

Universidade do Minho
Escola de Economia e Gestão

**The impact of financial literacy and behavioral biases
on financial behavior**

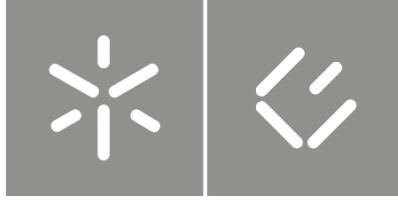
Maria Helena Santos Pereira

**The impact of financial literacy and
behavioral biases on financial
behavior**

Maria Helena Pereira

UMinho | 2022

abril de 2022



Universidade do Minho
Escola de Economia e Gestão

Maria Helena Santos Pereira

**The impact of financial literacy and
behavioral biases on financial behavior**

Dissertação de Mestrado
Mestrado em Gestão e Negócios

Trabalho efetuado sob a orientação da
Professora Doutora Cristiana Cerqueira Leal

abril de 2022

Direitos de autor e condições de utilização do trabalho por terceiros

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizados deste trabalho



Atribuição-NãoComercial-SemDerivações

CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Acknowledgements

Firstly, I would like to thank professor and supervisor Cristiana Cerqueira Leal, for all the help, availability, support and patience with all the questions and doubts, as well with all the knowledge shared. I am very grateful for my family and friends, that throughout this journey pushed me and motivated me the most. Without them, certainly this journey would have not been so pleasant.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

O impacto da literacia financeira e enviesamentos comportamentais no comportamento financeiro

Resumo: Esta dissertação estuda o comportamento financeiro dos indivíduos portugueses, nomeadamente a sua participação no mercado e a tendência para o efeito disposição (preferência para vender ativos com ganhos e manter em carteira ativos com perdas). Para tal, investiga de que forma a literacia financeira e enviesamentos comportamentais (aversão a perdas e excesso de confiança) afetam esses comportamentos. Primeiramente é explicada a literacia financeira e os seus determinantes. Posteriormente é estudado o comportamento financeiro (participação no mercado e efeito disposição), tal como outras dimensões do comportamento como é o caso da composição da carteira e a disparidade entre as preferências declaradas e o comportamento real. Para tal, são usadas como variáveis explicativas a experiência, importância dos valores mobiliários no património total, perfil de risco, aspetos cognitivos e enviesamentos comportamentais, tal como variáveis sociodemográficas, sendo que no caso dos comportamentos financeiros é também utilizada uma medida de literacia financeira. Por último é observado de que forma os resultados mudam em função da medida de literacia financeira adotada.

Os dados utilizados foram recolhidos pela CMVM – Comissão de Mercado de Valores Mobiliários em 2021 através de um questionário online respondido por 1850 indivíduos portugueses com mais de 18 anos. Para a análise são usadas regressões probit e probit ordenado dada a natureza das variáveis dependentes. Os resultados sugerem que a literacia financeira apresenta uma relação positiva com experiência, importância do investimento, enquanto ambos os enviesamentos comportamentais (aversão a perdas e excesso de confiança) apresentam uma relação negativa com a literacia financeira. Os indivíduos do sexo masculino são mais propensos a ter maior literacia financeira (mesmo quando contabilizadas as diferenças de excesso de confiança entre géneros). Existe uma maior disponibilidade a participar entre os indivíduos que são mais instruídos financeiramente e mais confiantes, demonstrando assim a importância da perceção que o indivíduo tem das suas capacidades. Quando observado o nível de diversificação da carteira, indivíduos com maior literacia financeira são também mais prováveis de apresentar maiores níveis de diversificação. Por último, o excesso de confiança, atua como um dissuasor ao envolvimento do investidor no efeito disposição (contudo ao utilizar-se uma medida alternativa para o efeito disposição isto deixa de ser observável). É também testado o impacto de diferentes medidas de literacia financeira nos resultados, sendo que a magnitude dos efeitos diminui consideravelmente aquando da inclusão de uma nova dimensão na definição de literacia financeira.

Palavras-Chave: Literacia Financeira; Comportamento Financeiro; Enviesamentos Comportamentais; Participação no Mercado; Efeito Disposição

The impact of financial literacy and behavioral biases on financial behavior

Abstract: This dissertation studies the financial behavior of Portuguese individuals, namely their participation in the market and the tendency for the disposition effect (preference to sell assets with gains and keep in portfolio assets with losses). To this end, it investigates how financial literacy and behavioral biases (loss aversion and overconfidence) affect these behaviors. Firstly, financial literacy and its determinants are explained. Subsequently, financial behavior (market participation and disposition effect) is studied, as well as other dimensions of behavior, such as portfolio composition and the gap between stated preferences and actual behavior. Therefore, the explanatory variables used are experience, the importance of securities in total wealth, risk profile, cognitive aspects, and behavioral biases, as well as socio-demographic variables, and in the case of financial behaviors, a measure of financial literacy is also used. Finally, it is observed how the results change depending on the financial literacy measure adopted.

The data used was collected by CMVM - Comissão de Mercado de Valores Mobiliários in 2021 through an online questionnaire answered by 1850 Portuguese individuals over 18 years old. Probit and ordered probit regressions are used for the analysis given the nature of the dependent variables. The results suggest that financial literacy shows a positive relationship with experience, importance of investment, while both behavioral biases (loss aversion and overconfidence) show a negative relationship with financial literacy. Males are more likely to have higher financial literacy (even when accounting for gender differences in overconfidence). There is a greater willingness to participate among individuals who are more financially literate and more confident, thus demonstrating the importance of the individual's perception of his or her abilities. When the level of portfolio diversification is observed, individuals with higher financial literacy are also more likely of displaying higher levels of diversification. Finally, overconfidence acts as a deterrent to investor's involvement in disposition effect (however, by using an alternative measure for the disposition effect, this is no longer observable). The impact of different measures of financial literacy on the results is also tested, and the magnitude of the effects decreases considerably when a new dimension is included in the definition of financial literacy.

