the household is a male, the probability to hold bank account and its amount are higher compared with a female. Besides, if he/she is married or has a partner, the responses are positively affected. The health variable estimate is significant and the sign indicates a positive relationship between good health and amount ownership.

Finally, we look at country-level variables. All of them positively affect the probability to have a bank account while only GDP is statistically significant investigating the amount of money in it. We see that probability estimate for the reference-household is 1. If we modify for a unit-change the explanatory variables the probability result 1 as well.

Stocks

Table 6.4 shows the distribution of stocks holding, by total investment amount percentiles and country.

Stocks	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**19,9%**	**19,6%**	**18,3%**	**15,4%**	**10,4%**	**83,6%**
Austria	0,6%	1,0%	1,0%	0,7%	0,4%	3,6%
Belgium	0,7%	1,7%	1,8%	1,7%	1,6%	7,6%
Czechia	2,8%	2,2%	2,3%	1,1%	0,2%	8,6%
Denmark	0,6%	1,1%	0,7%	0,9%	1,2%	4,6%
France	0,9%	2,3%	1,9%	1,8%	1,1%	8,0%
Germany	0,8%	1,5%	1,5%	1,6%	1,2%	6,5%
Greece	4,1%	0,9%	1,3%	0,7%	0,1%	7,1%
Italy	2,1%	2,3%	2,3%	1,3%	0,4%	8,3%
Netherlands	0,5%	1,6%	1,6%	1,8%	1,4%	6,9%
Poland	4,6%	2,1%	0,9%	0,4%	0,1%	8,0%
Spain	1,6%	1,8%	1,2%	0,8%	0,6%	6,0%
Sweden	0,5%	0,8%	1,3%	1,5%	0,9%	4,9%
Switzerland	0,3%	0,4%	0,6%	1,1%	1,2%	3,5%
Yes	**0,0%**	**0,4%**	**1,7%**	**4,6%**	**9,6%**	**16,4%**
Austria	0,0%	0,0%	0,0%	0,1%	0,1%	0,3%
Belgium	0,0%	0,0%	0,1%	0,4%	1,6%	2,2%
Czechia	0,0%	0,0%	0,1%	0,1%	0,0%	0,3%
Denmark	0,0%	0,2%	0,4%	0,9%	1,9%	3,4%
France	0,0%	0,0%	0,2%	0,4%	0,8%	1,4%
Germany	0,0%	0,0%	0,1%	0,4%	0,7%	1,1%
Greece	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%
Italy	0,0%	0,0%	0,1%	0,2%	0,2%	0,5%
Netherlands	0,0%	0,0%	0,1%	0,4%	0,9%	1,4%
Poland	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%
Spain	0,0%	0,0%	0,0%	0,1%	0,2%	0,3%
Sweden	0,0%	0,1%	0,5%	1,3%	2,1%	3,9%
Switzerland	0,0%	0,0%	0,0%	0,2%	1,1%	1,3%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.4 Distribution of stocks holding, by country and total financial investment percentiles

A small proportion of European households hold some stocks, only the 16%. Among them, this portion, the 10% holds an amount greater than the 80° percentile. Just 4% of households having some stocks invest within the 80° percentile and the remaining households prefer to invest sums under the 60° percentile. Among those who do not own stocks, the majority invest has a zero or negative sum, and the number of households not having stocks decreases with the increasing of investment.

Stocks	Amount	Participation	Probability estimates
Individual-level variables			
Sex	0.03 (0.06)	0.1 * (0.04)	0.014
Health	-0.02 (0.03)	-0.11 *** (0.02)	0.012
Partner	0.03 (0.07)	0.004 (0.05)	0.015
Employed	0.04 (0.08)	0.17 ** (0.06)	0.013
age	0.18 *** (0.04)	0.12 *** (0.03)	0.009
age 2	-0.001*** (0.0003)	-0.0008 *** (0.0002)	
Main house	0.1 *** (0.01)	0.08 *** (0.02)	0.59
Other real estate	0.03 . (0.01)	0.05 (0.07)	0.096
lhgtincv	0.43 *** (0.03)	0.39 *** (0.02)	0.016
riskpref	-0.46 *** (0.04)	-0.83 *** (0.03)	0.15
edu_years	0.01 (0.01)	0.09 *** (0.006)	0.016
nchild	-0.05 . (0.02)	-0.06 *** (0.02)	0.014
Country-level variables			
GDP	0.27 *** (0.07)	0.24** (0.08)	0.039
HDI	-0.1 (0.06)	0.84*** (0.09)	
Intercept	-0.4 (1.62)	-7.46*** (1.23)	
Intra-class correlation	0.13	0.04	
Probability estimate for reference household			1.5%

Table 6.5 - Model estimates for stocks

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 6.5 reports the findings of the multilevel logistic analysis for the binary variable: *to having/not to having some stocks*.

In the logistic multilevel model, once controlling for individual-level and group-level characteristics, the ICC reduces to 4%, while in the multilevel regression it reduces to 13%. The factors included in the analysis are able to explain a lot of the existing between-group variance. The predicted probability of having stocks for a 'typical' household is really low: a male, married, with 65 years old and retired from work, with a good general health, an average wealth, 10 years of education, having a child and average risk preferences and living in Poland, has a probability to hold some stocks is equal to 1.5%. This probability increases with increasing of health status, age, risk preferences, education, wealth. For employed people, it increases as well. Changing the risk attitude in high risk appetite, the probability estimate goes up to 15%. The risk attitude is the strongest predictor for ownership of stocks, maybe because of the high risk contained in holding stocks. The other controlling variables seem to be less important. As before, the household wealth is confirmed to have a strong relationship with the financial investment: real estate and incomes are positively correlated with it, in particular the variable representing the main house and the mortgage. Also in this case, there are some differences between the two model estimates. For example, years of education are significant for the choice between having or not having some stocks but it is no longer significant considering the invested amount. The number of children behaves similarly: it affects negatively the decision of investing or not in these risky products, but it is not significant for the invested sum. Output also demonstrates as age is always important to determine choices about stocks. General health status is another variable related to the participation decision, but it is not significant in explaining invested amount.

Interesting is how living in a "rich" country can positively affects the outcome variables: HDI is not significant to explain the investment amount but it is does for determining the decision *to invest/not to invest*. Well-being characteristics are strongly correlated with the process of participation decisions, while economic features is related with both the responses. The reference country is Poland, because of its lowest levels of GDP and HDI. We continue to observe as *country belonging* do make a strong impact on financial decisions.

Bonds

Table 6.6 shows the distribution of bonds holding, by total investment amount percentiles and country.

