Politecnico di Torino

Master's Degree in Management and Engineering

Financial Literacy, Education, and Investment Decisions of Italian Households

Supervisor:

Prof. Riccardo CALCAGNO

Candidate:

Nicol SOLARTE HERRERA

Abstract

This study aims to measure financial literacy and study its relationship to stock market participation and assets allocation. Using the cross-sectional Bank of Italy's Survey on Household Income and Wealth two types of empirical models were studied. According to the data, most individuals lack knowledge of fundamental concepts such as inflation and interest rates. Males, highly educated people, and individuals who are self-employed possess higher financial literacy. Regarding stocks and investment funds ownership, an important portion of households do not invest in this kind of assets. The individuals whose wealth is found in the last two quintiles are the ones who most participate in the market. In the first analysis, with probit models that incorporate the likelihood of stock market participation, financial literacy has been shown to have a positive and significant impact on the probability of participating. Likewise, additional characteristics associated with the choice to own stocks were also analyzed. The possession of a university degree, particularly in economics or statistical field, having a pension, being male and saving for retirement are all positively and significant related with the stock market participation. On the contrary, being a risk-averse person significantly reduces the probability of stockholding. The linear regression model, that were used to study the choice of fraction of financial assets invested in stocks and investment funds, shows that, similar to the participation model, the financial literacy, education and income are positive strong predictors, conversely, high risk aversion reduces the share of financial wealth invested in stocks.

Key words: *Financial Literacy, Stock Market Participation, Assets Allocation*

Summary

Introduction

Households represent a critical agent of the economy. They are the primary consumers and responsible for creating demand in markets, through the purchasing of goods and services for the day-to-day needs and preferences. They also represent the source of the labor force for businesses and industries. They are a significant source of government revenue by paying taxes. Likewise, households play a significant role in the financial economics. According to Tufano (2009), financial services and products used by households constitute a substantial portion of the financial industry in all advanced countries.

Every day households are faced with the task of making a number of decisions to achieve their objectives. Within this set of decisions, which ones related to the financial and welfare take a huge importance. Households need to manage means of payment (cash vs. credit card), forms of debt (personal vs. collateralized loans, fixed vs. variable rates), saving and investment vehicles, insurance contracts (accident property, health insurance), and financial intermediaries (financial advisory, money managers). The building of household wealth depends on this, on its financial decision-making ability. Substantially, it establishes how much money is saved, how household funds are invested, how the financial wealth is allocated, what types of investments are used, how much risk is assumed, and consequently how much return may be realized.

Now more than ever, more sets of financing and investment choices are available to households Guiso & Sodini (2013). The accessibility of financial information on the internet, and the expansion of the mutual fund industry have made it more affordable to access the equity market. However, with financial innovation, the complexity of financial instruments is expanding, and families must cope with new and more complicated financial products to keep up with the financial information. At the same time, the structural reform in the pension plan of Italy and other countries has caused an ongoing shift in decision power from the governments and employees toward private individuals Guiso & Jappelli (2009). Now, workers must decide both how much to contribute and how to allocate their pension wealth. Thus, individuals and families must assume more responsibility for their own financial well-being. In turn, this requires greater financial sophistication in both savings and investments management. The necessity of knowledge, proper literacy, and accurate information about the financial sector and how it works comes with much more importance than before. Households require, for instance, to have planning and computational capabilities as well as a good understanding of the trade-off that different

alternatives entail to optimally make financial-related choices. Furthermore, they can take care of their own decisions themselves, based on their knowledge and the information collected, or can rely on a third-party advisor. Alternatively, they can delegate to external experts the task of managing their finances.

How households choose when faced with their finances matters is influenced by numerous factors. Empirical evidence from developed countries shows that an important financial decision such as participation in the financial market is influenced by a variety of factors, including household financial wealth, age, education, financial literacy, risk aversion, trust in financial institutions, social interaction, homeownership, and social capital.

This study focuses on how financial literacy and education level affect the main financial decisions taken by the households in Italy regarding how they allocate their financial wealth. Particularly the participation in the stock market decision, the portfolio choice, and the equity diversification strategies. The first section is a review of other research works which argument is similar to the proposed in this study. Later in the dataset section, detailed description of the data, characteristics of the sample and statistical considerations are presented. In the following section, a descriptive statistic on the data used for estimation of multivariate models will be provided. Probit regressions regarding the decision to participate in the stock market and linear regressions aimed at estimating the fraction of financial wealth invested in stocks will be shown in the multivariate analysis section. Finally, conclusions and implications of the analyses will be discussed.

Literature Review

Financial literacy has been an important subject of study for many countries to measure the knowledge, behavior, and attitudes of its individuals toward this topic. There is a large empirical literature that evaluates and analyzes what are the effects that financial literacy has on the financial decision-making of individuals and households.

Lusardi & Mitchell (2014) in one of their recent studies, through economic models and experiments, confirmed the causal impact of financial literacy on economic decision-making and separately identified this effect from other factors, including education and cognitive abilities. They affirm that there are likely to be important benefits of greater financial knowledge, including savvier saving and investment decisions, better debt management, more retirement planning, higher participation in the stock market, and greater wealth accumulation.

Regarding household behavior, Lusardi & Mitchell (2007) compared the wealth holding across two courts of data from the Health and Retirement Study of US (HRS) in 2004 and 1992, they evaluated how successfully individuals plan for their retirement, whether financial literacy is associated with better planning, and whether retirement preparedness is associated with these behaviors. Finally, they found that those with low financial literacy are less likely to plan for retirement and therefore accumulate less wealth. Similar to this conclusion, in Italy, Fornero & Monticone (2011) found that as for retirement planning, financial literacy positively and significantly increases the probability of participating in a pension fund.

Financial literacy not only affects the behavior of households but also, it's an important matter at the moment of taking financial decisions such as selecting an optimal portfolio. Guiso & Jappelli (2009) studied the effect of financial literacy on portfolio diversification directly. Their source data was the 2007 Unicredit Customer's Survey which has indicators of portfolio choice, financial literacy, and many demographic characteristics of investors. They find that poor financial literacy is a main factor explaining lack of portfolio diversification and that the measures of financial literacy are strongly correlated with the degree of portfolio diversification.

Christelis, Georgarakos & Haliassos (2009) support the previous finding. Their results from the analyses of US Survey of Consumer Finances imply that financial sophistication and information have a significant role in facilitating investment and allocation of different stockholding choices.

The participation market is another important decision that individuals face when they are in adulthood and want to invest to increase their wealth. Van Rooij, Lusardi & Alessie (2011) used the DNB Household Survey to measure basic and advanced financial knowledge and its relation to the stock market. They found that there is a lack of financial literacy within Dutch households and concluded that this lack of financial literacy correlates with investors' decisions to participate in the stock market. Those who have low financial literacy are significantly less likely to invest in stocks.

Aside from the correlation between financial literacy and market participation. Cole & Shastry (2009), demonstrate, through an analysis of data from the U.S Census, that cognitive ability and education significantly increase participation. Those graduating from high school are significantly more likely to report higher income from retirement savings than those not graduating and, similarly, those with higher test scores are more likely to hold a wide variety of financial instruments, including stocks, bonds, mutual funds, savings accounts, tax-deferred accounts, and CDs.

From a European level perspective, a study that took data from Germany, France, Italy, Sweden and the UK, supports the hypothesis that the financial literacy of individuals affects their financial behaviors. Likewise, it states that people with more financial literacy could make better financial decisions compared with less financial literacy individuals (Nicolini & Haupt, 2019).

Regarding the Italian situation, previous research has documented a low level of financial literacy in the population, according to the Report on financial investments of Italian households, the financial knowledge remains low, as more than 80% of respondents answered incorrectly the advanced notions questions about finance. Fornero & Monticone (2011) also concluded, through an empirical analysis of the Bank Iltay's Survey on Households Income and Wealth, that most individuals lack knowledge of basic concepts such as interest and inflation and found that financial literacy has a positive and significant impact on the probability of pension plan participation.

Financial illiteracy is widespread and particularly acute among specific groups of the population, such as women, the elderly, and those with low educational attainment. (Lusardi & Mitchell, 2008). Particularly, women display much lower levels of financial literacy than the older population. In addition, women who are less financially literate are less likely to plan for retirement and be successful planners. These findings are unexpected not only because the literacy tests were quite simple, but also because the sample from the HRS 2004 consisted of people who had used credit cards and had obtained one or two mortgages.

As mentioned before, financial literacy is not the only determinant of the financial decision-making process. There are several different factors that affect these decisions. Johnston, Kassenboehmer & Shields (2016), used eight data waves from the Household Income and Labor Dynamics in Australia (HILDA) survey and concluded that non-economic dimensions, such as physical and mental health as well as education, cognitive ability, and personality traits, have an important role in determining financial decision-making. Cognitive ability is a significant predictor, with high ability associated with a high probability of decision-making responsibility; although, male cognitive ability is significantly more important than female cognitive ability. Similar results are found recently in the work done by Xu & Yao (2022), which examines the evidence from SHR. The authors concluded that conscientiousness, memory, and numeracy are favorable personal attributes for household financial decision-making. These attributes imply a relative advantage, so those financial decision makers at a disadvantage tend to have a lower total net worth and a lower financial net worth.

Dataset

The Survey on Household Income and Wealth (SHIW) is carried out every two years since 1977 by the Bank of Italy. It is the main valuable source of detailed information about the demographic, consumption, income, and savings of a representative sample of Italian households, as well as its economic and financial behavior.

With the aim of amplifying and extending the information collected, the SHIW has grown in scope including more variables such as the financial knowledge of the Italians, which was a section – answered by the household head- introduced in addition to the standard questionnaire for the first time in 2006 and it remains until now, except for the 2014 and 2012 wave. The formulation of the questions asked in this section has changed over the years, however, they remain the same focus on three main topics: interest, inflation, and risk.

Household, the reference person, and the aggregated data

Households are one of the main agents of any economy, it refers, according to the survey context, to a group of persons that normally reside in the same dwelling on 31 December of the year to which the survey refers and that contributed at least part of their income to the household. It also includes any members temporarily absent (e.g. on vacation, away for study, etc.) and any non-relatives living permanently in the home at 31 December of the reference year.

The reference person (R.P) is primarily responsible for or most knowledgeable about the household budget and is who provides the interviewer with all the information about the family. The other members are also interviewed, and they are categorized by their relationship with the R.P, for instance, spouse of R.P, child of R.P. Throughout this paper the reference person is also going to be mentioned as household's head.

Considering the fact of that in the survey not only the R.P is interviewed but also the other members, it's important to have it into consideration for understanding the distinction of the variable presented in the analysis. The variables available at individual level (such as age, sex, education) and the aggregated variables which are at household level (such as financial investments, income, and wealth). Given this, a strong assumption must be taken to make the statistical analysis. It's assumed the individual data of the R.P (such as age, sex, education and all the demographic variables) as the representative of the whole family.

Sample design

The target population focuses only on the officially resident households, excluding people living in institutions or those who are in the country illegally. The sample for the survey is drawn in two stages, with municipalities and households as, respectively, the primary and secondary sampling units. Before the primary units are selected, they are stratified by region and population size (primary sampling unit stratification). Starting from 2020 the households are stratified based on income and indebtedness.

Imputation and Replication

The missing answers affect the data quality of the survey, causing a significant degree of bias, making processing, and analyzing data more difficult. For this reason, imputed values were used to compute only the aggregated variables, such as income and wealth, where data is missing.

In order to properly obtain estimates of one or more population parameters, an assessment of the variability of the sample is necessary. The actual variability of the estimators can be determined only with techniques that consider both sample selection procedure and the nature of the estimators. For this reason, the Jackknife Repeated Replication (JRR) method was used, as it is reputed to have better statistical properties since it takes into account the structure imposed on the data by the complex nature of the sample design (stratification and two-stage sampling). This replication method obtains information on variability by reproducing the estimation phases on replicas of the original sample. For further information on the topic, please refer to Appendix 1.