Keywords: Behavioral Biases; Disposition effect; Financial Literacy; Market Participation

Index

1. Introduction	1
1.1 Relevance and overview	1
1.2 Main goals	1
1.3 Work Structure	2
2. Review of the literature	3
2.1 Introduction to Financial Literacy	3
2.1.1 The Benefits of Financial Literacy	3
2.1.2 Definition of Financial Literacy and related constructs	4
2.1.3 Determinants of Financial Literacy	6
2.1.4 Financial literacy in Portugal	10
2.2 Behavioral Aspects	11
2.2.1 Overconfidence	12
2.2.2 Loss aversion	13
2.3 Market Participation	14
2.3.1 Financial literacy and Market Participation	14
2.3.1.1 Portfolio Composition	14
2.3.2 Stock Market Participation	15
2.3.2.1 Financial Literacy and risk profile	15
2.3.2.2 Cognitive and Behavioral aspects (loss aversion and overconfidence)	15
2.3.2.3 Sociodemographic factors (gender, age, income, education)	15
2.4 Disposition effect	16
2.4.1 Disposition effect determinants	17
3. Methodology	19
3.1 Method	19
3.2 Hypotheses and models	19
3.3 Questionnaire design	24
4. Data	28
4.1 Data Collection	28
4.2 Data Treatment	28
4.2.1 Variables and open-answer questions	29
4.2.2 Inconsistencies	32

4.3 Descriptive Analysis	33
4.4. Internal Consistency analysis	40
4.5. Multicollinearity analysis	41
5. Empirical Results.....	45
5.1 Financial Literacy.....	45
5.1.1 Interaction term between gender and overconfidence	49
5.2 Market Participation.....	51
5.3 Portfolio composition	53
5.4 Gap between hypothetical and current portfolio.....	56
5.5 Disposition effect among investors.....	60
5.5.1 Disposition effect with reference point	60
5.5.2 Disposition effect without reference point	63
5.6 Robustness checks: Differences with different financial literacy measures	64
6. Conclusion	71
6.1 Limitations and avenues for further research	73
References	75
Appendix	81

Index Table

Table 1: Summary description of variables	31
Table 2: Summary statistics of age	33
Table 3: Summary Statistics regarding sociodemographic (gender, education, occupation, income, metro), risk profile and study field	34
Table 4: Summary Statistics regarding investors, product ownership, reasons to invest and not to invest , which facts made them start investing or would make them invest and experience level	36
Table 5: Summary Statistics regarding financial literacy, cognitive reflection test and numeracy.....	38
Table 6: Frequency regarding the level of objective financial literacy and self-perceived, cognitive reflection test, and experience	39
Table 7: Summary table regarding the percentage of Portuguese individuals who know financial questions across different studies.....	40
Table 8: Pearson Correlation Matrix	43
Table 9: Frequency distribution of finknow	45
Table 10: Marginal effects of ordered probit with finknow as dependent variable	47
Table 11: Marginal effects of ordered probit with finknow as dependent variable	48
Table 12: Marginal effects of probit with marketpart as dependent variable	52
Table 13: Frequency distribution of the number of different financial product	54
Table 14: Marginal effects of probit with portf as dependent variable – reduced model.....	54
Table 15: Marginal effects of probit with portf as dependent variable	55
Table 16: Summary Statistics of the components of the hypothetical portfolio	57
Table 17: Summary Statistics of pgap	57
Table 18: Marginal effects of ordered probit with pgap as dependent variable – reduced model	58
Table 19: Marginal effects of ordered probit with pgap as dependent variable	59
Table 20: Frequency regarding the questions of de (as it is the case of losses and gains)	61
Table 21: Marginal effects of probit with de as dependent variable	62
Table 22: Frequencies regarding disposition effect with (de) and without reference point (de1).....	63
Table 23: Frequencies regarding alternative measures of financial literacy – finbehav and perceived	64
Table 24: Summary results regarding alternative measures of financial literacy	70
Table A.1: Variance Inflation Factor	81
Table A.2: Marginal effects of the interaction term between gender and overc with finknow as dependent variable for each outcome	81
Table A.3: Marginal effects of ordered probit with finbehav as dependent variable	83
Table A.4: Marginal effects of the interaction term between gender and overc with finbehav as dependent variable for each outcome	84

Table A.5: Marginal effects of ordered probit with perceived as dependent variable	84
Table A.6: Marginal effects of the interaction term between gender and overc with perceived as dependent variable for each outcome	85
Table A.7: Marginal effects of ordered probit with marketpart as dependent variable [use of perceived and finbehav as independent variables]	86
Table A.8: Marginal effects of ordered probit with portf as dependent variables [use of finbehav as independent variable]	87
Table A.9: Marginal effects of ordered probit with portf as dependent variables [use of perceived as independent variable]	88
Table A.10: Marginal effects of ordered probit with pgap as dependent variables [use of finbehav as independent variable]	89
Table A.11: Marginal effects of ordered probit with pgap as dependent variables [use of perceived as independent variable]	90
Table A.12: Marginal effects of probit with de as dependent variables [use of perceived and finbehav separately as independent variable]	91
Table A.13: Marginal effects of probit with de1 as dependent variables [use of finknow, perceived and finbehav separately as independent variable]	93

1. Introduction

1.1 Relevance and overview

Every day, each one of us encounters many choices regarding multiple dimensions of our life, that may relate to the basic daily situations or even complex decisions such as financial aspects, that are in part influenced by the degree of financial literacy, behavioral biases among many other factors. Financial literacy among several advantages administers mechanisms that influence positively the willingness to future savings and better household budget management, contributes to a more in-depth knowledge of financial risks and avoids the occurrence of over-indebtedness and excessive interest rates in some financial products, as it is stated in Banco de Portugal (2011). The financial literacy level of the majority of the population is quite low, inclusively it may not be enough to secure reliable financial decisions, which translates into common and dangerous behaviors such as enormous household debt, smaller participation in savings, wrong perception of risk diversification, and unconscious and alarming misallocation of capital (Calgano & Monticone, 2015; Georgarakos & Inderst, 2011; Lusardi & Mitchell, 2011; Lusardi & Mitchell, 2014; Stolper, 2018). This lack of financial literacy and the emergence of new financial products or instruments with escalated complexity (as mentioned by Van Rooij et al., 2011) creates a hazardous cocktail.

Throughout the last decades, behavioral economics and finance have emerged, bringing a new approach over some of the main assumptions regarding traditional views of decision-making behaviors. The full rationality assumption is contradicted by some of the heuristics used by individuals since these biases are imperfect and may lead to sub-optimal decisions (Leal & Oliveira, 2021; Shefrin, 2002).