Bonds	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**20,0%**	**19,9%**	**18,9%**	**17,3%**	**14,7%**	**90,8%**
Austria	0,6%	1,0%	1,0%	0,8%	0,4%	3,8%
Belgium	0,7%	1,7%	1,9%	1,9%	2,3%	8,6%
Czechia	2,8%	2,2%	2,4%	1,2%	0,2%	8,8%
Denmark	0,6%	1,2%	0,9%	1,4%	2,3%	6,4%
France	0,9%	2,3%	2,1%	2,1%	1,7%	9,1%
Germany	0,8%	1,5%	1,4%	1,5%	1,0%	6,2%
Greece	4,1%	0,9%	1,3%	0,7%	0,2%	7,2%
Italy	2,1%	2,3%	2,0%	0,9%	0,3%	7,6%
Netherlands	0,5%	1,6%	1,7%	2,1%	2,1%	7,9%
Poland	4,6%	2,1%	0,9%	0,4%	0,1%	8,1%
Spain	1,6%	1,8%	1,2%	0,9%	0,7%	6,2%
Sweden	0,5%	0,8%	1,5%	2,3%	2,2%	7,3%
Switzerland	0,3%	0,4%	0,6%	1,1%	1,3%	3,6%
Yes	**0,0%**	**0,2%**	**1,1%**	**2,7%**	**5,3%**	**9,2%**
Austria	0,0%	0,0%	0,0%	0,0%	0,1%	0,1%
Belgium	0,0%	0,0%	0,1%	0,2%	0,8%	1,2%
Czechia	0,0%	0,0%	0,0%	0,1%	0,0%	0,1%
Denmark	0,0%	0,1%	0,2%	0,4%	0,9%	1,5%
France	0,0%	0,0%	0,0%	0,1%	0,2%	0,3%
Germany	0,0%	0,0%	0,1%	0,5%	0,8%	1,5%
Greece	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Italy	0,0%	0,0%	0,3%	0,6%	0,4%	1,2%
Netherlands	0,0%	0,0%	0,0%	0,1%	0,3%	0,3%
Poland	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%
Spain	0,0%	0,0%	0,0%	0,0%	0,1%	0,1%
Sweden	0,0%	0,0%	0,2%	0,5%	0,8%	1,5%
Switzerland	0,0%	0,0%	0,0%	0,2%	1,0%	1,2%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.6 Distribution of bonds holding, by country and total financial investment percentiles

For bonds, the proportion of households participating in financial market is still low. Nearly 91% of households don't hold any bonds. Among them, people tend to invest almost uniformly if we consider the amount of investment. Among households holding some bonds, the 9%, especially in Denmark and Sweden, invest large sums in financial market. Almost nobody has amounts lower than the 40° percentile. Danish and Swedish households who do not hold bonds fall mainly within the last class.

Bonds	Amount	Participation	Probability estimates
Individual-level variables			
Sex	0.08 (0.06)	0.06 (0.05)	2.1%
Health	-0.07 . (0.03)	-0.17 *** (0.02)	1.6%
Partner	0.002 (0.07)	0.14* (0.06)	0,02
Employed	0.13 (0.09)	0.22** (0.07)	1.8%
age	0.08 (0.04)	0.14*** (0.04)	1.1%
age 2	-0.0005 (0.0003)	-0.0008*** (0.0002)	
Main house	0.08 *** (0.02)	-0.04* (0.02)	0.2%
Other real estate	0.04 (0.03)	0.04* (0.02)	10%
lhgtincv	0.3 *** (0.03)	0.22*** (0.02)	2.4%
riskpref	-0.02 (0.04)	-0.38*** (0.03)	6.8%
edu_years	-0.01 (0.01)	0.06*** (0.006)	2.4%
nchild	-0.02 (0.02)	-0.1*** (0.02)	2.3%
Country-level variables			
GDP	0.39 *** (0.088)	0.35** (0.09)	5%
HDI	-0.23 . (0.08)	0.46*** (0.09)	
Intercept	4.56* (1.71)	-8.74*** (1.39)	
Intra-class correlation	0.26	0.04	
Probability estimate for reference household			2.3%

Table 6.7 – Model estimates for bonds

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 6.7 reports results of multilevel logistic analysis for the dichotomous variable *to have/not to have bonds* and multilevel regression for the continue variable representing the invested amount in bonds. Distribution of the participation decision with respect to households age is again hump-shaped, that is the probability to hold some bonds increases with age till a certain threshold and then it starts decreasing, while in multilevel regression it is not significant. In this way, age affects the ownership decision, but it does not how much to invest. The probability for our reference household to hold some bonds is about 2.3%. On the other hand it strongly decreases if we consider a single head of household, a younger person (50 years old rather than 65), or with a large value of own main house. On the other hand, it increases with the increasing value of the other real estates or education. It is worth noting that the estimate of the coefficient of risk preferences is not significant in multilevel standard model, while it is significant in explaining the decision of investing in bonds and, after a unit change in risk attitude (from no risk preferences to high risk ones), probability of holding some bonds increases by 4.5 percentage points. Educational level and number children are able to explain the probability of holding some bonds (increasing with the increasing of education and children), but they are not for investment amount. The same for job situation and marital status: people employed or with a partner are more likely to have bonds, *ceteris paribus*.

The strong correlation of real and financial assets is still confirmed. The relationship between main house worth and bonds investment is positive for investment amount, but it takes a negative sign with respect to ownership decision. Country's variables are significant. Living in a country with high ratings of GDP and HDI affects positively the choice between participating or not in bonds market while only GDP score influence positively the investment amount.

Mutual funds

Table 6.8 shows the distribution of mutual funds holding, by total investment amount percentiles and country.