Sampling weight

As mentioned before, the sample design must be considered to perform a statistical inference. It takes an important role, if it is ignored, generally it leads to inaccurate estimations of the variance, and the univariate and multivariate analysis can be biased Faiella (2008). Therefore, a survey weighting process was applied by the SHIW staff. Each person in the household is assigned with a sampling and population weight. Furthermore, there are another 279 replicate weights regarding the Jackknife method, needed for calculating the sample variance.

An in-depth analysis is available in Appendix 2, aimed at explaining the weight system, the reasons for the introduction and the statistical techniques that can be used to make appropriate statistical inference on data from SHIW.

Criteria for estimating descriptive statistics.

In order to get representative results for the total Italy population, as Faiella (2008) suggest in his accounting for sampling design in the SHIW study, a properly used of sampling weight is required for descriptive statistical analyses. Point estimates from unweighted analyses may be significantly skewed if some of the population is oversampled, if the survey has coverage issues or unit non-response is large, or if any of these characteristics are in any way connected to the study variables. Through the bias ratio, this will have an impact on the coverage of the confidence level statements.

Criteria for estimating multivariate analyses.

In statistical inference, the consideration of sampling features, sampling design and weighting, within the analysis and its impact on the results, has been assessed by several authors. Faiella & Gambacorta (2007) find that, the increasing variability induced by using weighted estimators is compensated for by the bias reduction. However, the use of weights, particularly in regression techniques, is controversial. Some authors, such as Nathan and Smith (1989), state that using survey weights in regression analysis involves a loss of efficiency. Moreover, (Hansen, Madow & Tepping (1983) and Lohr (2022) suggest using all the design features in regression models when the sample size is large and the sample size helps to mitigate the possible loss of efficiency due to survey weights. Regarding the SHIW, Faiella (2008) analyzed how the sampling design affects variance estimates and inference using the data collected by the survey wave 2002 One of its main conclusions was that using survey weights in regression analysis gives design unbiased parameters that are robust to model misspecification. Finally, Fiella suggested that the increase in bias associated with unweighted estimates is not compensated for by the decrease in the standard error. In this study, as many informed authors and knowledgeable people about the use of complex surveys suggest, the replicate weights and the sampling weight are going to be used for the statistical analysis.

Measurement of financial literacy

Regarding the financial literacy assessment, Lusardi & Mitchell devised a module on this topic for the SHR 2004 U.S. Their questions were designed to evaluate individuals' understanding of fundamental financial concepts such as the working of interest compounding, the impact of inflation, and risk diversification. This test has been widely used in other surveys, including the SHIW. In the 2020 wave, inside the saving and investments section, an additional module with

three questions on financial literacy was included as follows, (the questions were answered by the individual identified as the reference person of the household)

i. Suppose you put 100 euros into a savings account with a guaranteed interest rate of 2% per year. You don't make any further payments into this account, and you don't withdraw any money. How much would be in the account at the end of 5 years, once the interest payment is made?

    a. Less than 102 euros

    b. Exactly 102 euros

    c. More than 102 euros

    d. Don't know

    e. No answer

ii. Suppose you put 1,000 euros into a savings account with a guaranteed interest rate of 1% per year. Suppose furthermore inflation stays at 2 per cent. In one year's time will you be able to buy the same amount of goods that you could buy by spending today1,000 euros?

    a. Yes

    b. No, less than I could buy today

    c. No, more than I could buy today

    d. Don't know

    e. No answer

iii. In your opinion, the purchase of shares of one company usually provides a safer return than buying shares of a wide range of companies through a mutual fund?

    a. True

    b. False

    c. Don't know

    d. No answer

The first two questions regarding Compound interest, and Inflation, respectively, indicate whether respondents command the key economic concepts fundamental to saving. The third question evaluates knowledge of risk diversification, crucial to informed investment decisions.

Note that *Don't know* (DK) is always a respond option. This means the respondents are not forced to give a (random) answer and should therefore minimize guessing.

Variables

Base on the literature presented in the section 2, the independent variables for the multivariate models are grouped as follow:

Social and demographics:

I. Age: as previous studies concluded, age has a relationship with one of the main financial decisions that face a household, market participation. This increases strongly with age. Age is also linked with another kind of decisions such as borrowing. Studies have found that borrowing decisions as well as the selection of interest rates vary with age.

II. Gender: dummy variable whose value is equal to one if the reference person is a man, and zero if is woman. Many studies confirm gender differences in financial decision-making, financial literacy, and risk attitude. Women tend to be less knowledgeable about financial matters Baglioni (2018) and participate less in the stock market van Rooij, Lusardi, Alessie (2011).

III. Married or in a civil relationship: Marital status and children can dramatically change the financial behavior of a person. The saving, life insurance, wealth accumulation and asset allocation are affected Love (2010). Based in Guiso (2003) and Bertaut (2016) it is expected that marriage increases stock market participation.

IV. Education: is a dummy variable whose value is equal to one if the reference person obtained a 3/5-year degree or a postgraduate certification, zero in all other cases. Some studies suggest that the level of education has a positive relationship with participation in the stock market, as well as affecting saving and investment behavior.

V. Work status: two dummy variables are created, one for respondents who are retired and the other one for self-employed individuals. Those who are self-employed have a different risk profile compared to those employed or retired, they are already exposed to high risk in the labor market and may therefore be less likely to hold stocks Heaton & Lucas (2016).

VI. Financial literacy: is measured in two ways: first, a dummy variable is used, it takes the value of one when the respondent (the reference person of the household) can answer correctly all three questions asked in the financial knowledge module of the survey. Second, a categorical variable for the number of correct answers (taking values from 0 to 3).

VII. **Economic Degree:** Thanks to the survey's wealth of data, it's also possible to take additional variables into consideration as this one. A range of different fields of diploma degree is available on the survey. Particularly, for this case, the Economic or Statistic diploma was selected as it represents a higher economic education and interest for this field. Therefore, a dummy variable takes the value of one if the reference person has a diploma on Economic or Statistic and zero in all other cases.

VIII. **Internet Usage:** Likewise, the SHIW accounts for the internet or email usage of households. It is expected that the use of internet generates a positive effect on participation.

Economic variables

IX. **Net disposable Income:** it is a numeric variable; the logarithm of net disposable income is considered to avoid skewness. This variable excluded the income from the financial investments made in stocks or investment funds, which are clearly correlated with stock market participation.

X. **Net wealth:** Likewise, a logarithm is applied to the net wealth of the reference person, which like the net income, does not include the stocks or investment funds. Several studies have confirmed the strong implication of economic factors over the financial decision of households, particularly in the holding stocks likelihood.

Additional variables:

XI. **Risk aversion:** SHIW also measures the financial risk-taking of the respondent, providing 4 choices from the highest to the lowest risk-taking. As the literature suggests, the main investment decisions, such as the heterogeneity of portfolios and stock holding depend on the risk attitude of the investor Dohmen, Falk, Huffman, Sunde, Shupp, Wagner (2021).

XII. **Saving reasons:** in the SHIW there are six options available for the question regarding the reasons for saving. To this study, three options were selected as dummy variables: saving for major reasons, saving for retirement, and saving for education. It is expected that those who understand the importance of saving for retirement will be more inclined to engage in the stock market,

Finally, as it's mentioned before, the households face to more than one single financial decision, in this study two dependent variables will be used to analyze how the financial literacy, the education

and other demographic and economic variables affect the main financial decision that represent the dependent variables, which are: stock market participation and assets share invested in stocks.

I.   Share invested in stocks: Is the ratio of stocks and investment funds values to total financial wealth.

II.   Stock market participation takes the value of one whether the household hold any stock or investment fund.

Detailed information about the variables is presented in the Annex 1

Sample selection

The broad purpose of this analysis is to depict the influence of financial literacy, education and other demographic variables in the main investment choices that take the Italian households. To this end, the analysis is restricted to the households that are in the condition to make investment decisions even if modest in value. Therefore, the selected sample is composed by households with a positive total net worth and total financial assets greater than zero.

The most recent survey, the 2020 wave, is used as the primary and only source of data. It is a single imputed data set with regression imputation and with JRR replicated weights. This wave covers 6239 households composed of 15198 individuals. However, the selected sample covers 5911 households. For all the statistical calculations, the STATA program and its available survey functions were used, as well as the Jackknife option for the definition of the sample design and the estimations. (For more information about these functionalities, please refer to Jackknife estimation for survey data manual available on the official site of (STATA)).

Descriptive Statistics

Financial Literacy, evidence

The following figures show the current evidence about the financial literacy of the total Italian households and of the selected sample. Figure 1 represents the answers to the interest compounding question, showing that slightly more than half sample (52%) give a correct answer, while 17% do not know. Among those giving a wrong answer (28%), 18% answer "Exactly 102 euros" which represents a relatively less serious mistake than "Less than 102 euros" (10%).

The answers to the inflation test are shown in Figure 2, where 61% answer correctly, while 22% do not know and almost one quarter of the respondents give an incorrect answer (14%). Finally, in the question about risk (Figure 3) 57% give a correct answer, while one third of respondents don't know and a few provide an incorrect answer (10%). In all the three questions the proportion of sample who do not provide an answer is around 2,5%.

Regarding the overall performance, see Figure 4, about one third of the Italian households give three correct answers or provide at least one correct. About 10% do not know the answer to all tests and 1,18% respond wrongly. As can be seen in the figures, the answers provided by the entire sample are pretty similar to the answers provided by the selected sample.

Comparing the answers on all the three tests, Italians seem more familiar with interest and inflation than with risk. This may be related, among other factors, to the personal experience with inflationary environments, as a large share of the sample experienced relatively high inflation during their prime age (in the 80's). The risk question has the second highest proportion of correct answers; however, it also has the highest proportion of DK compared with the inflation and risk question. This lack of knowledge on risk diversification can be explained due to only 23% of households have at least one financial asset other than a bank or post office saving account, which reflect the lack of diversification in its financial wealth.
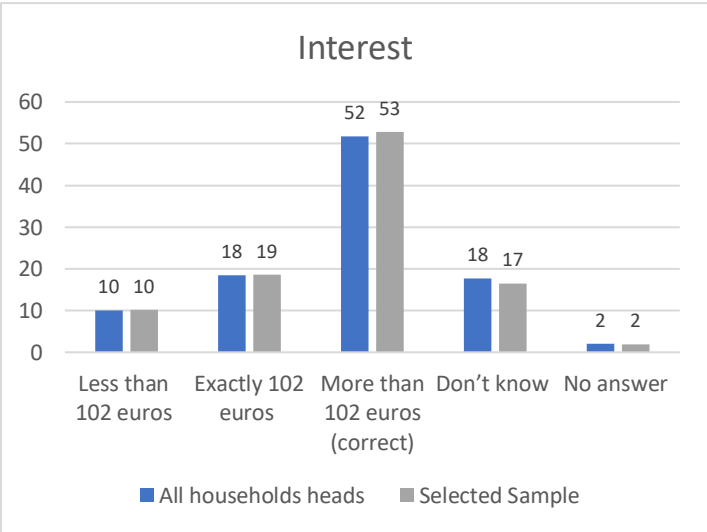