1.2 Main goals

This research intends to draw a better understanding of financial literacy and financial behaviors (as it is the case of market participation and disposition effect) in the Portuguese population. A higher understanding allows for more suited practices to enhance healthy and optimal financial decisions. Therefore, at first instance, it dives into financial literacy in order to identify its determinants. Afterward it studies the propensity to participate in capital markets, and as a sub-set it also investigates portfolio diversification and the gap between individual's preferences and actual behavior. Lastly, it studies disposition effect in investors in order to try to understand which factors affect the occurrence of this bias. Since there is no consensual definition of financial literacy, in the end it is also tested how the results of this work alter depending on the measure used. To better conduct this research and acknowledging that a quantitative research strategy is used, there are general goals that are aimed, namely:

- Establish financial literacy determinants
- Explain financial behavior (market participation, portfolio diversification, gap between investor's preferences and actual behavior in portfolio composition and investor's disposition effect) through:
 - Financial literacy
 - Behavioral biases (loss aversion and overconfidence)
 - Demographic characteristics
- Understand how the results change depending on the adopted measure of financial literacy

In order to fulfil these goals, this research uses secondary data retrieved by CMVM in 2021 through an online self-completion questionnaire on the Portuguese population. Moreover, this sample is composed by 1850 inquiries with a minimum age of 18 years old. In accordance with the written hypotheses, several regressions are held to test if there is evidence able to reject or not the initial statement. Thus, probit and ordered probit regressions are applied, since the dependent variables are either binary or categorical.

1.3 Work Structure

This dissertation is constituted by six sections. This section presents a short introduction and in the following section it is the literature review, where a state of the art is presented through different author's contributions and lenses to the discussion of financial literacy and financial behavior. Then the section of methodology explains which research strategy and technique to collect data is used, along with hypothesis to be tested. Afterwards, it is the data section that helps describing the sample. In the fifth section are the empirical results, where the hypothesis previously held are tested and discussed. Finally, the conclusions of this dissertation are presented, as well as possible limitations faced and future research.

2. Review of the literature

2.1 Introduction to Financial Literacy

Financial literacy, or the enhancement of this concept, is a worldwide ambition shared by most countries, that gains even more emphasis when considering disconcerting situations as excessive overindebtedness by households, the subprime crisis, low rates of saving, financial scandals that rock confidence in the financial market and misapplication of savings, among others. Financial choices have increased its complexity, however the hazards of it have followed the same trend - the risks of financial choices tend to enlighten society, especially in cases of crisis, that until that point, sub-optimal decisions passed unseen (Hung et al., 2009). In the Portuguese case, this concept has become widely known, especially since the international crisis of 2008, which not only hampered the economic situation, it also made households revise their financial choices. Acknowledging the negative results of the lack of financial literacy, there is still no “one-fits-all” solution that has been able to disseminate this phenomenon - despite several policies taken and state interventions to improve financial literacy, financial policies to affect positively financial literacy appear to only justify 0.1% in the variation of financial behaviors, inclusively in individuals with lower income the effects are even smaller (Fernandes et al., 2014). Thus, new measures, perspectives and tools are needed to mitigate it, since not only it jeopardizes the process of decision-making of individuals, moving them into sub-optimal choices as mentioned by Calcagno and Monticone (2015), but it also has the destructive power to aggregately shatter an economy and disable or retard its recovery.

2.1.1 The Benefits of Financial Literacy

Financial literacy becomes even more meaningful and essential when handling economic or financial shocks that rock any economy, since those that present stronger financial literacy were less likely to encounter themselves with unfavorable income distress and present higher amounts of savings even in a moment of crisis (Klapper et al., 2013). Hung et al. (2009) affirm that individuals with lower levels of literacy are presumed to be more probable to incur in financial errors, dubious or inadvisable practices, but are also more likely to poorly handle economic shocks. Reversely, when financial literacy is correctly used and thought of as being part of the human capital of the individual, it may foster the length of the positive utility derived from consumption (Huston, 2010). Additionally, it provides tools that affect positively the prospective of future savings, provides a better personal and household budget management, mitigates the possibility of engaging in unauthorized techniques, contributes to a more in-

depth knowledge of financial risks and avoid the possibility of over-indebtedness and excessive interest rates in some financial products, as it is stated in Banco de Portugal (2011). Individuals that are financially knowledgeable not only contribute to their own financial health and wealth, but they also have an important role in the development of the society that they are integrated (Hilgert et al., 2003). Klapper et al. (2013) also affirm that it can stabilize economic parameters and indicators even when facing macroeconomics disrupters, which can be even more relevant in economies that do not have a very stable economic cycle and that its economy easily reacts to these shocks negatively.

2.1.2 Definition of Financial Literacy and related constructs

To deeply understand financial literacy, firstly one have to be able to define and measure it. However, there is no consensus or universal measure when it comes to financial literacy, thus the on-going debate persists derived from different views of definition and measurement, that hamper the comparison of results and possible solutions that could be drawn from those (Hung et al., 2009). Furthermore, derived from the vastness of different definitions adopted, the results of these researches may also be endangered (Remund, 2010). Inclusively, Martins et al. (2021) state that as new dimensions are added to the construct of financial literacy, it appears to exist more significative differences depending on which measure is used.

Thus, financial literacy among various definitions comprises different perspectives and concepts in each one of them - Hung et al. (2009) discloses these variations in definition, stating that some definitions focus either on knowledge, on the capacity to apply this expertise, on the knowledge perceived by the individual, on optimal behavior or even experiences. The most frequent definition uses only knowledge/understanding, namely basic financial concepts as it is the case of compounded interest, nominal and real values and diversification (these tend to use “the big three” questionnaire to evaluate financial literacy, that was introduced by Lusardi & Mitchell, 2011). This metric of financial literacy has not been fully tested and may be biased and influenced by other factors e.g numeracy, that also affect financial outcomes, as Schmeiser and Seligman (2013) point out.