Mutual funds	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**20,0%**	**19,8%**	**18,7%**	**16,3%**	**12,6%**	**87,4%**
Austria	0,6%	1,0%	1,0%	0,7%	0,3%	3,6%
Belgium	0,7%	1,8%	1,9%	1,9%	1,9%	8,2%
Czechia	2,8%	2,2%	2,3%	1,1%	0,2%	8,6%
Denmark	0,6%	1,2%	1,0%	1,5%	2,3%	6,7%
France	0,9%	2,3%	2,0%	1,8%	1,2%	8,1%
Germany	0,8%	1,5%	1,4%	1,6%	1,3%	6,5%
Greece	4,1%	0,9%	1,3%	0,8%	0,2%	7,2%
Italy	2,1%	2,3%	2,3%	1,3%	0,5%	8,4%
Netherlands	0,5%	1,6%	1,6%	1,9%	1,6%	7,3%
Poland	4,6%	2,1%	0,9%	0,4%	0,1%	8,1%
Spain	1,6%	1,8%	1,2%	0,9%	0,7%	6,1%
Sweden	0,5%	0,8%	1,2%	1,3%	1,0%	4,8%
Switzerland	0,3%	0,4%	0,6%	1,0%	1,4%	3,7%
Yes	**0,0%**	**0,2%**	**1,3%**	**3,7%**	**7,4%**	**12,6%**
Austria	0,0%	0,0%	0,0%	0,1%	0,2%	0,3%
Belgium	0,0%	0,0%	0,1%	0,2%	1,2%	1,6%
Czechia	0,0%	0,0%	0,1%	0,1%	0,0%	0,3%
Denmark	0,0%	0,0%	0,1%	0,3%	0,9%	1,3%
France	0,0%	0,0%	0,1%	0,4%	0,7%	1,2%
Germany	0,0%	0,0%	0,1%	0,4%	0,6%	1,1%
Greece	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Italy	0,0%	0,0%	0,1%	0,2%	0,2%	0,4%
Netherlands	0,0%	0,0%	0,1%	0,2%	0,7%	1,0%
Poland	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%
Spain	0,0%	0,0%	0,0%	0,1%	0,1%	0,2%
Sweden	0,0%	0,1%	0,5%	1,4%	2,0%	4,0%
Switzerland	0,0%	0,0%	0,0%	0,2%	0,8%	1,1%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.8 Distribution of mutual funds holding, by country and total financial investment percentiles

As the most of risky assets, the proportion of household holding mutual funds is low as well. Overall, it is lower than 13%. Among them, more than 7% invest very large amounts in financial market. Instead, among those who do not hold this financial asset, the number of households without mutual funds decreases with the increasing of the total financial investment.

Mutual funds	Amount	Participation	Probability estimates
Individual-level variables			
Sex	-0.16 . (0.05)	0.01 (0.05)	17%
Health	-0.06 . (0.03)	-0.11 *** (0.02)	14%
Partner	-0.02 (0.06)	-0.05 (0.06)	18%
Employed	0.06 (0.07)	0.1 (0.07)	16%
age	0.12 * (0.04)	0.17*** (0.04)	8%
age 2	-0.0006 . (0.0003)	-0.001*** (0.0002)	
Main house	0.08 *** (0.02)	-0.03* (0.01)	3.5%
Other real estate	0.05 * (0.02)	0.03 (0.02)	40%
lhgtincv	0.39 *** (0.03)	0.39*** (0.03)	18%
riskpref	-0.31 *** (0.03)	-0.7*** (0.03)	20%
edu_years	0.01 (0.007)	0.06*** (0.006)	0,18
nchild	-0.12 *** (0.02)	-0.03 . (0.02)	16%
Country-level variables			
GDP	0.02 (0.1)	-0.1 (0.09)	33%
HDI	0.12 (0.09)	1.25*** (0.1)	
Intercept	0.53 (1.53)	-9.65*** (1.33)	
Intra-class correlation	0.31	0.04	
Probability estimate for reference household			17%

Table 6.9 – Model estimates for mutual funds

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Estimates of parameters and standard errors are shown in Table 6.9. Comparing the estimates obtained in both models, we can note some large differences between the same variable estimates in the two models. Basically, ownership and investment amount are explained by different factors. Just for risk preferences and incomes the sign and the significance are the same. Risk aversion has still a negative influence on both the outcomes and the relation between the family incomes and the investment/participation is still positive.

General health level is strongly significant in explaining mutual fund purchasing, while it is not significant in explaining investment amount. Marital status and current job situation are never significant variables. Real estates confirm their positive relation with financial assets investments, but they do not with respect to the decision of investing in mutual funds. Education is able to explain just the participation decision, with the increasing of probability of purchasing mutual funds on the increasing level of education of the household head. Number of children is significant only for the investment amount, in particular, a child more makes decreasing the investment by 12%. HDI is significant in explaining holding and probability of having mutual funds: it is higher in countries where wellness is better. GDP is never significant.

Differently from the previous financial asset, the ICC for the standard multilevel model remains large (around 31%), even after controlling for the set of variables, leaving some not explained between group variance. For a reference household, the probability of having some mutual funds is 17%. If we change reference age (50 years old), this probability decreases till 8%. If we consider the maximum of real estate value found, it grows up till 40%, while increasing main house value it strongly decreases.

Retirement accounts

Table 6.10 shows the distribution of retirement accounts holding, by total investment amount percentiles and country.

Retirement account	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**20,0%**	**18,8%**	**16,6%**	**13,8%**	**10,0%**	**79,2%**
Austria	0,6%	1,0%	0,9%	0,7%	0,3%	3,6%
Belgium	0,7%	1,7%	1,7%	1,4%	1,6%	7,1%
Czechia	2,8%	1,5%	1,1%	0,3%	0,0%	5,7%
Denmark	0,6%	1,2%	0,9%	1,2%	1,0%	4,9%
France	0,9%	2,2%	1,6%	1,1%	0,6%	6,3%
Germany	0,8%	1,5%	1,4%	1,7%	1,2%	6,6%
Greece	4,1%	0,9%	1,3%	0,7%	0,2%	7,2%
Italy	2,1%	2,3%	2,3%	1,4%	0,6%	8,6%
Netherlands	0,5%	1,6%	1,6%	2,0%	1,9%	7,6%
Poland	4,6%	2,1%	0,9%	0,4%	0,1%	7,9%
Spain	1,6%	1,8%	1,1%	0,7%	0,4%	5,6%
Sweden	0,5%	0,8%	1,3%	1,3%	0,8%	4,6%
Switzerland	0,3%	0,4%	0,6%	0,9%	1,3%	3,4%
Yes	**0,0%**	**1,2%**	**3,4%**	**6,2%**	**10,0%**	**20,8%**
Austria	0,0%	0,0%	0,1%	0,1%	0,2%	0,4%
Belgium	0,0%	0,1%	0,3%	0,7%	1,6%	2,6%
Czechia	0,0%	0,7%	1,3%	1,0%	0,2%	3,2%
Denmark	0,0%	0,1%	0,2%	0,7%	2,1%	3,0%
France	0,0%	0,1%	0,6%	1,1%	1,3%	3,0%
Germany	0,0%	0,0%	0,1%	0,3%	0,6%	1,1%
Greece	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Italy	0,0%	0,0%	0,0%	0,1%	0,1%	0,2%
Netherlands	0,0%	0,0%	0,1%	0,2%	0,4%	0,7%
Poland	0,0%	0,0%	0,0%	0,1%	0,0%	0,2%
Spain	0,0%	0,0%	0,1%	0,2%	0,4%	0,8%
Sweden	0,0%	0,1%	0,5%	1,4%	2,2%	4,2%
Switzerland	0,0%	0,0%	0,1%	0,3%	0,9%	1,3%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.10 Distribution of retirement accounts holding, by country and total financial investment percentiles

Just 20% of our households hold a retirement account. The majority of them live in Sweden, Czech Republic and Denmark. Half of households having retirement accounts have also a large investment overall. People preferring to not have a retirement accounts, instead, tend to not invest in finance. Italy is the country where the percentage of households without any retirement accounts is the highest.