Figure 1 Answers to the interest compounding question
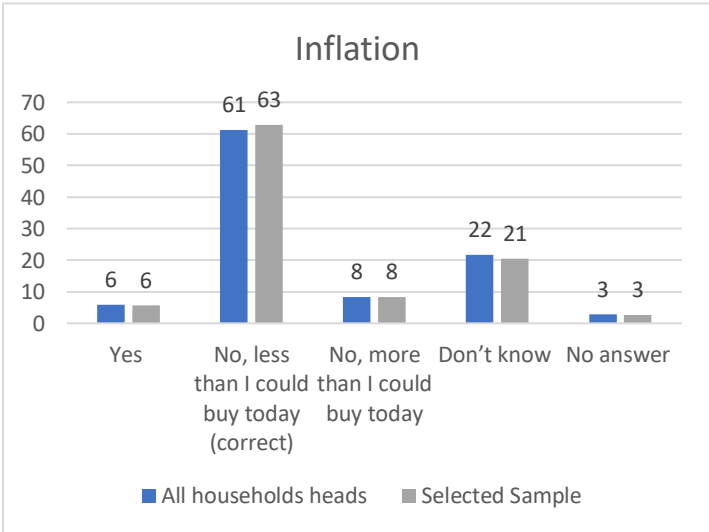
Figure 2 Answers to the inflation question
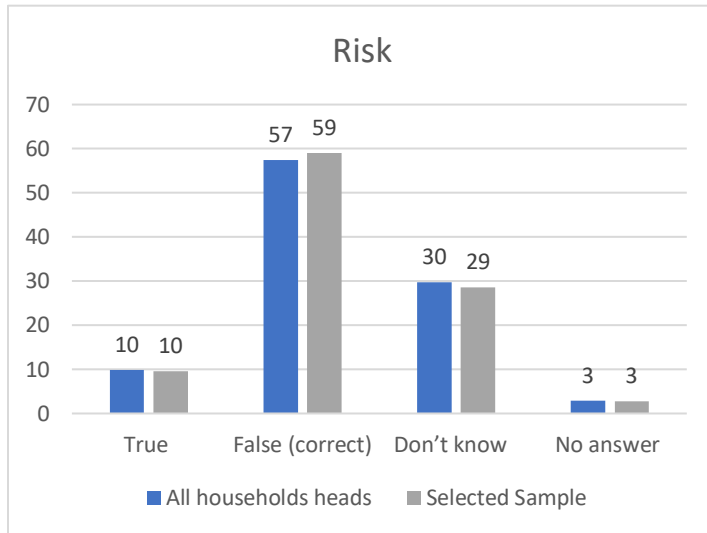
Figure 3 Answer to the risk question.
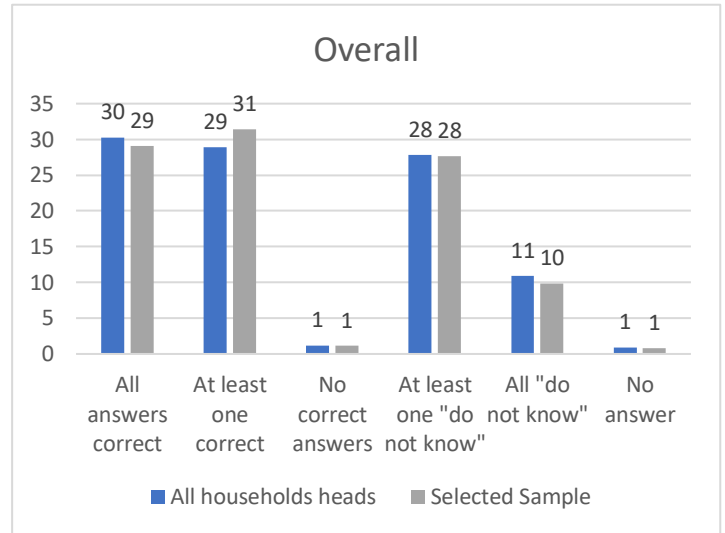
Figure 4 Overall performance

TABLE I. shows financial literacy performance by socio-demographic characteristics. The financial literacy age profile is hump-shaped for all three questions, with a peak in performance in the age class 31-40 years. As expected, the age classes with the highest share of correct answers report the lowest proportion of DK. This pattern is the same for the overall performance.

Regarding the performance by gender, men are more knowledgeable than women in all questions. Both genders show a better performance in the inflation question with respect to the other ones. Women have a very high proportion of DK, showing a major lack of literacy in the question regarding the risk in purchasing shares (37%). These findings are similar to those reported by Lusardi & Mitchell (2008) , and the findings in another literacy survey Lusardi & Mitchelli (2007).

As expected, financial literacy is strongly monotonically with the level of education, this is true when looking at both correct and DK answers. Those with the lowest level of financial literacy are concentrated in the lowest educational categories: no education and primary school. Conversely, those with higher education: upper secondary school, university degree, and postgraduate, have better financial literacy. In overall, only 15% of those with at most primary education (representing 18% of household's heads) could answer all three questions correctly, compared to 54% among those with a university degree.

The Self-employed (including small business owners, owners of members of family business, and members of liberal professions such a lawyer, architects, and so forth) display better knowledge

than employees. This behavior can be explained due to self-employed households are wealthier and more used to managing their personal/business finances. The non-employed (including unemployed, homemakers, students and so on) have the lowest performance in all measures. In overall, the 21,43% of no-employed individuals responded correctly to all the questions compared to 50% and 35% of self-employed and employed, respectively.

TABLE I Performance by socio-demographic characteristics

	Interest		Inflation		Diversification		Overall	
	Correct	DK	Correct	DK	Correct	DK	Three correct	At leat 1 DK
Age								
Age <= 30	55,82	13,59	58,67	20,05	55,24	31,97	35,2	26,02
Age 31-40	63,19	8,84	65,48	13,39	64,45	22,3	36,48	22,36
Age 41-50	55,73	12,69	64,37	19,08	61,34	24,66	31,36	25,29
Age 51-65	54,09	13,16	62,6	20,29	63,08	24,71	34,01	26
Age >=65	44,03	26,36	58,07	27,04	50,46	37,21	24,84	32,14
Gender								
Male	57,86	11,76	66,64	17,34	63,77	23,77	35,38	24,57
Female	43,49	25,8	54,05	27,8	48,91	37,87	23,37	32,17
Education								
No education	15,95	57,18	32,21	53,58	31,74	60,96	2,64	43,63
Primary (Isced 1)	32,3	38,52	48,53	35,63	37,56	48,54	15,23	36,73
Lower sec (Isced 2)	47,46	18,29	57,77	24,35	50,25	34,23	22,89	31,68
Upper sec (Isced 3)	61,25	7,58	69,2	13,91	69,01	19,09	38,9	23,42
Degree + (Isced 5+)	72,84	3,94	74,79	8,81	80,73	11,64	53,65	13,91
Occupational Status								
Employed	58,75	9,57	65,86	16,88	63,5	23,72	35,44	24,94
Self-employed	70,62	5,94	72,03	10,12	77,96	13,54	49,87	15,62
No-employed	41,96	26,26	55,36	28,22	47,79	38,2	21,43	33,07

Source: SHIW 2020 - Weighted data. Sample: all households (N=6236)

The structure of household portfolio

TABLE II reports the ownership of financial and real assets (the break-down structure presented for the financial assets is based on the adoption of classification utilized by the Bank of Italy on its official SHIW 2020 report. For additional information about this structure refer to Annex 2). The focus is on the different forms of savings and investments asked for during the survey. As can be seen, the concentration of financial assets ownership is on the bank deposits for all the wealth categories. Managed investments, which include investments founds, ETFs and saving managed,

are the second most owned financial assets. In contrast, the securities issued abroad are the last owned assets by households, reflecting that most of Italians only invest in domestic financial assets.

Additionally, TABLE II shows that financial market participation is strongly correlated with wealth. The assets ownership increases with the level of wealth, for most assets, it doubles with each additional wealth level. In particular, the proportion of households with stocks is 20,19% in the last quintile of wealth, against 0,74% in the poorest quintile, for managed investments the proportion increases from 1,73% to 35,03%. Also, it is possible to observe not only the correlation with financial market participation but also with financial literacy. The financial assets ownership of the individuals with advanced financial literacy, that is those who respond correctly to all tree questions, is significantly greater than those who do not know at least one answer from the financial knowledge module of the survey.

Regarding real assets, the ownership is slightly higher than the financial assets and increases with the level of wealth and financial literacy. The affluent are also more likely to invest in risky real assets, particularly in other real statements (49,45%) and business (44,47%).

In overall, only around 25% of selected households have at least one financial asset other than bank deposit or post saving office account, mostly in the form of managed investments. The mean financial wealth of these households is around €230.000 compared to €68.000 of those who only hold bank deposits. The ownership of assets increases with wealth and financial sophistication, the real assets ownership is overall higher than financial assets.

TABLE II Ownership rate of real and financial assets (%)

	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile	Selected sample	Correct answers	At least 1 DK
Financial Assets								
Deposits	100	100	100	100	99,98	100	100	99
Managed investment schemes	1,73	4,71	7,06	15,29	35,03	13,45	25,21	4,45
Equity shares and participating interest	0,74	2,5	3,42	8,6	20,19	7,48	14,34	1,83
Private Sector bonds	0,64	0,94	3,07	9,09	17,66	6,64	10,05	3,18
Government securities	0,55	2,53	4,64	9,45	11,72	6,08	9,33	3,61
Securities issued abroad	0,5	0	0,83	0,91	5,01	1,53	2,79	0,51
Other	0	0,41	1,1	2,11	3,75	1,56	2,63	0,8

Real Assets								
Primary Residence	5,56	78,71	94,12	94,89	96,2	77,31	83,37	70,52
Other real estate	4,38	15,01	20,15	36,2	68,36	24,87	44,82	19,89
Business equity	5,14	7,75	10,36	18,77	44,77	18,13	22,15	11,16
Valuables	79,42	79,77	85,11	90,67	92,97	85,97	87,98	83,84

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911)

The very different portfolio composition presented previously is reflected in the divergence in the proportion of financial wealth owned by households. TABLE III shows the composition of wealth, separately for financial and total assets. Statistics are computed as the average proportion of each type of assets with respect to the total financial wealth and total wealth. Households belonging to the poorest quintile primarily have deposits account as the principal source of their financial wealth, the deposits represent more than 90% of the total financial wealth for the first three quintiles. The fraction of managed investments, private-sector bonds and Italian government securities rises gradually across the central net wealth classes. The wealthiest 20 per cent of households are the ones most likely to directly own private-sector bonds and to entrust the management of a significant proportion of their financial assets to investment professionals. Regarding stocks, the evidence shows that less than 1% of the financial wealth of those in the bottom half comes from investments in stocks, and only 4,2% for the last quintile, reflecting that even the wealthiest Italians tend not to hold stocks directly by themselves. Once again, financial literacy seems to have a significant impact on financial wealth composition, the main difference between the most financially savvy people and people who lack financial knowledge is on managed investments and shares.

The second part of the table reports the composition of total assets. In contrast to the composition of financial wealth, large difference by wealth categories is seen: while for the rich over 86% of wealth is invested in real estate (including primary residence) and financial assets, for the less wealthy total assets are mostly invested in financial assets and valuables.

TABLE III Composition of household wealth

	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile	Correct answers	At least 1 DK
As a fraction of total financial assets (average)							
Deposits	97,43	94,51	90,72	79,56	64,84	75,45	92,62
Managed investment schemes	1,29	3,07	3,67	7,31	17,11	12,32	2,27
Equity shares and participating interest	0,18	0,71	1,05	2,5	4,2	3,04	0,47
Private Sector bonds	0,25	0,34	1,19	3,63	5,53	3,47	1,46
Government securities	0,4	0,95	1,83	4,92	3,53	3,06	1,73
Securities issued abroad	0,14	0	0,38	0,16	1,13	0,52	0,25
Other	0,3	0,4	1,14	1,91	3,65	2,13	1,2
As a fraction of total assets (average)							
Financial Assets	60,37	14,23	10,3	13,02	15,64	19,71	23,39
Real estate	9,79	79,89	85,14	80,7	70,69	70,13	65,18
Business equity	2,43	3,69	2,85	4,53	12,29	7,07	3,39
Valuables	27,41	2,18	1,7	1,75	1,37	3,09	8,04
Liabilities	7,23	10,93	6,57	3,66	3,69	9,29	3,6

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911)

For the purposes of the descriptive analysis of household portfolios, real and financial assets can also be categorized based on risk: risky financial assets include listed and unlisted stocks, investment funds, ETFs, managed savings, and corporate bonds. Classifying investment real estate and business wealth as risky assets. A more comprehensive definition of total risky assets includes risky financial assets, business equity, and investment real estate. This approach is inspired in the detailed paper of Guiso & Japelli (2005). Thus, TABLE IV reports statistics for risky financial assets and risky total assets. The first part of the table (Ownership) reports the proportion of Italian households who hold at least one of the risky financial assets mentioned above. As seen, it highlights large differences in ownership: in the central net wealth group almost 3% report owning at least one financial risky asset, against almost 45% in the top quintile. Regarding the ownership of total risky assets, this increases compared to the risky financial ones, since it includes the possession of another real estate different from the primary house and the entire or partial business ownership. Another way to reflect the participation of households in the risky financial

market is by accounting for the amount invested in the risky assets. This measure is reported in the second part of the table (Asset share). As seen, the fact that the top 20% has larger assets shares invested in risky financial assets reflects not only the different amounts invested by this group, but also the fact that households in the 0–80% group participate less in financial markets. Other important differences are presented for the total risky assets: the ratio between total risky assets and total assets is around 7,5% for the central's wealth quintiles and almost 39% for the wealthiest 20 per cent of households, indicating the richest segment of the sample has much greater inclination to invest in businesses and real estate.