According to Huston (2010), many empirical research used financial literacy, knowledge, and education as synonyms, even though these are distinctive. Thus most of the research up to 2010 did not define financial literacy, and those that did, regularly concentrated only either in ability or knowledge, rarely linking these two. Huston (2010) specifies that if the application element is not present, that is, if the individual cannot use his knowledge in practical financial decisions, then we are discussing financial knowledge rather than financial literacy. However, if the individual lacks specific skills, that may

influence his literacy, nonetheless, it can be overcome with the usage of certain tools. When measuring financial literacy, it is expected to include either the knowledge component, but also the application one – when focusing only in measurement of knowledge and ability and cognition, not addressing both these concepts, it will alter the effectiveness of financial literacy projects (Van Rooij et al., 2011). OECD (2019) assumes an additional component into a three dimensional definition, comprising knowledge, as well as behaviors and attitudes. Consequently, financial knowledge is not enough by itself to conceive financial literacy if there is no “mobilisation of cognitive and practical skills, and other resources such as attitudes, motivation, and values” (p.41), which implies that the theoretical knowledge of financial concepts shall be associated and followed by mathematical, reading and interpreting skills, as well as an active attitude towards financial matters. The combination of self-perceived and actual knowledge as a measure for financial literacy seems to explain better the relation of financial literacy and behavior, than if test scores were only evaluated. Thus using test score measures to evaluate financial literacy and subsequent financial behaviors may underestimate the impact of financial literacy in behavior (Allgood & Walstad, 2016).

For instance, the Portuguese case illustrates adequately this difference as it is observable in Plano Nacional de Formação Financeira (2020), that is based on an international questionnaire of financial literacy developed by the *International Network on Financial Education*. In the Portuguese case, these components (knowledge, attitudes and behavior) perform differently. Even though attitudes and financial behavior set Portugal in the 5th place, financial knowledge presents itself as a weaker point situating in 16th place, demonstrating that these may have different scenarios, thus if financial literacy was observed together, the components would hide other important dimensions.

Additionally, OECD (2019) presents an interesting perspective, stating that literacy can and should be introduced at younger ages, but also be seen as a life-cycle learning process since it should not be thought of as a barrier to be overpassed as if there was a fixed and optimal quantity to be acquired. Jappelli and Padula (2013) demonstrate to us that the choices of consumption and saving of an individual, according to the intertemporal model, will depend on the perception of future benefits. Even though there is a positive relationship between investments in literacy and returns, it is accompanied by not only monetary resources, but also timewise and commitment costs. As with most investments, it depreciates throughout time and it has a cost associated. The depreciation will be different between individuals, derived from different characteristics and the initial endowment.

As it was mentioned, there is a connection between the application and the accumulation of financial knowledge, thus financial literacy influences decision-making and according to Calgano and Monticone

(2015), when financial literacy is not observable that may call into question the reliability of financial decisions that will tend towards sub-optimal choices – lower levels of financial literacy are related with a higher possibility of making financial decisions non-optimal.

2.1.3 Determinants of Financial Literacy

Financial Literacy is highly influenced by socio-demographic and economic factors as it is presented in nearly all literature, inclusively some researchers use these factors as proxies for the evaluation of financial literacy (e.g it is the case of Dhar and Zhu (2006), stating that those variables are proxies since there is a correlation between better knowledge of markets and accessibility to information. Therefore, in the following section, it will be presented factors as age, gender, level of education and income that, not only affect the initial endowment and investment in literacy but in some cases, it is even possible to almost create a picture with sociodemographic information of groups with lower levels of literacy. Other factors as numeracy and cognitive reflection test, experience, and risk profile appear to also be linked with financial literacy.

2.1.3.1 Age

When analyzing socio-demographic factors, age affects literacy in a non-linear way and according to Lusardi and Mitchell (2011), it follows a U-shaped inverted pattern. This means that younger individuals will demonstrate lower levels of literacy, but as they age it will rise, reaching its peak amidst their life and after that, it will start decreasing once again, displaying a lower level at old age. This conclusion is in accordance with most of the research that argues that youngsters and elders exhibit the lowest levels, and middle-aged individuals have better results (Abreu & Mendes, 2010; Hung et al., 2009; Van Rooij et al., 2011).

2.1.3.2 Gender

The differences in gender financial literacy patterned in literature – having the man upper advantage in this matter (Abreu & Mendes, 2010; Hung et al., 2009; Lusardi & Mitchell, 2011; OECD, 2016; Martins et al., 2021; Sebastião et al., 2021; Potrich et al., 2015; Van Rooij et al., 2011) – according to Hsu (2011), can be explained by household distribution, being the man associated with financial management matters. Interestingly, since the woman tendentially lives longer, as she approaches widowhood and starts taking over the financial matters, its financial literacy rises, inclusively at most of the times catching up to 80% of the levels of financial literacy of the man before widowhood.

Moreover, Lusardi and Mitchell (2011), show that women are also more susceptible to answer “ I do not know” when compared with men, and this finding is repeatable throughout the globe. Bucher-Koenen et al. (2021) also find this higher susceptibility from women, acknowledging that when this option is missing, women frequently pick the right option. Indeed, these authors state that nearly 38% of the divergence of financial literacy between genders may be justified by lack of confidence from women, thus women do present lower levels of knowledge when taking into consideration confidence, however women are more knowledgeable than what they may think of themselves. Even though there is still a disparity between genders, it decreases considerably when removing the option. Munóz-Murillo et al. (2020), contradicts the vast literature, stating that gender does not explain financial literacy once cognitive ability variables are controlled.

2.1.3.3 Education and study field

When it comes to the education level, higher levels of education are associated with a higher level of financial literacy (Abreu & Mendes, 2010; Disney & Gathergood, 2013; Hung et al., 2009; Lusardi & Mitchell, 2011; Sebastião et al., 2021). However, Lusardi and Mitchell (2011) highlight that even when individuals have high levels of education, financial literacy may still be low and furthermore it is stated that this variable should not be used as a proxy for financial literacy.

The study field also impacts financial literacy as Sebastião et al. (2021) state that individuals who present as study field an area related to finance present a higher level of financial literacy (objective knowledge), moreover they also believe that their financial knowledge is higher than what they truly own. Van Rooij et al. (2011) also demonstrate the impact of the study field, in particular individuals who studied economy, presented higher levels of literacy in the long-run.