Estimates of the models for retirement accounts are presented in Table 6.11.

Retirement accounts	Amount	Participation	Probability estimates
Individual-level variables			
Sex	0.13 ** (0.04)	0.01 (0.04)	3.7%
Health	-0.09 *** (0.02)	-0.07 *** (0.02)	3.2%
Partner	0.18 *** (0.04)	-0.03 (0.05)	38%
Employed	-0.09 (0.06)	-0.28*** (0.06)	4.8%
age	0.11 *** (0.03)	0.08* (0.03)	6%
age 2	-0.0006 . (0.0002)	-0.001*** (0.0002)	
Main house	0.08 *** (0.01)	0.13*** (0.02)	98%
Other real estate	0.04 *** (0.01)	0.06* (0.02)	28%
lhgtincv	0.18 *** (0.02)	0.54*** (0.02)	4%
riskpref	-0.23 *** (0.23)	-0.37*** (0.03)	10%
edu_years	0.03 *** (0.005)	-0.07*** (0.005)	3.5%
nchild	-0.04 * (0.02)	-0.01 (0.02)	
Country-level variables			
GDP	0.7 *** (0.19)	-0.34*** (0.08)	3.6%
HDI	-0.49 . (0.21)	0.51*** (0.08)	
Intercept	4.68*** (1.13)	-6.77*** (1.23)	
Intra-class correlation	0.40	0.04	
Probability estimate for reference household			3.7%

Table 6.11 - Model estimates for retirement accounts

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

therobability of having a retirement account for the reference household as previously defined, is 3.7%. In standard multilevel regression, fixed effects of all level-1 variables are statistically significant, except for head of house current job situation of the head. Risk preference still appears to have a strong effect on the occurrence of this financial asset holding. Going from a risk aversion preference to a high risk preference the probability to decide to invest in retirement account increases by 6.3 percentage point, leading a probability of more than 10%. Respondent gender is significant in explaining investment amount: males tend to invest more than females. In explaining investing decision, job situation is significant, but its sign is apparently strange: a not-employed head is more likely to own this asset compared to an employed person. However, given the features of this financial product, not employed people may just use this investment to complement their public pension in the future.

The country predictors at level-2 reach both the statistical significance only in the multilevel logistic regression, while HDI doesn't affect the continue outcome.

Contractual savings for housing

Table 6.12 shows the distribution of contractual savings for housing holding, by total investment amount percentiles and country.

Cont. Savings housing	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**19,9%**	**18,7%**	**17,2%**	**16,7%**	**17,1%**	**89,7%**
Austria	0,6%	0,6%	0,4%	0,3%	0,1%	2,0%
Belgium	0,7%	1,8%	1,9%	2,0%	2,6%	9,0%
Czechia	2,8%	1,7%	1,2%	0,4%	0,1%	6,1%
Denmark	0,6%	1,2%	1,1%	1,8%	3,1%	7,9%
France	0,9%	2,1%	1,7%	1,3%	1,0%	6,9%
Germany	0,7%	1,4%	1,2%	1,2%	1,2%	5,7%
Greece	4,1%	0,9%	1,3%	0,8%	0,2%	7,3%
Italy	2,1%	2,3%	2,3%	1,5%	0,6%	8,8%
Netherlands	0,5%	1,6%	1,7%	2,2%	2,3%	8,2%
Poland	4,6%	2,1%	0,9%	0,4%	0,1%	8,0%
Spain	1,6%	1,8%	1,2%	0,9%	0,8%	6,3%
Sweden	0,5%	0,9%	1,7%	2,7%	2,9%	8,7%
Switzerland	0,3%	0,4%	0,6%	1,2%	2,2%	4,7%
Yes	**0,0%**	**1,3%**	**2,8%**	**3,3%**	**2,9%**	**10,3%**
Austria	0,0%	0,4%	0,6%	0,5%	0,4%	1,9%
Belgium	0,0%	0,0%	0,1%	0,2%	0,5%	0,8%
Czechia	0,0%	0,5%	1,2%	0,9%	0,2%	2,8%
Denmark	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%
France	0,0%	0,2%	0,5%	0,9%	0,9%	2,5%
Germany	0,0%	0,1%	0,4%	0,8%	0,7%	1,9%
Greece	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Italy	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Netherlands	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Poland	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%
Spain	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Sweden	0,0%	0,0%	0,0%	0,0%	0,1%	0,1%
Switzerland	0,0%	0,0%	0,0%	0,0%	0,0%	0,1%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.12 Distribution of contractual saving for housing holding, by country and total financial investment percentiles

The percentage of people holding contractual savings for housing is lower than the previous assets (10%). Almost nobody in Greece, Italy, the Netherlands and Spain hold this asset. For the others, if the household holds it, its total financial investment is relatively high and falls in the class associated with the largest percentile.

C. sav. for housing	Amount	Participation	Probability estimates
Individual-level variables			
Sex	0.2 *** (0.05)	0.02 (0.05)	79%
Health	-0.06 . (0.03)	-0.004 (0.03)	79%
Partner	0.09 (0.07)	-0.01 (0.06)	79%
Employed	-0.07 (0.08)	0.34*** (0.07)	73%
age	0.03 (0.04)	0.02 (0.04)	85%
age 2	-0.0002 (0.0003)	-0.0004 (0.0003)	
Main house	0.12 *** (0.035)	0.3*** (0.03)	100%
Other real estate	-0.02 (0.05)	0.23*** (0.03)	100%
lhgtincv	0.12 *** (0.03)	0.12*** (0.02)	79%
riskpref	-0.07 (0.05)	-0.14*** (0.04)	85%
edu_years	0.007 (0.008)	-0.06*** (0.01)	78%
nchild	0.003 (0.02)	-0.002 (0.02)	79%
Country-level variables			
GDP	0.42 (0.29)	-0.8*** (0.09)	40%
HDI	-0.68 . (0.33)	-0.9*** (0.09)	
Intercept	7.12*** (1.57)	-2.97* (1.39)	
Intra-class correlation	0.42	0.04	
Probability estimate for reference household			79%

Table 6.13 – Model estimates for contractual savings for housing

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Model estimates in tables 6.13 consider contractual savings for housing.