TABLE IV Ownership and asset composition of risky assets

	1st quintile	2nd quintile	3rd quintile	4th quintile	5th quintile	Correct answers	At least 1 DK
Ownership							
Risky financial assets	2,59	6,97	9,45	23,04	44,89	33,27	6,29
Risky total assets	11,69	26,08	34,39	57,48	88,24	64,85	28,27
Asset Shares							
Risky financial assets/Total financial Assets	1,67	3,81	5,19	10,92	23,28	16,78	2,99
Total risky assets/Total Assets	3,81	7,17	7,9	15,62	38,55	23,34	8,55

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911)

Based on the current discourse, it appears that one of the best options for explaining stock market participation and portfolio heterogeneity is the amount of household resources as well as financial literacy.

To summarize, the evidence regarding the composition of household portfolios within the SHIW emphasizes several aspects that can be summed up like this. To begin with, the financial wealth of household is highly concentrated on deposits and the household portfolios are highly heterogeneous. Second, a lot of people don't make direct or indirect stock market investments through investment accounts or managed savings, especially those with modest wealth. Third, one of the most significant factors in dividing household portfolios into different categories is probably household resources, which have a strong correlation with the ownership of risky financial instruments. However, as stated by Guiso & Japelli (2005), since wealth may be correlated with risk aversion, transaction costs, information costs, and other economic variables, the correlation between wealth, asset ownership, and asset amounts may be the result of multivariate factors.

Among these variables, financial literacy highlights, as it has shown to have influenced stock market participation decisions and portfolio diversification strategies.

A detailed descriptive analysis of financial literacy, wealth, and other economic and demographic characteristics in relation to stock market participation is below.

Financial Literacy and Market Participation

In the selected sample, 14,61% of households own stock or investment funds. Thus, as in many other European countries, many households do not participate in the stock market. This figure, however, hides major differences among demographic groups. As reported in TABLE V stock ownership increases slightly with age, except for the oldest households, those older than 65. Those between 51 and 65 years old are the ones who hold the most stocks or mutual funds. Stock market participation is much lower among women than men, a finding also reported in other studies see also Haliassos & Bertaut (1995) and consistent with the sharp differences in literacy between women and men Lusardi & Mitchell (2008). Regarding the education level, stock ownership increases sharply with education. As expected, no one without education holds stocks or mutual funds. Only a small fraction of those with low education own stocks. However, even the large majority of those with a university degree do not participate in the stock market. Thus, factors other than educational level can prevent stock ownership. Note that the results are similar when considering the level of financial literacy. Only 27% of households that display advanced literacy knowledge, that is those who responded to all the answers correctly, participate in the stock market. As the previous evidence showed, even people with high levels of education do not always score highly on financial literacy. This implies that education is not always a suitable proxy for literacy and that models of portfolio choice may need to take both factors into account to explain behavior toward stocks. Stock market participation increases strongly with net disposable income and wealth, they have almost the same pattern. The net disposable income is the sum of payroll income, pensions and net transfers, net self-employment income and property income minus interest payments. Meanwhile, wealth is the sum of real assets, deposits, savings, bonds, government securities and trade credit or credit due from other households minus financial liabilities. Almost more than one third richest individuals in Italy don't hold direct stocks or mutual funds, this is at the core of the much-mentioned stock market participation puzzle. Even in wealthy households, stock investment is surprisingly uncommon, indicating the potential importance of

non-economic costs. Furthermore, stocks are complex assets, and many households may not know or understand stocks and the workings of the stock market. This is one explanation for stock ownership deficiency that has not yet received enough attention in the literature.

TABLE V Stock market participation across subgroups (%)

Age		Financial Literacy Level (1)	
Age <= 30	14,79	Advanced	27,19
Age 31-40	15,66	Low	4,33
Age 41-50	16,63		
Age 51-65	18,44		
Age >=65	11,24	**Net disposable income**	
		1st quintile	1,81
Gender		2nd quintile	4,68
Male	18,68	3rd quintile	9,48
Female	8,95	4th quintile	17,13
		5th quintile	36,85
Education			
No education	0		
Primary (Isced 1)	3,31	**Net wealth**	
Lower sec (Isced 2)	10,54	1st quintile	2
Upper sec (Isced 3)	16,93	2nd quintile	5,48
Degree + (Isced 5+)	33,36	3rd quintile	7,87
		4th quintile	17,96
Occupational Status		5th quintile	36,09
Employed	16,26		
Self-employed	25,49		
No-employed	10,27		

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911). (1) Advanced level refers to the reference person who responds correctly to all the three questions of the financial module. Low level refers to the reference person who doesn't know the answer ow at least on question pf the module.

Explanatory Variables

Correlation Matrices

Before making a detailed description of the potential explanatory variables, two correlation matrices were made to look into potential relationships between the independent and dependent variables. A tetrachoric correlation was used for the dummy variables (Figure 5), and a point biserial Pearson correlation for numeric variables (Figure 6) .

The choice of tetrachoric correlation as statistical method is based, mainly, on its suitability for binary variables and its ability to overcome assumptions violated by traditional correlation

measures. Prior research in related fields such as determinants of stock market participation and risk management has successfully employed tetrachoric correlation to explore analogous relationships Kaustia, Conlin, Luotonen (2023) Noyan & Şimşek (2012). As can be seen in Figure 5 participation in the stock market and university degree, advanced financial literacy (that means that respondent answers all the three questions correctly), economic degree and use internet are highly correlated. With tetrachoric correlation coefficients (rho) ranging from [0.2; 0.4], other variables showed a positive correlation with the dependent variable, such as: the fact of being male, being self-employed, and having a less/medium financial risk aversion. Conversely, being financially strong risk-averse is negatively correlated with stockholding, rho=-0,35.

The perfect negative correlation observed between certain variables in the tetrachoric correlation matrix, such as pensioner and self-employed dummy variables is because naturally they are mutually exclusive. Meaning the presence of occurrence one automatically excludes the possibility of the other happening. An individual who has a pension can't belong at the same time to the labor force or being self-employed. The same happens among the variables of saving for major reasons, saving for retirement, and saving for education, as well as among the three financial risk-taking profile variables (high, medium, and low). Despite the negative correlation observed, these variables can still be suitable candidates for a probit regression, they provide a clear distinction between the possible outcomes, enhancing the interpretability of the model.

	Y	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Y	1.0000														
A	0.2695	1.0000													
B	0.1868	0.4971	1.0000												
C	0.4417	0.1153	0.1345	1.0000											
D	-0.1288	0.0315	-0.1757	-0.3676	1.0000										
E	0.2791	0.3288	0.1866	0.4436	-1.0000	1.0000									
F	0.4240	0.1990	0.1214	0.3594	-0.2071	0.2659	1.0000								
G	0.4482	0.1783	0.1547	1.0000	-0.3862	0.3863	0.3549	1.0000							
H	0.4650	0.2426	0.3631	0.6042	-0.6925	0.5810	0.4209	0.4395	1.0000						
I	-0.0371	0.0501	0.0046	0.1106	-0.2959	0.1176	0.0596	0.0934	0.2189	1.0000					
J	0.0956	0.1079	0.2126	0.1933	-0.1397	0.0822	0.0687	0.1135	0.2374	-1.0000	1.0000				
K	0.1214	0.0299	-0.0568	-0.0982	0.2162	-0.0290	0.0151	-0.0119	0.2148	-1.0000	-1.0000	1.0000			
L	0.3040	0.2615	0.0703	0.2162	-0.1226	0.2794	0.1384	0.2806	0.2011	0.0993	0.0399	-0.0117	1.0000		
M	0.3020	0.1260	0.1358	0.2585	-0.1612	0.1329	0.0274	0.2085	0.2652	0.0062	0.0543	-0.0044	-1.0000	1.0000	
N	-0.3464	-0.1505	-0.1434	-0.2856	0.1742	-0.1715	-0.0432	-0.2603	-0.2824	-0.0187	-0.0590	0.0057	-1.0000	-1.0000	1.0000

Figure 5. Tetrachoric correlation matrix. Source: SHIW 2020 Unweighted data. Sample: Selected households
Dark green ≥ 0.4; 0.2≤ light green ≥ 0.5; -0.2 ≤ light red ≤ 0.4; dark red ≥ 0.4

Owning a stock or mutual fund	Y	Pensioner	D	Internet	H	High financial risk	L
Male	A	Self-employed	E	Save for major porposes	I	Medium financial risk	M
Married	B	All answers correct	F	Save for education	J	Low financial risk	N
University Degree	C	Economic or statistical degree	G	Save for retirment	K		

Furthermore, a point-biserial correlation which is a special case of the Pearson correlation was used to measure the relationship between the independent continuous variables (age, income, and wealth) and the stockholding which is a dichotomous variable (Y in the figure). As one might anticipate, participation in the stock market is highly correlated with income and wealth. Contrarily, age has a weak correlation, even negative. This can be explained due to an important proportion of respondents, 40%, are over than 65-years-old, and this is the class age with the lowest percentage of market participation.

	Y	Age	Age2	Ln_Inc~e	Ln_Wea~h
Y	1.0000				
Age	-0.0296	1.0000			
Age2	-0.0420	0.9897	1.0000		
Ln_Income	0.3767	-0.0841	-0.1049	1.0000	
Ln_Wealth	0.3273	0.0544	0.0336	0.6742	1.0000

Figure 6. Point biserial Pearson Correlation. Source: SHIW 2020 Unweighted data
Sample: Selected Households

Y: Owning a stock or mutual fund   Age2: Age squared

Explanatory variables analysis

This section described a synthesized version of the previous analysis in order to present the explanatory variables using weighted averages and proportions available in TABLE VI. All the statistics consider only the selected sample, which are the households with strictly positive net worth and financial assets. The abbreviation SH and NSH will be used to refer to stock holding and not stock holding, respectively.

Regarding the mean age of the reference person, it seems that the SH are younger than the NSH. This observation coincides with the conclusions of Agarwal, Driscoll, Gabaix & Laibson (2007), who found that sophisticated financial choices peak around 50 years old.

Analyzing the percentages of males and females inside the selected sample, as found before, the males are the majority in the stock market compared to females. Less than 30% of those who own stocks are women.

With respect to the marital status of the reference person, it seems that the SH, on average, cohabitate more with a spouse or a partner compared to NSH. When a person gets married, she or he could experience dramatic changes in their financial position; wealth increases or divides, and

spending needs changes in relation to housing. On the other hand, a young individual could have a completely different investment profile based on priorities and needs which clearly differ from those of a married person.

The percentage of graduates differs substantially between SH and NSH. Inside the stock market participants, those households' heads that have a degree or postgraduate qualification are more than half respect to those that are outside the market. It's also important to highlight that only 16% of all households' heads with positive net and financial wealth in Italy have, at least, a university degree. This represents a significant low proportion that must be examined by the relevant authorities. However, they are more than one third in the class of stockholders. This suggests that graduation leads to benefit from the equity risk premium more than people with lower educational qualifications.