2.1.3.3 Income

Then, income does also present to be related to financial literacy, namely lower levels of income are accompanied by lower levels of financial literacy (Abreu & Mendes, 2010; Disney & Gathergood, 2013; Hung et al., 2009; Sebastião et al., 2021). Inclusive, Potrich et al. (2015) specifies that an additional level of income increases by 6,32% the likelihood of entering the group with high financial literacy. Family income also follows the same direction with an increase of 3,73% of reaching that group. These previous authors go further stating that income is considered one of the most important variables when assessing and clarifying financial literacy. In contrast, the measure of financial literacy used alters the impact of income, and income only appears as a positive determinant being statistically significant,

only when envisioning financial literacy as a two-dimension concept of knowledge and behavior and when financial literacy is defined as a combination of three dimensions - knowledge, behavior, and attitude (Martins et al., 2021).

2.1.3.4 Area of residence

Higher levels of literacy are registered in individuals that live in both metropolitan areas and that have superior incomes. (Sebastião et al., 2021). Depending on the different proxies of financial literacy adopted, according to Martins et al. (2021), the relevance of this variable varies. For example living in the metropolitan area of Lisbon has a positive influence in all models excluding the model with the subjective measure of financial literacy. According to Abreu and Mendes (2010), living in Porto metropolitan area translates in a higher likelihood of having higher financial literacy. Sincal and Marôco (s.d, as cited in Rosa, 2021) state that there are regional differences in the financial literacy performance, favoring individuals from the following NUTS, Leiria and MédioTejo, whereas individuals from Alto and Baixo Alentejo and Região Autónoma dos Açores underperform significantly.

2.1.3.5 Cognitive abilities: numeracy and cognitive reflection

Financial literacy can also be linked with cognitive abilities as numeracy and cognitive reflection test. Firstly, according to Agarwal and Mazumder (2013), individuals when encountering choices, their cognitive ability has the capability to influence how various information is processed. For example when choosing a portfolio, not only financial knowledge alters the perception of information, but also other cognitive abilities as it is the case of memory or computational capacity.

Furthermore, it is stated that mathematical capacity has the strength to decrease the probability of sub-optimal choices occurring, being a powerful component for financial decision-making. Jappelli and Padula (2013) and Lusardi and Mitchell (2011), explain the importance of mathematical capacity through mathematical test scores (as it is the case of PISA), specifically factors as numerical skills at a younger age seem to be preponderant in the individual 's financial literacy. Regarding numeracy, that is often measured jointly and as part of financial literacy, Hung et al. (2009) explain that numeracy should be seen as a different construct from financial literacy- even though these are intertwined, numeracy is a more extensive ability that, even though it assumes an important role in financial literacy, it goes beyond financial topics and it is applied in various areas.

In terms of cognitive reflection, and according to Kahneman (2013), psychology literature has demonstrated that when individuals are asked to perform a challenging cognitive assignment, they have

a higher likelihood of not being able to resist the temptation. The cognitive reflection test (abbreviated as "crt") introduced by Frederick (2005), measures the propensity for individuals to entail in deeper reflection in order to overcome intuitive incorrect answers that are associated with poor processing. It is commonly known for its three questions, although they have been widely spread, becoming well known for some individuals. As new items were considered, the measure of reliability increases and displays strong indicator of performance for functions who demand rational reflection and disposition for thinking. Furthermore, Toplak et al. (2014) state that crt should be seen as a measure of rational thought, rather than a predictor. Cognitive ability refers to "computational power that is available to the individual", whereas crt looks for the intensity of processing. Failing at rational responding in crt, may also derive from lack of knowledge in certain areas as financial literacy, mathematics/statistics, etc. According to Muñoz-Murillo et al. (2020), individuals who perform better in cognitive abilities have a superior level of financial literacy, indicating that these individuals may present an easier learning process when it comes to financial concepts.

2.1.3.6 Experience

Interestingly, several types of experience impact the level of financial literacy of each individual, whether it regards the pension system adopted of their country (Jappelli & Padula, 2013), the financial knowledge of siblings or family and financial literacy (Van Rooij et al., 2011), or even at a more macro level regarding their country past financial distress (Lusardi & Mitchell, 2011). When it comes to actual experience in financial practices, experience assumes a preponderant role in financial knowledge for some authors, although controversial in which direction it occurs. Hilgert et al. (2003) explain that overall as knowledge increases, it appears to exist refinements in financial practices from individuals. However the authors point out that this causality may not only exist with this direction. Thus, knowledge can be enhanced through experience and practices, not only by the individual itself but also from others that surround him that influence him. Moore (2003) agrees with this last causality hypothesis direction, stating that as financial experience increases and individuals display appropriate financial behaviors, the individuals will have better financial literacy, thus financial literacy rises derived from experience.

2.1.3.7 Risk profile

Risk profile seems to be connected with financial literacy, and as Martins et al. (2021) suggest, individuals who are less risk averse consider themselves more financially wise which, and in this case, it is also aligned with higher levels of financial literacy (objective knowledge). As actual knowledge

increases, the probability of ownership of standard products also increases for both genders. However, perceived knowledge only has a positive relationship with standard products for men. As for sophisticated products, perceived literacy is relevant for men and women (Bannier & Neubert, 2016). Women tend to be more risk averse than men as observed in Nilesh et al (2021) and Barber and Odean (2001). Moreover, Bannier and Neubert (2016), affirm that there is no relationship between risk tolerance and advanced investments for women, contrarily to what is observed in men. Women need high levels of actual financial knowledge when compared to men for investment decisions. Thus, risk tolerance in women is only related with the investment in standard products, whereas in men it is also connected with sophisticated products (as risk tolerance decreases, the likelihood of owning risky products also diminishes). The marital status also influences risk, and single individuals perform riskier choices than married as observed in Barber and Odean (2001); as well as the age of the individual. As age increases, the individual's willingness to take risks decreases, being possibly explained by experience, limited time to reverse possible losses from riskier choices or asset accumulation as retirement draws near (Yao et al., 2011).