The sign of the significant coefficients in logistic model are positive, explicating for instance the tendency for employed households to hold contractual savings for housing with a bigger probability. Instead, the coefficient for the same variable in the other model is not significant.

Higher levels of education lead to higher probability to own this asset. A unit-increase in years of school, raises that probability of one percentage point.

Country-level variables are significant just for dichotomous response. Economic and wellbeing scores have negative effect on probability of holding contractual savings for housing. If we consider another country, such as Czech Republic, the probability for a Czech reference household to hold contractual savings for housing decreases till 40%, compared to 79% of a Polish household.

For this kind of financial asset, as well as for the retirement account, the ICC of the final model for the invested amount remains very large (40% or more). After controlling for some level-1 and level-2 variables, a very large part of the variability between-groups is still unexplained. Hence, for these two financial assets, some other and different information is needed to explain the heterogeneity across countries. This is also surprising noting that controlling for the same set of variables, the model for the ownership of these financial assets shows a very low ICC (4%).

Life insurances

Table 6.14 shows the distribution of life insurances holding, by total investment amount percentiles and country.

Life insurances	<20°	20°-40°	40°-60°	60°-80°	>80°	Total
No	**19,8%**	**17,7%**	**16,2%**	**13,0%**	**10,2%**	**76,8%**
Austria	0,6%	1,0%	0,8%	0,5%	0,1%	3,1%
Belgium	0,7%	1,7%	1,6%	1,5%	1,8%	7,4%
Czechia	2,8%	2,2%	1,9%	0,6%	0,0%	7,5%
Denmark	0,6%	1,2%	1,0%	1,4%	1,4%	5,6%
France	0,9%	2,1%	1,8%	1,5%	1,3%	7,6%
Germany	0,8%	1,5%	1,2%	1,1%	0,7%	5,2%
Greece	4,1%	0,9%	1,3%	0,7%	0,1%	7,1%
Italy	2,1%	2,3%	2,1%	1,2%	0,4%	8,1%
Netherlands	0,5%	1,5%	1,3%	1,3%	1,1%	5,7%
Poland	4,5%	0,5%	0,2%	0,1%	0,0%	5,3%
Spain	1,5%	1,8%	1,2%	0,8%	0,4%	5,7%
Sweden	0,5%	0,7%	1,1%	1,3%	1,3%	5,0%
Switzerland	0,3%	0,4%	0,5%	1,0%	1,5%	3,7%
Yes	**0,1%**	**2,4%**	**3,8%**	**7,0%**	**9,8%**	**23,2%**
Austria	0,0%	0,0%	0,2%	0,3%	0,4%	0,9%
Belgium	0,0%	0,0%	0,3%	0,6%	1,4%	2,4%
Czechia	0,0%	0,1%	0,5%	0,6%	0,2%	1,4%
Denmark	0,0%	0,0%	0,1%	0,5%	1,8%	2,4%
France	0,0%	0,2%	0,3%	0,7%	0,6%	1,8%
Germany	0,0%	0,1%	0,4%	0,8%	1,1%	2,4%
Greece	0,0%	0,0%	0,0%	0,1%	0,0%	0,2%
Italy	0,0%	0,0%	0,2%	0,3%	0,2%	0,8%
Netherlands	0,0%	0,1%	0,4%	0,8%	1,3%	2,6%
Poland	0,1%	1,6%	0,7%	0,3%	0,1%	2,9%
Spain	0,0%	0,0%	0,0%	0,2%	0,4%	0,6%
Sweden	0,0%	0,2%	0,6%	1,4%	1,7%	3,8%
Switzerland	0,0%	0,0%	0,1%	0,3%	0,7%	1,1%
Total	20,0%	20,0%	20,0%	20,0%	20,0%	100,0%

Table 6.14 Distribution of life insurances holding, by country and total financial investment percentiles

Almost one fourth of the households purchase life insurances. Sweden is the first country, followed by Poland, The Netherlands, Denmark and Belgium, with a proportion of holding life insurances above 2.4%. Among households having these financial assets, the 9% take an investment above the 80° percentile. The rest prefer not to purchase life insurances. They are uniformly distributed among total asset amount percentiles.

Life insurances	Amount	Participation	Probability estimates
Individual-level variables			
Sex	0.19 *** (0.04)	0.17*** (0.04)	75%
Health	-0.07 *** (0.02)	0.013 (0.017)	79%
Partner	0.07 (0.04)	0.2*** (0.04)	75%
Employed	-0.03 (0.05)	-0.21*** (0.05)	82%
age	-0.07 * (0.03)	-0.01*** (0.03)	89%
age 2	0.0003 (0.0002)	-0.0004 . (0.0002)	
Main house	0.03 *** (0.01)	0.08*** (0.01)	99%
Other real estate	-0.02 (0.01)	0.01 (0.02)	84%
lhgtincv	0.1 *** (0.02)	0.13*** (0.01)	79%
riskpref	-0.07 * (0.03)	-0.28*** (0.03)	89%
edu_years	0.02 *** (0.005)	-0.04*** (0.005)	77%
nchild	0.007 (0.01)	0.07*** (0.01)	79%
Country-level variables			
GDP	1.95 *** (0.27)	0.45*** (0.07)	82%
HDI	-1.6 *** (0.27)	-0.43*** (0.07)	
Intercept	12.59*** (0.99)	3.56*** (1.00)	
Intra-class correlation	0.43	0.04	
Probability estimate for reference household			78%

Table 6.15 – Model estimates for whole life insurances

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The low value of ICC for the multilevel logistic regression confirm the ability of these variables in explaining the probability of holding life insurances. Instead, ICC in standard multilevel regression is relatively high, expressing the persistent existence of not explaining group variance.

Really notable is the result of the high probability to have a life insurance for a reference household: it is about 78%. We can look at the strong positive relationship between life insurance holding and main house value, as well as for the household income. Conversely, the relationship between other real estate value and life insurances holding is not significant. The investment distribution with respect to respondent's age is different from the previous assets estimates, with a convex shape, explaining households tendency of investing less with the increasing age till a certain threshold and then they start investing more with the increasing age. However, this is not surprising thinking to the features of this financial asset, as underlined in table 1.2.