Based on the study made by Alessie, Lusardy & Rooij (2011) which included the retired persons in their empirical work. The pensioner variable was added to account for the fact that some households may be in the decumulation phase of their life cycle and enjoy a pension. This led to retirees allocating their pension savings to stocks or mutual funds. The respondents which have a pension are 40%, they are more in the class of NSH than SH. However, this could be explained by the fact that, on average, the pensioners are 74 years old, and this is the class age with the lowest stock market participation.

Likewise, a dummy for self-employed was also added to account for those who are already exposed to high risk in the labor market and may therefore be less likely to hold stocks Heaton & Lucas (2016). Although self-employees are around 15%, among the equity holders' class there are almost one third of them. The reason could be the fact that they have higher net wealth compared to those with a different occupational status.

Regarding financial literacy. Although household reference persons who respond to all the questions correctly are around one third, there are more than 50% in the class of equity holders. This suggests that financial literacy represents an advantage and confidence in the market over people with lower financial knowledge.

As demonstrated by the findings of Christiansen, Joensen, Rangvid (2007), who found that controlling for economics education is more crucial than controlling for educational attainment, on the basis that informational advantages economist are more likely to hold stocks than otherwise

identical investors. It is decided to include an additional variable related to the type of degree, economic or statistical diploma. Although the data shows that only 4% of individuals possess this title, they are around 12% in the class of equity holders. Demonstrating that having an advantage of economic knowledge is related to participation in the market.

Finally, based on the data, seems that SHs have a much higher probability of using the internet than NSHs. However, it is challenging to both estimate and prove a causal relationship between internet use and stockholding.

TABLE VI Demographic explanatory variables for stock market participation

Stock market participant	Age	Male	Married	Degree (5+)	Pensioner	Self-employed	FL all correct	Economic degree	Internet usage
Stock holders	57	74%	67%	38%	32%	27%	58%	12%	90%
Not Stock holders	60	55%	51%	13%	41%	13%	26%	2%	67%
Total	59	58%	54%	16%	40%	15%	31%	4%	70%

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911)

TABLE VII presents the net wealth (including real assets) and the net income as the main economic explanatory variables. It should be mentioned that neither wealth nor income include stocks and mutual funds, which are clearly correlated with stock market participation.

It is possible to observe the average net worth and net income households by quintiles groupings and, in the last three rows, the same variable assumed to be SH or NSH. The inequality among Italian families is evident. According to the official report of the SHIW 2020 provided by the bank of Italy, the Gini index[1] was 68,2% in 2020 and it was practically unchanged compared to previous years (2016) (Banca d'Italia 2020).

Regarding the population analyzed between stockholders and non-stockholders, the data shows that the average net worth between the two categories is different. Shareholder's wealth is around four times that of NSH. Similarly, the net income of those who participate in the stock market is twice compared to those who do not.

---

[1] The Gini index is a composite indicator of the degree of inequality, it is expressed in percentage values, ranges from 0 (perfect equality) to 100 (maximum inequality)

	Net Wealth	Net Income
1st quintile	7.956	11.060
2nd quintile	73.350	20.030
3rd quintile	148.326	28.203
4th quintile	265.204	41.570
5th quintile	1.129.511	95.516
Stockholers	830.612	75.920
Not stockholders	261.949	34.691
Total	345.035	40.715

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911)

Finally, TABLE VII presents additional variables regarding the financial risk attitude and the reasons for saving. Over 50% of those surveyed say they take on a low level of risk (low returns, with no risk of losing the invested capital). While only 2% say they prefer investments that offer very high returns, but with a high risk of losing part of the capital. To conclude, a large majority of NSHs (57%) expressed a desire to avoid taking financial risk, while among the SHs, 66% have a medium financial risk-taking profile.

TABLE VIII Additional explanatory variables for stock market participation

Stock Market Participant	Financial risk taking attitudes			Saving Reason		
	High Risk	Medium Risk	Low Risk	Major purposes	Old-age provision	Education
Stock Holder	4%	66%	29%	9%	36%	20%
No Stock Holder	1%	41%	57%	9%	27%	19%
Total	2%	45%	53%	9%	29%	19%

Source: SHIW 2020 - Weighted data. Sample: selected households (N=5911)

Multivariate Analysis

Probit Models: decision to participate in the stock market

TABLE IX reports the results from probit regression where the dependent variable takes the value of one if an individual holds stocks either directly or through investment funds. Three different models were used: a basic model that includes the traditional variables which are comparable to the majority of earlier studies (Model A), a second model in which the measure of financial literacy

is added (Model B) and a third model in which, in addition the financial literacy variable, secondary demographic and behavioral variables are also included (economic degree, internet usage, reasons for saving and financial risk attitudes).

According to the basic model, income and wealth are the most important predictors of participation in the stock market, followed by gender, education, and participation in a pension plan. These covariates are statistically significant and increase the probability of holding stocks. Conversely being married reduces the probability of participating in the market.

Model B demonstrates that even with a wide range of demographics factors included in the first model, financial literacy also matters for stock ownership. Individuals with an advanced level of financial knowledge are also more likely to participate in the stock market.

Given the presence of more variables in the third model, some covariates such as married and university degree which were statistically significant are no longer significant. Furthermore, the estimate of advanced literacy does not change much after accounting for these additional factors. Respondents who possess an economic or statistical degree, have access to the internet, and save for retirement are more likely to hold stocks. Contrarily individuals who have a financial risk-adverse attitude are less probability to participate in the market.

Overall, even after considering the additional variables such as the specific economic knowledge (economic degree) or the attitude toward financial risk, the estimates show that financial literacy influences stock market participation above and beyond the impact of the conventional determinants of stock ownership.

TABLE IX Probit model for stock market participation. One, two or three asterisks means that the predictor is significant at the ten, five and one percent level respectively. Models estimated considering weighted data.

	Model A			Model B			Model C		
	Coef.	*P-value*	*AME*	*Coef.*	*P-value*	*AME*	*Coef.*	*P-value*	*AME*
Male	0.259**	0.002	*0.047***	0.235**	0.005	*0.042***	0.206**	0.013	*0.036***
Married	-0.137**	0.037	*-0.026***	-0.106*	0.098	*-0.019**	-0.087	0.192	*-0.015*
University Degree	0.276**	0.001	*0.056***	0.225**	0.008	*0.044***	0.136	0.113	*0.025*
Pensioner	0.265**	0.041	*0.050***	0.301**	0.022	*0.056***	0.349**	0.008	*0.063***
Self-employed	-0.060	0.583	*-0.011*	-0.073	0.493	*-0.013*	-0.078	0.466	*-0.013*
Age	0.018	0.337	*-0,002*	0.017	0.340	*-0.002*	0.006	0.710	*-0.001*
Age2	-0.000	0.124		-0.000	0.121		-0.000	0.426	
Ln (Income)	0.602***	0.000	*0.112****	0.559***	0.000	*0102****	0.500***	0.000	*0.088****
Ln (Wealth)	0.143***	0.000	*0.027****	0.132***	0.001	*0.024****	0.115**	0.005	*0.020****
Advanced Literacy				0.444***	0.000	*0.087****	0.434***	0.000	*0.083****
Economic Degree							0.276**	0.039	*0.054**
Internet Usage							0.338**	0.029	*0.056***
Saving for major reasons							-0.096	0.394	*-0.016*
Saving for education							-0.068	0.419	*-0.012*
Saving for retirement							0.283**	0.002	*0.053***
High financial risk							0.093	0.697	*0.017*
Low financial risk							-0.378***	0.000	*-0.067****
Constant	-9.565***	0.000		-9.138***	0.000		-8.263***	0.000	
c	0,7984			0,8073			0,8232		
Observations	5899			5899			5899		

Average Marginal Effects

So far, it was possible to only comment about the sign and statistical significance of regressors since in a probit regression the coefficients cannot be interpreted directly as in the linear regression models, for this reason, it's fundamental to compute the average marginal effects. The marginal effects formally are defined as partial derivatives of the regression equation with respect to each variable in the model for each unit in the data. Average marginal effects are simply the mean of these unit-specific partial derivatives over some sample, they are a way of presenting results as difference in probabilities, which is more informative and easier to understand and interpretate than shifts in the standard cumulative normal function, which is the base of the probit

model. The above table also reports the average marginal effect (AME)[2] for each one of the regressors of each model.

Having said this, now the focus is put on the AME presented in TABLE IX. Regarding the model A, which accounts for the traditional variables, the education is the factor with the highest impact over the decision to hold stock, in fact, on average, individuals with a university degree are 5,6 percentage points more likely than those with a different education level to participate in the stock market. Conversely, on average the married peoples' probability of holding stocks is 2,6 percentage points lower than it is for those who are in a different marital status. In model B, with the introduction of the financial literacy variable, the statistical significance of some variables changes. Individuals who are knowledgeable about the basic concepts of interest, inflation and risk are 8,7 percentages points more likely to hold stocks than those who don't know about these concepts. In the last model, controlling for not only the traditional variables and financial knowledge, but also for additional factors such as reasons for saving and financial risk attituded. Financial literacy continues to be the non-economic factor with the highest effect over the market participation decision. Internet usage and saving for retirement also have a significant and positive impact, the probability of holding stocks is around 5 percentage points higher compared to an individual who does not use internet and save for another different reasons. It's important to highlight the economic degree indicator, it is more significant than having a degree in a different field, in fact, on average, the market participation probability of individuals who have a degree in economics or statistics is 3 percentage points higher than it is for those who possess another degree. On the other hand, strong financial risk-adverse people are 6.7% points less likely than individuals with a different risk aversion position to holding stocks.

Marginal effects at representative values

The objective of this research is to focus explicitly on the effect of a change in the educational status and financial literacy level on the likelihood of holding stocks. For this reason, another regression model is proposed, it is similar to model C but with some slight modifications, the age, income, and wealth are not taken like continuous variables as in model C, instead they are accounted as indicators, dichotomous variables for age groups and for quintiles of wealth and income. Additionally, the education and financial literacy variables are more specific, the new

---

[2] The AME are calculated with the margins **postestimation** command of Stata. They are interpreted as "On an all other things equal basis, a one-unit change in the independent variable is associated with an AME value percentage-point change in the probability of Participating in the Stock".

model accounts for every educational status and for each answer responded correctly to the financial literacy questionnaire, both are taken as categorical variables. The new model is reported in Annex 3.

The predictors with a high statistical significance are the same as in the previous models. Male, pensioner and saving for retirement conserve its significance level at 5%, however, the usage internet loses its predictive power. Regarding the education variable, the power prediction changes according to the group, the first education group: the lower secondary school, is significance at 10% compared with the other two education groups which are significance at 5% (like the university degree variable in the previous models). Thus, possessing a certain level of education significantly change the impact on the probability of holding stocks. As shown in Figure 7, the participation probability of individual with secondary school, upper secondary school, and university degree are, on average, 5.3, 5.8 and 10 percentage points higher than it is for people with primary school, respectively.
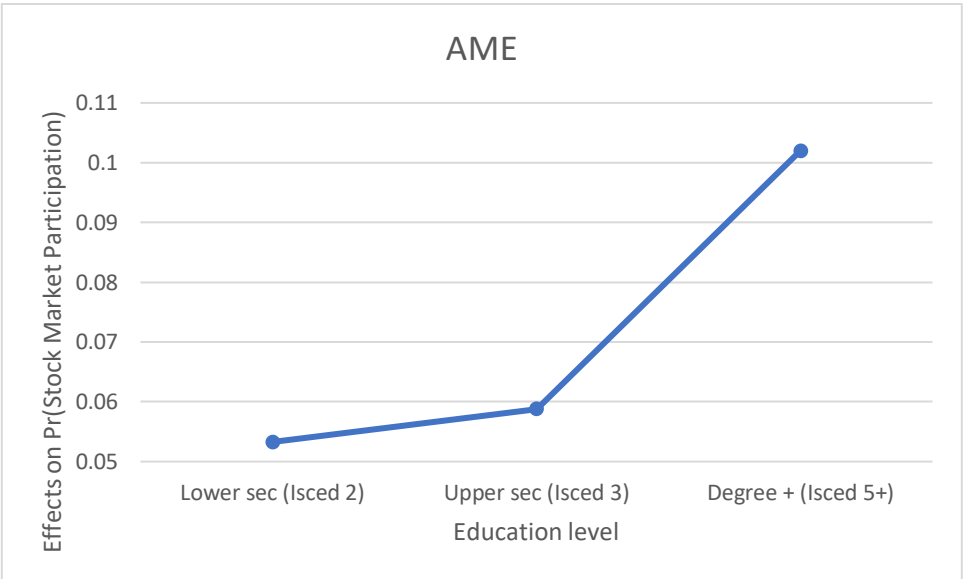
Figure 7 Average Marginal Effects of Education variable

On the other hand, the categorical variable that takes the number of correct answers is statistically significant only for value three, having responded to just one or two questions correctly is not predictive enough. Thus, only being completely knowledgeable about the financial concepts of interest, inflation, and risk, which means answers to the three questions correctly, is, on average,

9,6 percentage points more likely to participate in the stock market than those with partial financial knowledge.