2.1.4 Financial literacy in Portugal

The increasing awareness of financial literacy importance is visible through several implemented mechanisms in Portugal. It is the case of *Plano Nacional de Formação Financeira*, the incorporation of financial contents in school education through the course of "cidadania e desenvolvimento", as well as in the respective textbooks and in teaching training (OECD, 2020). *Plano Nacional de Formação Financeira*, among many actions, throughout the years has carried away several reports that allow an intertemporal comparison of Portuguese financial literacy in several of its domains - so far three reports have been conducted in 2010, 2015 and 2020. When comparing 2010 to 2015, there is an improvement when it comes to financial literacy, however in 2020 the average score of financial concepts decreased when compared to the 2015 results (for example in 2015, 76.7% of the inquires correctly responded half of the questions, whereas in 2020 only 59.1% were able to do so). Nonetheless, despite this drop in the average score, there are more individuals in 2020 who successfully answer all questions (8,9%) (Plano Nacional de Formação Financeira, 2021). Observing the success rate in the questioned financial concepts that are also present in this research, in 2015 the concept of inflation (87%), relationship between risk and return (81%), risk and diversification (72.4%) were widely known, whereas the compounding interest (39.5%) and relationship between interest rates and bond prices drastically fell short (only 6.5% were able to correctly answer it) (Plano Nacional de Formação Financeira, 2016). OECD

(2016) also presents results regarding Portuguese financial knowledge in 2015, demonstrating that very similar results to the result from the national report (compounding interest=41%, inflation=87%, relationship between return and risk=82% and relationship between risk and diversification=73%). In 2020, all rates dropped – inflation to 74.4%, relationship between return and risk to 71.6%, relationship between risk and diversification to 45.1% and compounding interest with 31% (Plano Nacional de Formação Financeira, 2016).

According to Lusardi and Oggero (2017), worldwide, 33% of the adults are considered literate (correctly respond 3 out of 4 standard questions). In Portugal, only 26% of the adults are financially literate, therefore displaying a worrying concern. Portugal underperforms relatively to the worldwide average but also to the European average (52%), being surpassed by the majority of the countries except for some of the last countries entering EU. Portuguese millennials present a more positive, but still low, number with 38% of millennials being considered financially literate. OECD (2020) considers that overall, Portuguese youngsters, who will become the next future generation, have minimum knowledge to deal with daily financial challenges (Portugal is ranked 7th in financial literacy performance).

2.2 Behavioral Biases

Throughout the last decades, behavioral economics and finance have emerged as a contradiction of some of the main assumptions regarding traditional views of decision-making behaviors (Whitehead et al., 2014). Furthermore, as acknowledged in Leal and Armada (2005), behavioral finance, as a new line of thought introduces important factors that are lacking in the efficiency theory, namely psychological and behavioral aspects that influence decision-making and incorporate a more realistic assumption regarding the "full rationality", entailing limited rationality.

Derived from the vastness of daily choices and cognitive limitations, individuals rely on heuristics and shortcuts to increase the agility when facing different choices. Heuristics are defined as rules of thumb used by individuals to perceive information and make decisions. However, heuristics are imperfect, being associated highly with biases and the errors derived from these are systematic (Shefrin, 2002). If anomalies are independent between investors, as economists proclaim, then in equilibrium these should nullify its impact. However, as mentioned by Hirshleifer (2001), the problem arises with the assumption of independence in biases of individuals, since in fact these appear to be systematic biases that are shared by individuals. According to the same author even if individuals have positive incentives, perform training, or even do it repetitively, still, biases are not fully eliminated. Additionally,

not always the best outcomes derive from those biases that are projected most of the time unconsciously (Leal & Oliveira, 2021).

In the following work, two behavioral biases are thoroughly analyzed, namely overconfidence and loss aversion. These two biases are going to be used, initially in explaining financial literacy, as well as in comprehending market participation and investor's disposition effect. Overconfidence is a cognitive bias and loss aversion an emotional bias (Pompian, 2006), therefore allowing a broader contribution of the two types of biases when explaining behavior. The choice falls back on the fact that these behavioral biases seem to be the most fitting when considering the key areas of this research (market participation and disposition effect), as well as their dissemination therefore allowing a better literature comparison.

2.2.1 Overconfidence

Overconfidence is defined as “unwarranted faith in one's intuitive reasoning, judgments, and cognitive abilities” (Pompian, 2006, p.72). Additionally, Abreu and Mendes (2010) explain overconfidence as a conviction by investors that their choices are exceeding the average. This translates in individuals assuming that they have full information, think highly of their capabilities - some even assume that they can beat the market - trade excessively, and generally hold undiversified portfolios (Abreu & Mendes, 2010; Pompian, 2006; Shefrin, 2002). Gaudecker (2015) explains that when under diversification is observable, the individuals that suffer from higher losses do not display basic literacy (measured by numerical and concepts questions) or do not tackle this lack of literacy with external advice - overconfidence may be one of potential cause from this pattern. Despite lack of knowledge by these investors which would make them susceptible to higher losses or higher risk, they tend to be overconfident in their capabilities and judgement, more than the rest of individuals. This also explains why wealthier investors have smaller losses, since these will less likely have low financial literacy and still trust in their own judgment.

Moreover, it should be stressed out that regarding this low level of literacy that impacts our actions, Disney and Gathergood (2013) introduce a striking finding, notably the existence of self-awareness of weaknesses in financial knowledge by those that have weak financial literacy. Hayat and Anwar (2016) state that the increase of financial literacy has a positive impact on the overconfidence of the investors that way increasing the appearance of it. These contradict some unrealistic optimism and overconfidence of the investors that is explained in Thaler and Sunstein (2008) and also in Takeda et al. (2013). These authors assume that as financial literacy rises, the bias of overconfidence decreases, so financial knowledge acts as a deterrent from biased decisions in the financial market. In addition, Baker et al.

(2019) and Gerth et al. (2021), diverge from these two positions, acknowledging that there is no found relationship between overconfidence and financial literacy. Controversially, pursuant to Disney and Gathergood (2013), even though these individuals lack financial literacy, they do not endeavor in actions to overcome and offset this blindness in a complex market, even though cheap or free tools as newspapers or websites.

Other factors as sociodemographic appear to also have an impact when measuring this relationship. Özen and Ersoy (2019), show that the bias of overconfidence is superior in individuals that are either employees of financial enterprises or financial clients and these tend to highly avoid regret. Additionally, other variables that also affect the existence of this bias are the conclusion of a college degree, the level of income and the asset's total, which decrease its appearance (Takeda et al., 2013). Sebastião et al (2021) goes further in the education aspect, affirming that those who present as study field an area related to finance, more likely have a higher perceived knowledge when compared to the actual knowledge, even though these individuals do present high actual financial knowledge. Baker et al. (2019) affirm that gender also affects overconfidence since men display more overconfidence when compared to women. Sebastião et al. (2021) and Nilesh et al. (2021) contradict this thought, demonstrating that women tend to overestimate their financial knowledge.