The risk attitude behaves as in the previous asset: each increase in risk preferences scale (from no risk preferences to high risk preferences) corresponds, on average, to 7% increasing in response variable, *ceteris paribus*: risk aversion has a strong negative effect on life insurance investment and on purchasing decision. Education affects positively on investment amount, but negatively on the participation decision. Still discordant from the other assets, life insurance holding is positively influenced by number of children. The probability of deciding to hold a whole life insurance grows with the number of children. This probably a consequence of the fact that, if parents have some children, they want to ensure them against any unforeseen future.

Second-level factors are significant: living in a country with a high score of GDP raises, on average, both the probability and the invested amount.

7. Heckman selection model

In the previous chapter, we compared two different models, the first describing the decision of investing in a financial asset and the second describing the invested amount for such asset. We found that significant relationships in the first model often were not the same for the second model estimation. In this chapter we want to understand it may exist some correlation between the decision of investing and the subsequent invested amount. Heckman selection model consider jointly these two problems, underlying the correlation between the decision and the investment equations.

We are interested in the relation between the financial product ownership and investments and their characteristics. The selection equation models the decision of purchasing financial assets. We can argue that those who purchase financial products may differ each other and, in turn, this implies different invested amounts.

Heckman selection model consists of two equations, one equation describing the relation between an outcome of interest (financial investment) and a vector of covariates, and the second, the decision equation, describing the relation between a binary participation decision h_i and another vector of covariates.

The sample selection occurs when the values of the dependent variable are restricted to a range of values. In this case, the dependent variable is only observed for a subsample. However, there is information about the whole sample.

Regression equation is given by:

$$w_i = \ x'\beta + \ \varepsilon_{1i}$$

Participation decision is specified by the following:

$$h_i^* = z'\gamma + \ \varepsilon_{2i} \ , h_i = 1 \ if \ h_i^* > 0$$

$$h_i = 0 \; if \; h_i^* \leq 0 \; in \; this \; case \; w_i \; is \; not \; observed$$

Given the features of this dataset, x and z variables are the same in our analysis. No exclusion restrictions are needed to identify the model (Heckman, 1978), however it is a good practice to specify at least one exclusion restriction.

The variable $w_i = \log y_i$ is the actual investment for unit i and w_i is not observed for households not participating in financial market. Binary variable h_i indicates if unit i has the asset. In other words, the regression equation determines the value of w_i. The decision equation explains whether an observation is in the sample or not (Verbeek, 2006).

The parametric form of the model assumes that:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim NID \left[0, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right]$$

The contribution of the observation i to the likelihood function is the probability of observing $h_i = 1$ or $h_i = 0$. The second equation describes the level of investment by individuals who decided to invest in that financial product. Then, the contribution of these units to likelihood is given by $f(w_i|h_i = 1)$.

The parameter that makes the model of Heckman different from the analysis of a regression model and a *probit* model separately, is the coefficient of correlation between the error terms of the two equations.

The maximization of the likelihood function produces consistent and asymptotically efficient estimators, having an asymptotic normal distribution.

If the error terms of the equations of decision and regression were uncorrelated, the equation of the financial investment could be consistently estimated using OLS. Instead, if the two error terms are correlated, the OLS estimator will be affected by bias due to sample selection.

7.1 Model estimation

For each country, a dummy variable is created. GDP and HDI are not included in the analysis, to avoid multicollinearity problems due to the presence of the country-dummies.

We choose to use the same factors in order to explain the ownership and the investment amount choice, as well. Participating decision is determined by the same variables affecting the investment amount.

	Bank accounts	Stocks	Bonds	Mutual funds	Retirement accounts	Savings for housing	Life insurance
Decision equation							
Gender	.033***(.024)	.05(.024)	-.035(.02)	.002(.026)	.042(.023)	.022(.030)	.086***(.021)
Health	-.08**(.012)	-.063***(.01)	-.091***(.013)	-.073***(.012)	-.059***(.011)	-.020(.015)	-.016(.010)
Partner	.104(.031)	.002(.030)	.013(.033)	-.005(.03)	.040(.027)	.087(.036)	.080**(.024)
Employed	.021(.031)	.065(.034)	.099 . (.038)	.008(.036)	-.281***(.032)	-.108(.042)	-.138***(.028)
Age	.311(.155)	.773***(.177)	.655**(.194)	.897***(.19)	.169(.176)	.210(.223)	-.902***(.151)
age2	-.021(.010)	-.05***(.011)	-.036*(.012)	-.05***(.012)	-.032*(.012)	-.024(.015)	.041***(.010)
Main house	.230***(.023)	.068***(.009)	-.007(.009)	.005(.008)	.10***(.009)	.035 . (.0)	.056***(.008)
Other real estate	.097*(.036)	.036 . (.012)	.024(.010)	-.003(.009)	.03 . (.013)	.010(.01)	-.005(.009)
Incomes	.097***(.008)	.221***(.016)	.171***(.018)	.192***(.018)	.148***(.013)	.129***(.016)	.097***(.009)
Riskpreferences	-.444***(.036)	-.789***(.025)	-.450***(.029)	-.762***(.02)	-.393***(.025)	-.274***(.033)	-.255***(.023)
edu_years	.399***(.035)	.399***(.033)	.296***(.036)	.298***(.034)	.316***(.031)	.236***(.041)	.168***(.028)
nchild	-.052***(.009)	-.036***(.009)	-.041***(.010)	-.016(.010)	-.011(.00)	.024 . (.011)	.020 . (.00)
Austria	.381***(.068)	.065(.08)	-.901***(.094)	.199(.086)	.84***(.096)	2.74***(.132)	.723***(.066)
Germany	.570***(.06)	.213 . (.067)	-.039(.054)	.376***(.068)	.804***(.082)	1.86***(.129)	.753***(.055)
Sweden	.104(.068)	.579***(.072)	-.543***(.069)	.891***(.075)	1.45***(.085)	-.171(.157)	.824***(.059)
Netherlands	.641***(.066)	.369***(.064)	-.919***(.069)	.243***(.068)	.38***(.084)	-.257(.180)	.676***(.054)
Spain	.247***(.053)	.202(.080)	-.790***(.086)	.093(.087)	1.01***(.08)	.264(.172)	.154 . (.065)
France	.869***(.068)	.32***(.064)	-1.01***(.070)	.389***(.066)	1.62***(.07)	1.96***(.128)	.396***(.055)
Denmark	.041(.073)	.442***(.072)	-.489***(.06)	-.173 . (.078)	.966***(.085)	-.488**(.172)	.225***(.063)
Greece	-1.07***(.048)	-.530***(.100)	-1.58***(.142)	-1.1***(.180)	-.665***(.158)	-.53(.31)	-.654***(.087)
Switzerland	.383***(.077)	.504***(.07)	-.006(.061)	.46***(.073)	1.05***(.085)	.038(.168)	.336***(.064)
Belgium	.761***(.065)	.604***(.062)	-.341***(.055)	.461***(.065)	1.32***(.078)	1.14***(.131)	.533***(.054)
Czechia	-1.12***(.054)	-.989***(.084)	-1.7***(.097)	-.878***(.090)	1.25***(.082)	1.76***(.132)	-.080(.061)
Poland	-1.5***(.051)	-.584***(.104)	-1.3***(.115)	-.712***(.124)	-.10(.106)	.312 . (.154)	.956***(.054)
Intercept	-.635(.583)	-6.21***(.678)	-5.13***(.749)	-6.35***(.729)	-2.79***(.653)	-4.17***(.838)	2.00***(.558)

Table 7.1 Heckman selection model _ Decision equations

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tables 7.1 and 7.2 list the estimates of Heckman selection model for each asset. Table 7.2 shows estimates for regression equation and table 7.1 shows decision equation outputs.