The specification of this new model also allows to perform specific analyses for assessing the change in the probability of certain variables at specific values of other variables. Therefore, the marginal effects at representative values for the most significant predictors are presented. The AME of male is 3,7%, which means that on average the males' probability of holding stocks is 3,7 percentage points higher than it is for women. However, it's useful also to know how this probability varies with the education status (the representative values). Figure 8 shows the variation in the participating probability of male and pensioner with different levels of education. As expected, the effect in the probability of holding stocks increases with the level of education. Possessing the same level of education, the participating probability of a pensioner man is higher compared to the probability of a man without pension, indicating that the characteristic of enjoying a pension plan has more impact than the fact of being male or female over the holding stocks decision at specific level of education.
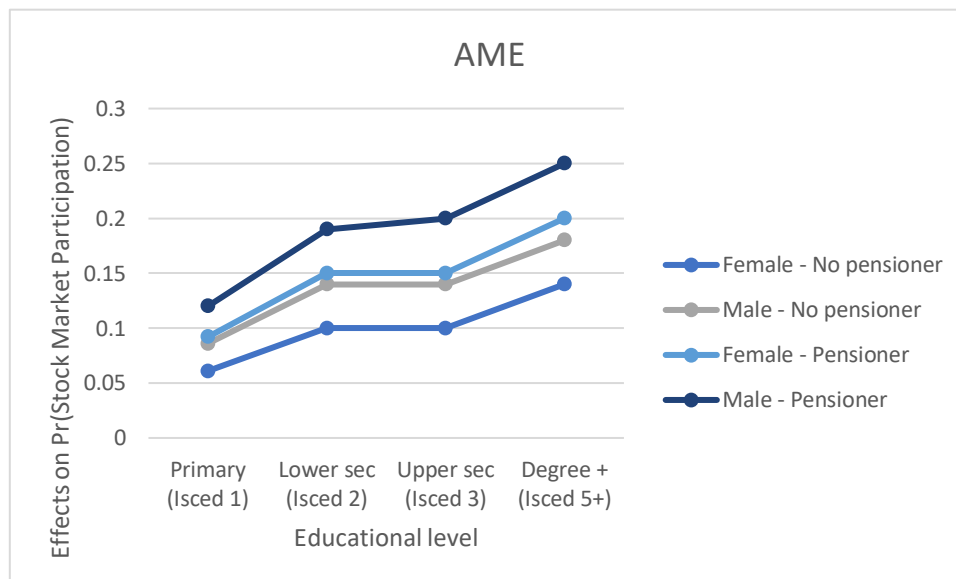
Figure 8 Average Marginal Effects of male and pensioned, at education levels

Another interesting analysis can be done over the effect that financial literacy variable has on the participating probability at different levels of financial risk attitude. Figure 9 shows that, even when a person has an advanced financial knowledge if he has a strong risk-adverse attitude, the likelihood of holding stocks is lower compared with those who also have a high financial literacy

but are less risk-adverse. Thus, the effect of financial knowledge differs greatly by risk attitude. It is less than 8 percentage points for risk-adverse and almost 13 percentage points for those less risk-adverse.
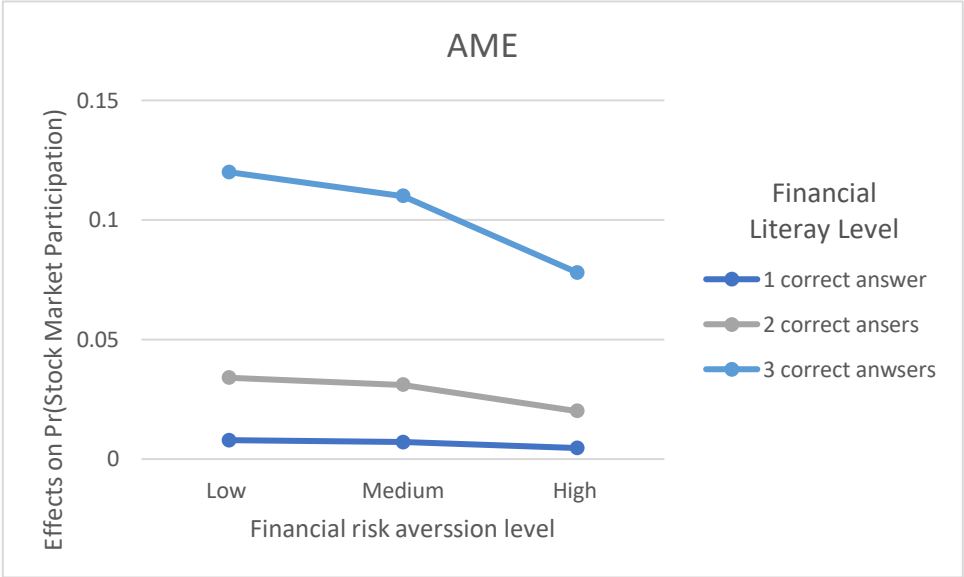
Figure 9 Average Marginal Effects of Financial literacy at Financial risk aversion level

Finally, the level of income and wealth are also some of the most significant predictors for the holding stock likelihood. Figure 10 presents the variation of the marginal effects of the quintile's income on the participating probability at different levels of financial literacy (1, 2 and 3 indicating the number of correct answers). The results show that even when the individual belongs to the last quintile, his likelihood of holding stocks varies a lot with his financial knowledge. The participating probability of highest-income people who respond correctly only one question is 11 percentage points higher than those with the lowest income, while for those who obtain all the three correct answers is almost 19 percentage points. Financial knowledge again proves to be a strong indicator over the decision to participate in the stock market.
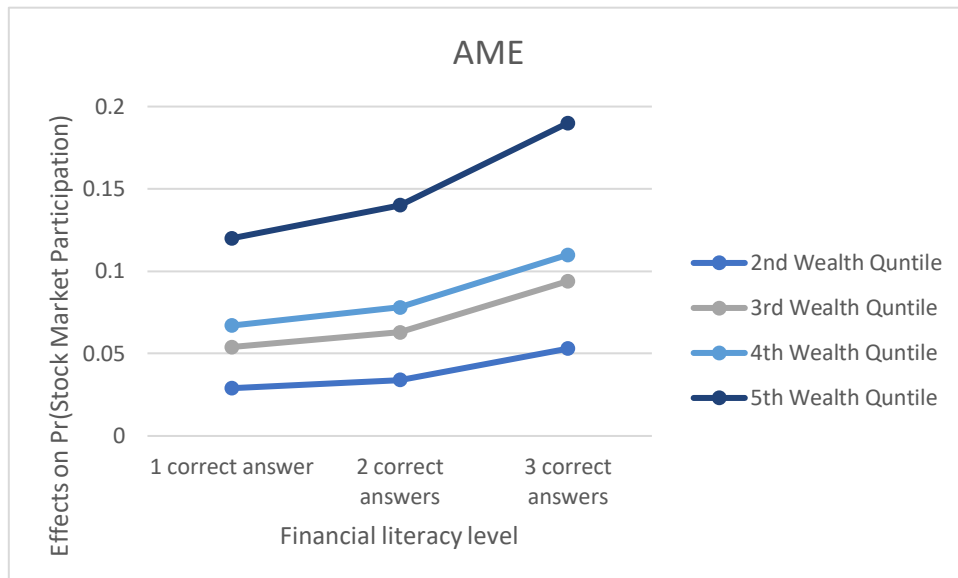
Figure 10 Average marginal effects of wealth quintile variables at Financial literacy level

Linear regression: asset share invested in stocks.

Having analyzed the choice to participate in the stock markets, the allocation choices of financial capital in stocks are studied. TABLE X reports the same structure used in the previous three models for the stock holding decision but this time for the ratio of value invested in stocks and investment funds over the total financial wealth, through a linear regression model. The basic Model A includes only the traditional variables, the Model B adds the financial literacy component, and the Model C incorporates additional characteristics including reasons for saving and risk aversion.

The results show that, similar to what happens in the participation, university degree, income, wealth and financial literacy are significant in explaining the risky share allocation of financial resources and are positively related. High financial risk aversion and being married are significant (except for the model B and C for the married variable) and have a negative effect on the share invested in stocks. Being male is significant at 10% when only traditional variables are considered and at 5% when the financial literacy characteristic is added. In the presence of additional factors, the gender variable loses its predictive power.

The household income and net wealth is always significant in the models and has a positive effect, the economic resources increase the investment in stocks or funds. Furthermore, unlike other studies the age in the Italian context, particularly for the sample of the SHIW, has not a significant impact over neither on the participation decision nor the share invested. On other hand, compared

to the stock market participation model, the fact of being pensioner loses its significance, occupational status seems not to be predictor for the amount invested in stocks, as well as the reasons for saving.

After having generally described the results of the regression models in terms of the sign regressors and the level of significance, the estimated effects will be analyzed from a quantitative point of view based on the Model C

In the regression model C, the asset share invested in stocks is 3,2 percentage points higher for investors with a university degree compared to individual with a different education. Particularly, the share of those with a degree in economics or statistics is 5,5 percentage points higher compared to individual who possess a degree in a different field. Regarding the household economic resources, it is estimated that a doubling of net disposable income corresponds to a 1.1 percent point increase in the share. Additionally, the financial literacy measure has always been a strong significant predictor for both the choice to participate and the share invested in stocks. The share of those who respond correctly to all the three questions of the financial module is 3 percentage points higher than those who don't know at least one answer. Finally, the asset share invested in stocks is 2,4 percentage points lower for low-risk tolerance investors compared to investors who are more tolerant with respect to financial risk.

TABLE X Linear regression model for share of wealth invested in stocks. One, two or three asterisks means that the predictor is significant at the ten, five or one percent level respectively. Models estimated considering weighted data.

	Model A		Model B		Model C	
	Coef.	p-value	Coef.	p-value	Coef.	p-value
Male	0.019**	0.041	0.016*	0.065	0.014	0.101
Married	-0.012**	0.049	-0.010	0.120	-0.009	0.168
University Degree	0.048***	0.000	0.044***	0.000	0.032**	0.018
Pensioner	0.014	0.438	0.016	0.368	0.018	0.317
Self-employed	-0.022	0.127	-0.023*	0.093	-0.026*	0.063
Age	0.001	0.779	0.001	0.790	0.000	0.867
Age2	-0.000	0.457	-0.000	0.463	-0.000	0.609
Ln (Income)	0.035***	0.000	0.032***	0.000	0.027***	0.000
Ln (Wealth)	0.012***	0.000	0.011***	0.000	0.010***	0.000
Advanced Literacy	-0.451***	0.000	0.032**	0.001	0.030**	0.001
Economic Degree					0.055**	0.018
Internet Usage					0.015	0.256
Saving for major reasons					-0.002	0.879
Saving for education					-0.001	0.938
Saving for retirement					0.016	0.157
High financial risk					0.047	0.272
Low financial risk					-0.024**	0.005
Constant	-0.470***	0.000	-0.433***	0.000	-0.357***	0.000
R-squared	0.0755		0.0810		0.0912	
Observations			5690		5690	

Conclusions

This study has contributed to the literature that attempts to look inside household financial decision-making by providing an insight into the choice of participation in the stock market and the decision about how much of the financial wealth to invest in the market. Using the 2020 cross-sectional data of the Survey on Households Income and Wealth from the Bank of Italy, particularly by exploiting the questions about financial literacy introduced in this survey from 2006, the financial literacy distribution in Italy population was investigated as well as its impact on the above financial decisions.