2.2.2 Loss aversion

Loss Aversion is a key point in the prospect theory and in Kahneman and Tversky (1979) an important conclusion is brought up namely how individuals oversee the evolution of their trades. According to Kahneman and Tversky (1979), loss aversion results from higher urge in avoiding losses than obtaining gains. Specifically, because the feeling of loss has a greater impact when compared to a gain in the same amount, more specifically about two and a half times. Therefore, individuals will avoid selling at a loss, thus they need at least 1.5 to 2.5 of gains when compared to losses in order to accept (Burton & Shah, 2013). Instead of individuals monitoring the evolution of their assets in regard to the market evolution, they do it with respect to gains or losses compared to the price of acquisition. This bias is also linked with the concept of “get evenitis”, that corresponds to the individual not wanting to sell at a loss with the hope of at least being able to get even before selling it. Consequently, individuals will take risks to avert suffering a loss.

This bias does not appear to only impact illiterate individuals, but also sophisticated investors (Shefrin, 2002). Inclusively Gerth et al. (2021), present a positive relationship between loss aversion and financial literacy. Female investors are also more likely to present higher levels of loss aversion, as

evidenced by Rau (2014) and as wealth equity increases, the individual decreases its loss aversion (Feng & Seasholes, 2005)

2.3 Market Participation

2.3.1 Financial literacy and Market Participation

Financial literacy appears to be positively linked to the presence in financial markets (Klapper et al., 2013; Sebastião et al., 2021; Yamori & Ueyama, 2021). Sebastião et al. (2021) demonstrate that this positive correlation occurs in actual knowledge, as well as with perceived knowledge. Individuals with high actual and perceived literacy are 21% more probable of owning financial instruments (stocks, bonds, mutual bonds, etc) than those with lower scores in actual and perceived, additionally individuals with high actual and perceived literacy are more likely in 13% to have stocks when compared with individuals with low perceived and high actual literacy (Allgood & Walstad, 2016).

2.3.1.1 Portfolio Composition

The components of the individual's portfolio can indicate lack or expertise in financial literacy, namely those that possess higher debt cost are more likely to be illiterate (Disney & Gathergood, 2013). The composition of individual's portfolios also varies within income/wealth, age, and education, that are also commonly used indicators of financial literacy, as previously explained. According to Calvet et al. (2007), individuals with financial wealth below average mainly own cash, mutual funds and individuals' securities, and as wealth increases real estate and risky assets increase its proportion, and the individuals located at the 10% richest invest largely on risky assets. The same authors explain that age interferes with wealth composition, that is observable through the differences in real estate (this has larger importance for younger households, reaching its maximum importance of 80% at age 40, and from that on it decreases its importance up to 50% for elders) and through cash (highest point is observed with youngsters linked with house buying, decreasing its significance as age goes by, especially for elders that reduce its saving and increase consumption). Other products remain fairly constant throughout individuals' life: mutual funds display evenly 10% and individual stocks present a proportion of 6%, that increases lightly for youngsters and middle-aged individuals (Van Rooij et al. (2011) considers that there is no relevant relationship between owning stocks and age. Stocks appear as an important product for individuals located at 10% richest, that invest strongly in direct stockholdings with a proportion higher than 50%. Mutual funds overpower stocks for the majority of portfolios (except for the richest individuals) and have an important role in the diversification of household's portfolios. Lastly, these authors state

that riskier investments have higher costs, that for individuals with higher education these costs may be smaller, especially if the investments is larger enough.

2.3.2 Stock Market Participation

2.3.2.1 Financial Literacy and risk profile

According to Van Rooij et al. (2011) there is a certain reluctance in holding stocks. Overall higher levels of financial literacy are linked with higher market participation in stocks (Bucher-Koenen et al., 2021; Lusardi et al., 2009; Van Rooij et al., 2011; Yamori & Ueyama, 2020). According to Xia et al. (2014), risk profile of the individual is important at explaining stock ownership, since as the individual becomes more risk averse, he/she is less likely to participate in the stock market.

2.3.2.2 Cognitive and Behavioral biases(loss aversion and overconfidence)

The willingness to participate in indirect market (stocks included in other financial products) and in the stock market is positively related with cognitive abilities as it is the case of numeracy according to Christelis et al. (2010). In terms of behavioral biases as loss aversion and overconfidence, loss aversion negatively affects strong market participation through uncertainty avoidance (Rieger, 2020). Yang (2019) also presents a negative relationship but assumes that loss aversion impacts stock participation through risk aversion – this negative correlation derives from the reluctance in realizing losses combined with the stock ownership being viewed as highly volatile. By contrast overconfidence presents a positive relationship with stock market participation (increasing by 20% the likelihood of participating), and under confidence a negative relationship (Xia et al., 2014). Christelis et al. (2010) do not find overconfidence as relevant when explaining stockholding, however their proxy for overconfidence is whether the individual was depressed last month.

2.3.2.3 Sociodemographic factors (gender, age, income, education)

In accordance with Yamori and Ueyama (2021), participation in the stock market does not differ much between gender, contrary to some evidence in the literature where women participate less in the stock market as it is portrayed in Bucher-Koenen et al. (2021) and Rieger (2020). Even by controlling sociodemographic variables, and the initial financial literacy used where the option “ I don ´ t know” was available, it appears to still exist a difference between genders in stock market participation, namely women are 4,61 percentage points less likely to (Bucher-Koenen et al., 2021). Rieger (2020) evidences the negative relationship between stock market participation and age, therefore favoring younger

individual (contrarily to Xia et al., 2014, that do not find any relevant relationship between these variables). Individuals with higher income are more likely to participate in the market as stated in Van Rooij et al. (2011), nonetheless it does not appear relevant according to Xia et al. (2014). Individuals with superior education are more likely to know the stock market functioning and asset pricing and invest in foreign stocks, according to Lusardi et al. (2009).