	Bank accounts	*Stocks*	*Bonds*	*Mutual funds*	*Retirement accounts*	*Savings for housing*	*Life insurance*
Regression equation							
Gender	.117***(.026)	.06(.05)	.095(.060)	-.104(.052)	.079(.042)	.182**(.053)	.097 . (.041)
Health	-.072***(.012)	-.038(.030)	-.077(.035)	-.085 . (.027)	-.031(.021)	-.061 . (.028)	-.054*(.02)
Partner	.060(.032)	.000(.068)	-.010(.070)	-.047(.063)	.106(.048)	.060(.068)	-.019(.047)
Employed	.076 . (.035)	.039(.083)	.112(.087)	.063(.074)	.154(.062)	-.040(.078)	.117(.057)
age	.944***(.177)	2.02***(.458)	.910(.458)	1.5***(.414)	.561(.331)	.223(.414)	-.014(.303)
age2	-.051***(.011)	-.117***(.030)	-.057(.029)	-.083*(.02)	-.00(.023)	-.010(.028)	.012(.020)
Main house	.091***(.010)	.111***(.014)	.066 . (.021)	.068***(.016)	.055***(.011)	.113**(.035)	-.004(.012)
Other real estate	.035*(.011)	.036*(.014)	.04(.027)	.056**(.017)	.033*(.012)	.016(.044)	-.011(.013)
Incomes	.166***(.012)	.42***(.059)	.257***(.054)	.338***(.046)	-.024(.0)	.09*(.036)	-.039(.020)
Riskpreferences	-.306***(.030)	-.919***(.162)	-.161(.117)	-.711***(.115)	-.007(.048)	-.144(.072)	.080(.047)
edu_years	.275***(.035)	.294(.109)	-.012(.104)	.245*(.08)	.080(.061)	.023(.084)	.026(.05)
nchild	-.068***(.009)	-.059(.025)	-.022(.025)	-.112***(.021)	-.032(.017)	.012(.022)	-.011(.016)
Austria	-.922***(.075)	-.161(.26)	-.075(.324)	-.96***(.214)	-.465(.239)	-1.28(.67)	-.105(.148)
Germany	-.677***(.062)	-.14(.202)	-.273 . (.11)	-.908***(.173)	-.615*(.206)	-.944(.572)	-.362*(.128)
Sweden	-.778***(.068)	-1.51***(.230)	-2.38***(.19)	-1.36***(.209)	-1.62(.216)	-2.6***(.506)	-1.52***(.133)
Netherlands	.090(.060)	.240(.206)	.248(.273)	-.551**(.171)	-.058(.211)	1.80*(.627)	-.162(.123)
Spain	.018(.066)	.388(.25)	-.052(.305)	.064(.232)	-.486(.219)	-.360(.575)	.47 . (.151)
France	-.630***(.059)	-.283(.203)	-1.03**(.296)	-1.04***(.171)	-1.19***(.212)	-.81(.581)	-1.08***(.125)
Denmark	-1.03***(.070)	-1.49***(.217)	-1.11***(.181)	-1.37***(.18)	-.699**(.209)	-2.2***(.561)	.38*(.133)
Greece	1.18***(.078)	-1.06**(.369)	-.654(.62)	-1.56 . (.690)	.691(.468)	-.184(1.28)	.205(.243)
Switzerland	.737***(.071)	.246(.22)	-.110(.123)	-.252(.184)	-.486(.213)	-1.20 . (.55)	-.105(.138)
Belgium	.05(.059)	.55(.221)	-.030(.141)	.105(.173)	-.96***(.210)	.913(.499)	-.338*(.124)
Czechia	-.903***(.072)	-3.11***(.336)	-2.37***(.493)	-3.13***(.264)	-2.95***(.21)	-1.5*(.563)	-.686***(.135)
Poland	-.31**(.096)	-.973 . (.418)	-1.40**(.501)	-2.24***(.402)	-.785*(.274)	-1.39*(.517)	-2.2***(.131)
Intercept	3.56***(.669)	-3.55(2.19)	3.93(2.14)	.168(1.82)	8.69***(1.25)	7.73***(1.8)	11.8***(1.10)

Table 7.2 Heckman selection model _ Regression equations

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

LR test of indep. equations	95.08	0.27	1.10	4.37	45.21	0.01	49.88
P-value	0	0.6042	0.2934	0.0366	0	0.9398	0

Table 7.3 Heckman selection model _ Likelihood ratio test

First, we look at the hypothesis of incorrelation between the error terms of the two equations (table 7.3). We accept the hypothesis of incorrelation for the following assets: bonds, stocks and contractual savings for housing. In other words, we cannot reject the null hypothesis of independence between participation decision and investment amount for these assets. This could be due to the reduced number of uncensored observations (number of households holding that asset).

For assets having no correlation between the two equations, we find many similarities between Heckman estimates and results from chapter 6. Stocks investments are mainly related with age, real estate values, incomes and risk preferences. The participation decision, instead, are influenced by almost factors included in the analysis, but in the standard regression (where there is no group-effect) gender, job situation and education are not significant. Bond investment is related just with incomes and main house value, while the positive decision of investing in bonds is due to a good health, age (hump-shaped), high risk preferences, education, few children and it relates with high incomes. In multilevel logistic regression having a partner and being employed are also significant in explaining probability of investing and this one is positively related with real estate values and negatively with main house value. Contractual savings for housing show an higher participation for male household heads, with a high value of own main house and incomes. Large investments are affected by high risk attitudes, being employed, high education and they are higher with respect to high worth of main house, real estate and household incomes.