The empirical results show that most individuals lack knowledge of basic financial concepts, even if they appear to be more familiar with inflation and interest compounding than with risk. Consistent with findings from other countries, women, those with less education and the unemployed display the worst performance. While the individuals with university degrees and postgraduation perform best.

Additionally, regarding assets allocation, the household portfolios exhibit wide heterogeneity, lack of stock market participation and lack of diversification. Only less than one quarter of Italian households has at least one financial asset other than bank deposit or post office saving account, mostly in the form of managed investment (investment funds and managed assets), suggesting a potential gap in the allocation of financial resources. Furthermore, a key observation from this study reveals that approximately more than 80% of the total household wealth is concentrated in real estate assets. This high proportion highlights the critical relevance that real estate investment has for Italian families in the construction of their wealth and also reveals the lack of financial market integration as a means to increase their wealth.

As for the effects, financial literacy has a positive and significant impact on both dependent variables studied, this increases the probability of holding stocks. It was shown that even for households belonging to the same wealth quintile, the participating probability varies at lot depending on the financial knowledge level (measured by the number of correct answers). This variation reinforces the idea that if some agents are better able to gather and understand information about investment opportunities and stock markets, their effective costs of stock market participation will be lower and consequently they will have a higher probability of participating in the stock market. According to the estimations, belonging to the same poorest quintile, the participation probability of individual with an advanced financial knowledge is 5,3 percentage points higher that it is for those who don't know about finance, whereas the probability to those who respond correctly just one question is 2,9 percentage points.

Educational status also has a strong prediction power on the market participation decision. The likelihood of holding stocks increases with the level of education, the most educated people have a higher probability of participation. Likewise, gender, occupational status, and saving for retirement are positively linked to participation, while being a high risk-adverse individual has a negative significant impact on the decision to hold stocks, as well as it reduces the share invested in stocks.

References

Baglioni, A., Colombo, L., & Piccirilli, G. (2018). On the Anatomy of Financial Literacy in Italy. *Economic Notes. Monte Dei Paschi Di Siena*, *47*(2–3), 245–304. https://doi.org/10.1111/ecno.12111

Bertaut, C. C. (2016). *Stockholding Behavior of U . S . Households : Evidence from the 1983-1989 Survey of Consumer Finances Author (s): Carol C . Bertaut Source : The Review of Economics and Statistics , Vol . 80 , No . 2 (May , 1998), pp . 263-275 Published by : The MIT P. 80*(2), 263–275.

Christelis, D., Georgarakos, D., & Haliassos, M. (2009). *Stockholding : From Participation to Location*. *June*.

Christiansen, C., Joensen, J. S., & Rangvid, J. (2007). *Are Economists More Likely to Hold Stocks ?*

Cole, S., & Shastry, G. K. (2009). Smart money: The effect of education, cognitive ability, and financial literacy on financial market participation. *Boston, MA: Harvard Business School*, *February*, 09–071. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1317298

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2021). Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey. *SSRN Electronic Journal*, *1730*. https://doi.org/10.2139/ssrn.807408

Faiella, I. (2008). Accounting for sampling design in the SHIW. In *Temi di discussione del servizio Studi* (Vol. 35, Issue 442). http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4616.2010.00331.x/full

Faiella, I., & Gambacorta, R. (2007). The weighting process in the SHIW. In *Temi di discussione del servizio Studi* (Vol. 35, Issue 442). http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4616.2010.00331.x/full

Fornero, E., & Monticone, C. (2011). *Financial literacy and pension plan participation in Italy, Center for Research on Pensions and Welfare Policies, Working Paper*.

Guiso, L., Haliassos, M., & Jappelli, T. (2003). Household Stockholding in Europe: Where Do We Stand and Where Do We Go? In *SSRN Electronic Journal* (Issue May 2018). https://doi.org/10.2139/ssrn.346542

Guiso, L., & Japelli, T. (2005). *Household Portfolio Choice and Diversification Strategies*. 1–45.

Guiso, L., & Jappelli, T. (2009). *Financial Literacy and Portfolio Diversification* (Issue January).

Guiso, L., & Sodini, P. (2013). Household Finance: An Emerging Field. In *Handbook of the Economics of Finance* (Vol. 2, Issue PB). Elsevier B.V. https://doi.org/10.1016/B978-0-44-459406-8.00021-4

Haliassos, M., & Bertaut, C. C. (1995). Who do so few hold stocks? *Economic Journal*, *105*(432), 1110–1129. https://doi.org/10.2307/2235407

Hansen, M. H., Madow, W. G., & Tepping, B. J. (1983). An evaluation of model-dependent and probability-sampling inferences in sample surveys. *Journal of the American Statistical Association*, *78*(384), 776–793. https://doi.org/10.1080/01621459.1983.10477018

Heaton, J., & Lucas, D. (2016). *Portfolio Choice and Asset Prices : The Importance of Entrepreneurial Risk Author*. *55*(3), 1163–1198.

Johnston, D. W., Kassenboehmer, S. C., & Shields, M. A. (2016). Financial decision-making in the household: Exploring the importance of survey respondent, health, cognitive ability and personality. *Journal of Economic Behavior and Organization*, *132*, 42–61. https://doi.org/10.1016/j.jebo.2016.09.014

Kaustia, M., Conlin, A., & Luotonen, N. (2023). What drives stock market participation? The role of institutional, traditional, and behavioral factors. *Journal of Banking and Finance*, *148*, 106743. https://doi.org/10.1016/j.jbankfin.2022.106743

Laibson, D., Gabaix, X., Driscoll, J., & Agarwal, S. (2007). The Age of Reason: Financial Decisions Over the Lifecycle. *American Law & Economics Association Papers*, 1–52.

Lohr, S. (2022). Sampling Design and Analysis. In *Duxbury Press* (Vol. 185, Issue Supplement_2). https://doi.org/10.1111/rssa.12820

Love, D. A. (2010). The effects of marital status and children on savings and portfolio choice. In *Review of Financial Studies* (Vol. 23, Issue 1). https://doi.org/10.1093/rfs/hhp020

Lusardi, A., & Mitchell, O. S. (2007). Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, *54*(1), 205–224. https://doi.org/10.1016/j.jmoneco.2006.12.001

Lusardi, A., & Mitchell, O. S. (2008). Planning and Financial Literacy: How Do Women Fare? In *NBER Working Paper No. 13750*. http://www.nber.org/papers/w13750

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, *52*(1), 5–44. https://doi.org/10.1257/jel.52.1.5

Lusardi, A., & Mitchelli, O. (2007). Financial literacy and retirement preparedness: Evidence and implications for financial education. *Business Economics*, *42*(1), 35–44. https://doi.org/10.2145/20070104

Nicolini, G., & Haupt, M. (2019). *The Assessment of Financial Literacy : New Evidence from Europe*. *2009*, 1–20.

Noyan, F., & Şimşek, G. G. (2012). Tetrachoric Correlation as a Measure of default Correlation. *Procedia - Social and Behavioral Sciences*, *62*(2002), 1230–1234. https://doi.org/10.1016/j.sbspro.2012.09.210

STATA. (n.d.). *Jackknife estimation for survey data manual*. Https://Www.Stata.Com/Manuals/Svysvyjackknife.Pdf.

https://doi.org/10.4324/9781003279693-7

Survey on Italian Households Income and Wealth. (2020). In *Banca d'Italia*. https://doi.org/10.4324/9781003137818-49

van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449–472. https://doi.org/10.1016/j.jfineco.2011.03.006

Xu, Y., & Yao, R. (2022). Financial Decision-Making Responsibility and Household Wealth Accumulation Among Older Adults: A Comparative Advantage Perspective. *Journal of Financial Counseling and Planning*, *33*(1), 3–23. https://doi.org/10.1891/JFCP-19-00075

Annex 1 Definition of variables

OWNING STOCK OR MUTUAL FUND: This is the dependent variable of probit modes, it's a dummy variable equal to one if the households have investment funds, ETFs, shares in listed companies or shares in unlisted companies. Zero otherwise.

STOCKS AND INV FUNDS/FINANCIAL ASSETS: This is the continuous dependent variable of linear regression models. It's the ratio of the value of investment funds, ETFs, shares in listed and unlisted companies and total financial assets.

AGE: age of the reference person.

AGE2: squared of age of the reference person.

MALE: dummy variable equal to one if the gender of reference person was male, zero otherwise.

MARRIED: dummy variable that accounts for the marital status of the reference person, equal to one if he/she is married or in a civil relationship.

UNIVERSITY DEGREE: dummy variable equal to one if the reference person has one of the following educations: 3-year university degree/higher educational diploma, 5-year university degree or postgraduate qualification. Zero otherwise.

LOWER SEC(ISCED2): dummy variable equal to one if the reference person has one of the following educations: lower secondary school certificate or vocational secondary school diploma (3 years of study). Zero otherwise.

UPPER SEC (ISCED3): dummy variable equal to one if the reference person has a upper secondary school diploma, zero otherwise.

PENSIONER: dummy variable equal to one if the working status of reference person is pensioner, zero otherwise.

SELF-EMPLOYED: dummy variable equal to one if the working status of reference person is self-employed, zero otherwise

ECONOMIC DEGREE: dummy variable equal to one if the university degree of reference person is in economics or statistics, zero otherwise.

FINANCIAL LITERACY: dummy variable equal to one if the reference person responds correctly to all the three questions of the financial module from the survey. Zero otherwise.

INTERNET USAGE: dummy variable equal to one if the respondent reports that he/she or anyone in the household uses the internet or email, zero otherwise.

SAVING FOR MAJOR REASONS: dummy variable equal to one if the respondent reports that major purchases (residences, vehicles, furniture, etc.), investments, paying off debts was one of household most important reasons for saving, zero otherwise.

SAVING FOR EDUCATION: dummy variable equal to one if the respondent reports that education/economic support/legacy to children, grandchildren was one of household most important reasons for saving, zero otherwise.

SAVING FOR RETIREMENT: dummy variable equal to one if the respondent reports that old-age provision was one of household most important reasons for saving, zero otherwise.

INCOME: household net disposable income for the previous calendar year. Includes payroll income (net wages, salaries, and fringed benefits), pension and net transfers (pension, arrears, financial assistance, scholarships, alimony and gifts, and income pension scheme), net self-employment income, income from real estate (actual rents and imputed rents) and income from financial assets (interest on deposits, government securities and other securities except on stocks and investment funds, net of interest payments).

LN OF INCOME: natural logarithm of INCOME.

WEALTH: household net wealth for the previous year. Includes real assets (real estate, business equity and valuables) financial assets (deposits, government securities, other securities, trade credit or credit due from other households) net financial liabilities (liabilities to banks and other companies, trade deb and liabilities to other households).

LN OF WEALTH: natural logarithm of WEALTH

HIGH FINANCIAL RISK: dummy variable equal to one if the respondent report that prefers investments that offer very high returns, but with a high risk of losing part of the capital, zero otherwise.

NO FINANCIAL RISK: dummy variable equal to one if the respondent report that prefers investments that offer low returns, with no risk of losing the invested capital.

Annex 2: Breakdown of financial assets

DEPOSITS

- Bank or postal deposits (current, saving accounts or deposit books)
- Repos / PO savings certificates

MANAGED INVESTMENT SCHEMES

- Managed savings
- Funds or ETFs

EQUITY SHARES AND PARTICIPATING INTEREST

- Shares of listed companies (at their market value at end- 2020
- Shares in companies limited by shares – srl - and shares of unlisted companies (at their estimated realizable value at end 2020)

PRIVATE SECTOR BONDS

- Bonds issued by Italian firms
- Bonds issues by Italian banks

GOVERNMENT SECURITIES

- BOTs (T-bills)
- BTPs (T-bonds) / inflation-indexed BTPs (T-bonds
- Other (CTZs, CCTs et al.)

SECURITIES ISSUED ABROAD

- Foreign deposits
- Other foreign financial assets (Foreign government securities, Foreign bonds, foreign shares, etc..)

OTHER

- Foreign bonds, foreign shares, etc..)
- Loans to coperatives (social loans, etc.)
- Other financial assets (options, futures, royalties, etc.)

Annex 3: Probit model for stock market participation with all variables as categorical

Probit regression. One, two and three asterisks means that the predictor is significant at the ten, five or one percent level respectively. Models estimated considering weighted data, SHIW 2020.

	Coef.	P-value	Margin Effect	p-value
Male	0.212**	0.010	0.037**	0.009
Married	-0.064	0.339	-0.012	0.338
Education (ref: Primary)				
Lower sec (Isced 2)	0.346*	0.078	0.053*	0.056
Upper sec (Isced 3)	0.373**	0.045	0.058**	0.026
Unisersity degree	0.592**	0.001	0.102***	0.000
Pensioner	0.273**	0.040	0.050**	0.048
Self-employed	-0.109	0.276	-0.019	0.261
Financial literacy (ref: no correct answer or DK)				
1 correct answer	0.040	0.770	0.006	0.769
2 correct answers	0.164	0.225	0.026	0.205
3 correct answers	0.523***	0.000	0.096***	0.000
Internet Usage	0.231	0.162	0.040	0.134
Saving for major reasons	-0.097	0.391	-0.017	0.378
Saving for education	-0.047	0.573	-0.008	0.569
Saving for retirement	0.277**	0.002	0.052**	0.002
Financial risk (ref: medium risk)				
High financial risk	0.124	0.591	0.023	0.607
No financial risk	-0.398***	0.000	-0.072***	0.000
Age (ref: age<=30)				
Age 31-40	-0.214	0.307	-0.042	0.338
Age 41-50	-0.202	0.324	-0.040	0.358
Age 51-65	-0.219	0.307	-0.043	0.343
Age >=65	-0.301	0.287	-0.057	0.314
Income (ref: income<=15,911€)				
2nd Income quintile (15,911 - 23,915)	0.320	0.248	0.039	0.204
3rd Income quintile (23,949 - 34,003)	0.515**	0.041	0.071**	0.014
4th Income quintile (34,015 - 50,606)	0.601**	0.013	0.087***	0.001
5th Income quintile (50,615 - 2,264,763)	0.878***	0.000	0.147***	0.000
Wealth (ref: wealth <5,513€)				
2nd wealth quintile (5,513 - 110,500)	0.428**	0.041	0.054**	0.026
3rd wealth quintile 110,500-195,000)	0.372*	0.066	0.045**	0.034
4th wealth quintile (195,000-377,623)	0.685***	0.001	0.100***	0.000
5th wealth quintile (377,670 - 1,62e+08)	0.978***	0.000	0.167***	0.000
Constant	-2.906***	0.000		
Observations	5817		5817	

Appendix 1: Imputation and Replication

Financial surveys are subject to non-participation rates or inconsistency in revealed data. The most common reason for non-participation is the unwillingness on the part of the household. In about a quarter of cases the households cannot be contacted. Regarding the missing answers, these are due to the reticence on the part of the respondents, difficulties in replying to the question or in providing a timely response to variables such as fringe benefits of employees, the value of financial assets and some monetary information on occupational pension plans and insurance policies.

To reduce the effect of the missing values, SHIW staff imputed realistic value only to compute the aggregated variables (in a single dataset). Doing this by the regression imputation methos, which consist in using regression models on the basis of other available information. In order to avoid an excessive concentration around average values, a random component is added, extracted from a normal variable with a mean of zero and a variance equal to that of the residuals in the regression model. This preserves the mean and the variance of the data actually measured.

On the other hand, regarding the sampling variance as an important step to have accurate estimators, the Jackknife Repeated Replication (JRR)method is suggested, which takes into account the properties of sample design. As mentioned, the SHIW design uses a two-stage stratified sample with the stratification of the primary sampling units (municipalities) by region and demographic size and by income and indebtedness for the secondary sampling units (households). A graphic representation is below.

```
                        ┌─────────────────────┐
                        │    Sample Design    │
                        └─────────────────────┘
                 ┌───────────────┴───────────────┐
       ┌───────────────────┐         ┌───────────────────┐
       │     1st Stage      │         │     2nd Stage      │
       │  Municipalities    │         │    Households      │
       │ Primary Sampling Unit│       │ Secondary Sampling Unit│
       │       (PSU)        │         │                    │
       └───────────────────┘         └───────────────────┘
            ┌──────────────────┐          ┌──────────────────────┐
            │  Stratified by:  │          │    Stratified by:    │
            │    * Region      │          │  * Household Income  │
            │  * Population size│         │* Household Indebtedness│
            └──────────────────┘          └──────────────────────┘
```
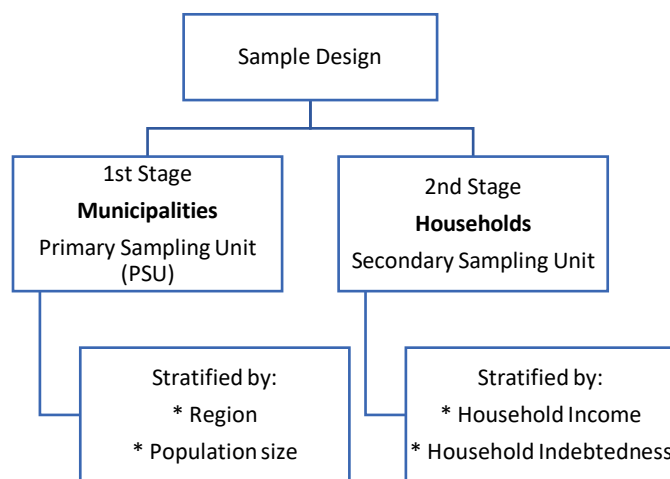
Figure 11 Sample design model

Within each stratum, the municipalities are selected by including all municipalities with a population of more than 40,000 inhabitants (self-representative units- SRU) and by random selection of smaller towns (No self-representative units - NSRUs) with probability proportional to the resident population.

Then in order to calculate the standard error for the SHIW a design must be chosen that is consistent with the sample unit selection process, and defined by the replication method, but does

not create computational difficulties. So, in the first place, all municipalities with more than 40,000 inhabitants (SRU) are placed in a separate stratum because they make no contribution to the randomization of the sample in the first stage. The sample households in each of these municipalities are then divided into two random groups. The remaining municipalities are grouped in their original stratum to give two municipalities per stratum. At the end of this process there are about 350 "pseudo" strata containing two "pseudo" first-stage units.

The JRR variance is finally calculated using the following steps:

 1. the number c of replications is equal to the number of "pseudo" strata

2. in each replicate the weight of the first "pseudo" primary sampling unit is set equal to zero and the sampling weight of the other is raised by a factor to compensate the weight of the cancelled unit

3. this weight is used to calculate, for each replicate, the relevant estimators $\hat{\theta}_{(i)}$

4. since the design for variance estimation contains two units per stratum, the estimate of the standard error is calculated as the square root of the sum of the square deviations between the estimate of the replications and the estimate on the total sample $\hat{\theta}_{(i)}$

$$stderr_j = \sqrt{\sum_{i=1}^{c} (\hat{\theta}_{(i)} - \hat{\theta})^2}$$

A useful way of assessing how far the estimation variance calculated to take account of the complexity of the sample design diverges from the one assuming simple random sampling is to measure the ratio between the two: for the generic estimator ˆθ the design effect (deff) is:

$$deff(\hat{\theta}) = \frac{V(\hat{\theta})_{compl}}{V(\hat{\theta})_{ccs}}$$

Appendix 2 Sampling weight

Starting from the 2020 survey, sampling weight construction was revised to adjust for the introduction of household stratification in the second stage of survey design. More specifically, to make the sample more representative across the income distribution, non-panel households were stratified in ten household income groups per geographical macro-area (North-East, North-West, Centre, South and Islands). Moreover, in order to improve the usability of the survey for studying financial vulnerability, a sample of indebted households was added to the selected sample and was stratified according to five debt size groups for each type of debt (non-performing and performing loans). As a result, for the 2020 survey the new sample design weights were calculated as follows:

a) The design weight $w_{hik}^{(0)}$ is calculated by multiplying the inverse of the probability of selecting municipality i of stratum h (i.e. the weight of first-stage units) and the inverse of the probability of selecting a stratum k household (second-stage stratification) residing in municipality i of stratum h (i.e. the weight of second-stage units).

$$w_{hik}^{(0)} = \left(\frac{1}{m_h} \frac{P_h}{P_{hi}} \right) \frac{N_{hik}}{n'_{hik}} \qquad h = 1 \dots H, \; k = 1 \dots K$$

where P_h and m_h are the resident population and the number of sample municipalities in stratum h (first stage), respectively. P_{hi} is the resident population in municipality i of stratum h. N_{hik} and n'_{hik} are the number of resident households and the number of selected households (theoretical sample), respectively, in municipality i of first-stage stratum h belonging to second-stage stratum k.

b) The adjustment for total non-response $w_k^{(1)}$ is obtained by multiplying $w_{ik}^{(0)}$ by the inverse of the response rate of stratum k to which each household belongs

$$w_k^{(1)} = w_{hik}^{(0)} \frac{n'_k}{n_k}$$

Where n'_k and n_k are total selected households (theoretical sample) and total respondents (actual sample) in second-stage stratum k.

c) The weight $w_k^{(1)}$ is adjusted for panel attrition (i.e. non-response from household units who participated in previous surveys) and to replicate the panel's optimal share, estimated at approximately 50 per cent of the sample, resulting in the weight $w^{(2)}$

d) Lastly, $w^{(2)}$ is adjusted based on additional information (calibration), using external data that are correlated to key economic variables in order to improve estimator accuracy. More specifically, weights are adjusted to replicate the population demographics in terms

of gender, age (seven classes), geographical area (three classes), municipality size (four classes), level of education (two classes) and household composition (five classes), resulting in the final weight $w_j^{(3)}$

$$w_j^{(3)} = w^{(2)} \gamma_j$$

where γ_j is the adjustment factor for class j of the stratification variable γ.

On the other hand, sample design variables (i.e. stratum and clusters ID) are often related to geographical information (e.g. the IDs of the municipalities) that is usually not disseminated in the public dataset due to confidentiality. For this reason, SHIW staff have disseminated replication weights, thus avoiding the inclusion of design variables in the public dataset to balance confidentiality protection and the users' possibility to properly compute variance.

In order to generate replication weights the sample is broken up into subsamples, called replicates. In the case of the SHIW, following the variance computation model design devised and applying the JRR method, the replication weights are computed as follows. For each of the 350 "pseudo" strata (h) with two "pseudo" PSU (j) the replicated weight is:

$$w_{hji}^r \begin{cases} 0; & \text{if the } i - th \text{ belongs to PSU } 1 \\ w_{h2i} \dfrac{\sum_{j=1}^{2} \sum_{i=1}^{n_j} w_{jih}}{\sum_{j=1}^{2} \sum_{i=1}^{n_j} w_{jih} - \sum_{i=1}^{n_1} w_{h1i}}; & \text{if the } i - h \text{ belongs to PSU } 2 \\ w_{hji}; & \text{for all the other units} \end{cases}$$