For assets having a correlation between participation decision and investment amount decision, we find out that male heads are more likely to hold a bank account and they tend to have an higher amount, compared to females. Ownership and invested amount are also strongly correlated with households incomes. Incomes relate with mutual funds, retirement accounts and life insurances as well. A good general health status is positively related to bank accounts, mutual funds, retirement accounts decisions and, at the same time, bank accounts and life insurances amounts. Having a partner influences only the participation decision for whole life insurances. Age affects only mutual funds, with the usual concave parabola shape

and it takes a convex parabola shape in explaining life insurances participation decision. The invested amount is instead explained by age for bank accounts and mutual funds and this relation is hump-shaped. Life insurance purchasing is more common among the younger people of our sample, suggesting the intention of these people to protect his/her family towards the end of their life savings by investing in this type of financial product. Again, as we expect, we see that the number of children positively affects the probability of holding life insurances and negatively the others. Main house value has a strong positive relation with bank account, retirement accounts and life insurances decision and with investment amount for all of them. Other real estate value is significant just for bank accounts participation decision and for bank accounts, mutual funds and retirement accounts invested amount. Risk preference is not significant in explaining retirement account and life insurance investment.

GDP and HDI scores used in the chapter 6 well explain the relations between financial assets holding and country of residence. As we expect, in this model, countries where GDP and HDI are lower, show smaller probability and investment compared to Italy, *ceteris paribus*. Investment in bonds is usually low, also for countries with high GDP and HDI, compared to the Italian one, confirming the tendency of Italians to invest more in bonds rather than stocks.

Conclusions

The main objective of this thesis was to determine whether the fact of living in a particular European region rather than in another one could affect financial participation of European households. Besides, we wanted to investigate the influence of demographic and socio-economic variables at the individual and at the group levels. Various statistical models have been implemented. First, a multilevel regression model for the total amount held by households covering all financial assets. Second, the same model was compared with a logistic multilevel model to look at the differences between the determinants of the decision of investing and how much to invest. Finally, a Heckman selection model has been used to jointly consider the participation decision and the investment amount decision.

To test the meaningfulness of using a two-tier structure in a sample where the units were grouped by European region, we have used some statistical techniques implemented in particular by Bliese. These techniques aim to verify the existence of differences between the actual groups of the sample and some pseudo-groups, created specifically to define if the variability between groups and the variability within each group among the units were the result of a random process or if the agreement between the units belonging to the same group was due to internal homogeneity and the disagreement between the groups was due to heterogeneity. The result has been to find a high variability between groups and low variability within the same group such as to proceed with the deployment of two-level models, households and regions of residence.

Now we can answer to the three research questions defined at the beginning. After controlling for some socio-economic characteristics, living in different countries leads to differences in financial assets ownership and invested amounts. With respect to these

variables, countries play a significant role in the European landscape. Answering to the first question, we can say that country affects the European households financial portfolio choices. Answering to the second question, the results of the statistical models have generally confirmed some theories in the literature, such as the strong relationship between financial investment and household wealth, incomes or the influence of socio-demographic characteristics such as health, gender of the respondent or level of education. A significant factor is the financial risk preference: as we expected, risk aversion strongly affects households choices. Age was found to be a strong significant variable (whose relationship with the financial investment has a parabolic shape with concave side down) in explaining financial investment. Country-level Index (GDP) shows that economic level has a strong group effect on European households total financial investment. Moreover, a good health, having a partner or a high value of real property and family income, a high level of education and being resident in an economically advanced country is positively related with the financial investment. On the other hand, the number of children negatively affects it. Finally, we can answer to the third question: the comparison between the decision of participating and the invested amount for the different financial products led to the conclusions that socio-economic and demographics factors at the households level, together with economic and well-being factors at the region-level affects in different ways the ownership decision and the investment amount about finance in the European landscape. Different conclusions are found for each implemented model, so we refer to the previous chapter for more particular findings, more accurate for each research problem.

An interesting suggestion for future analysis in this field and with these objectives would be to implement a multilevel Heckman selection model.

References

Albright and Marinova (2010), Estimating multilevel models using SPSS, STATA, SAS and R

Alzugaray (2011), Welfare asset limits and savings behavior

Bickel (2007), Multilevel Analysis for Applied Research: It's Just Regression!

Bliese (2012), Multilevel Modeling in R (2.4). A Brief Introduction to R, the multilevel package and the nlme package

Campbell (2006), Household finance - The Journal of Finance, Vol. 61, No. 4, pp. vi+1553-1604

Campbell and Cocco (2003), Households risk management and optimal mortgage choice - The Quarterly Journal of Economics, Vol. 118, No. 4 (Nov., 2003), pp. 1449-1494

Cocco, Gomes and Maenhout (2005), Consumption and Portfolio Choice over the Life Cycle - The Review of Financial Studies Vol. 18, No. 2

Curcuru, Heaton, Lucas and Moore (2004), Heterogeneity and Portfolio Choice: Theory and Evidence - Handbook of Financial Econometrics

Guiso , Haliassos and Jappelli (2002), Household portfolios

Guiso, Haliassos and Jappelli (2003), Household stockholding in Europe: where do we stand and where do we go? - Economic Policy April 2003 pp. 123–170

Heckman (1978), Dummy endogeneous variables in a simultaneous equation system – Econometrica 46, 931-959

Hox (2002), Multilevel Analysis: Techniques and Applications

Hox, Multilevel modeling: when and why – P. 147-154 in I. Belderjahn, R. Mathar & M. Schader (Eds.) Classification, data Analysis and data highways. New York: Springer Verlag

Maas C., Hox J. (2005), "Sufficient sample sizes for multilevel modeling", /Methodology: European Journal of Research Methods for the Behavioral and Social Sciences/, 1: 86–92.

Paccagnella O. (2011), "Sample Size and Accuracy of Estimates in Multilevel Models New Simulation Results", /Methodology: European Journal of Research Methods for the Behavioral and Social Sciences/, 7: 111–120.

Pelizzon and Weber (2008), Are Household Portfolios Efficient? An Analysis Conditional on Housing – Journal of financial and quantitative analysis Vol. 43, No. 2, June 2008, pp. 401–432

Skrondal and Rabe-Hesketh (2004), Generalized latent variable modeling

Snijders and R.J. Bosker (1999), Multilevel Analysis

Soede, Vrooman, Ferraresi e Segre, Disparità nei sistemi di welfare - Unequal Welfare States

Verbeek (2006), Econometria

Wolff (2006), International Perspectives on Household Wealth

This paper uses data from SHARE wave 2 release 2.5.0, as of May 24th 2011. The SHARE data collection has been primarily funded by the European Commission through the 5th framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from 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 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions)."