2.4 Disposition effect

Disposition effect relates with how individuals change their action depending on facing gains or losses through a S-shaped function, according to Shefrin (2002). Usually, individuals will sell their gains too soon but when facing losses they will hold it too long. This action of keeping an asset when performing losses relates with the negative sentiment associated with the loss that the individual is trying to avoid. Even though taking a loss carries an advantage legal-wise (that is why in some research the disposition effect is not observable at the end of the fiscal year - e.g Odean, 2019), many individuals experience more pain in the loss than the tax benefit derived from the recognition of the loss. This bias derives in part by emotional elements as it is the case of regret, therefore if the individual that faces a loss realizes that loss, immediately the investor will feel the negative impact, whereas when postponing the loss realization, the individual still has the possibility of reversing this scenario or the individual may extend its denial period. Even before realizing the loss, the individual may already be anticipating the regret of that action, as explained by Fogel and Berry (2010) and Li and Yang (2012).

According to Kahneman (1982), despite many beliefs that individuals tendentially are risk-averse and will only choose risky options if there is a high monetary compensation for that risk, there seems to exist a propensity for risk-loving options especially when facing a loss with certainty compared with a high probability of a loss with even greater magnitude. Thus, gains and losses do not portray the same behavior regarding risk – gains tend to be accompanied by risk-averse actions, whereas losses are followed by risk-seeking choices. This is observable through the S-Shaped value function that features concavity when handling gains and convexity with losses and displays how a loss with the same magnitude as a gain has greater effect. Additionally, gains and losses are evaluated in terms of changes, that is relatively to their reference point. Therefore, this line of action does not follow traditional expected utility theory (assumes that the utility derived is the weight sum of the product of the probability of that event happening and utility of each outcome). However, according to Da Costa et al. (2008), if individuals remove the reference point and substitute it into price changes, the investor stops seeing it as gaining or losing, therefore affecting the prevalence of the disposition effect.

2.4.1 Disposition effect determinants

2.4.1.1 Financial Literacy and experience

Regarding the relationship between financial literacy and the occurrence of disposition effect, Baker et al. (2019) concludes that there is a negative relationship between financial literacy and disposition effect, thus as literacy increases, the predisposition to this bias diminishes. Dhar and Zhu (2006) also reach the same conclusion, although it is used demographic factors as proxy for financial knowledge. Interestingly, Feng and Seasholes (2005), affirm that the hesitancy in realizing losses, observable in disposition effect, is eliminated when investor sophistication and experience in trading, are observable jointly. Separately, individual sophistication decreases by 67% the existence of the disposition effect, and experience in trading diminishes by 72% this bias. Experience in trading appears to lead to better control in withstanding impulses, especially in the case of selling and buying as it is the case in disposition effect. However, the other side of the coin in disposition effect does not happen, since when realizing gains early, these two factors together are not enough to eliminate this bias, being only responsible for a reduction of 37%. Seru et al. (2010) finds a similar conclusion in which investors indeed are less likely by 2% to present disposition effect as experience increases. This research emphasizes that the learning process associated with experience incorporates two types of learning: investor improving as the individual gains experience, but also the investor no longer trades when the individual perceives himself with deficient skills/abilities – this last type is often not accounted, which may overestimate the impact of learning.

2.4.1.2 Behavioral biases (loss aversion and overconfidence)

In terms of the studied behavioral biases it is found a negative correlation between loss aversion and realizing capital losses, therefore favoring the engagement in disposition effect (Rau, 2014). In terms of overconfidence, overconfident individuals present a more robust disposition effect, as a result of the emotional side when for example realizing gains. Overconfident individuals will attempt to retrieve gains earlier as a result of the pride and the success feeling that comes with it, reversely to realizing losses since this action is accompanied by feelings of shame (Chu,2012). This positive relationship is also featured in Ho (2021).

2.4.1.3 Sociodemographic factors (gender, income, occupation and age)

Demographic factors also affect the existence of it. In terms of gender, men are more prone to realize losses (Feng & Seasholes, 2005). Rau (2014) contradicts this affirmation, stating that female

investors present higher disposition effect, as a result of their higher loss aversion (the gender difference in disposition effect is explained by the losses component, where women with a higher loss aversion are more unwilling to realize losses). Whereas Da Costa et al. (2008) assume that if the reference point is used than there are no differences between gender, however if the reference point changes from the initial bought price into the last presented price women are less likely to present disposition effect. Individuals with superior wealth or income and in “professional occupations”, that is defined by working either in technical or administrative functions seem less affected by disposition effect (Dhar and Zhu, 2006). Age is also relevant since older individuals are less likely to engage in disposition effect, despite their weakened cognitive abilities compromising their financial skills (Korniotis & Kumar, 2011)

3. Methodology

3.1 Method

Posterior to the revision of literature that provides the state of art of the general theme, it must be presented which research strategy is used, as well as techniques for collecting data and a thorough explanation related to the research, including the target population and sample. Any research underlies philosophical assumptions and paradigms. There are two central paradigms in management research, in particular a positivistic or quantitative approach and an interpretative or qualitative paradigm, that will present differently four main assumptions: ontological, epistemological, axiological, and methodological.

In this research, it is used a quantitative or positivist paradigm on the account of large sample to reduce biases and sampling errors, quantitative methods and hypotheses testing and formulation, in order to generalize the results of the sample to the rest of the population (Bryand & Bell, 2007; Coutinho, 2013; Esterby-Smith et al. 2018). This study aims to understand the level of financial literacy in Portugal and the subsequent behaviors that derive from it, thus a quantitative study appears to be more suitable to the specific objectives previously mentioned. Additionally, the research design will relate to a cross-sectional study, since it will provide a deep insight into the financial literacy of the sample and it was collected at only one point in time (Coutinho, 2013).

In order to study the proposed goals, several hypotheses are created and tested through various regressions. The opted regression models for this research are probit and ordered probit, since the studied dependent variables are either binary (in this case it is implemented a probit model) or categorical variables (associated with the use of ordered probit). After the computation of either probit or ordered probit, it is calculated the marginal effects, otherwise the presented coefficients cannot be interpreted in terms of magnitude.

3.2 Hypotheses and models

Several hypotheses are tested in this research, aligned with the literature results. In order to have a better understanding of the following presented hypothesis, [table 1](#) summarizes all variables used as well as their meaning. The percentage of securities in overall wealth is simplified and presented as “investment importance”.

H1. Financial Literacy is positively influenced by experience, risk profile, investment importance, cognitive and behavioral biases, and sociodemographic variables (age – U-shaped inverted pattern, gender, income, study field and area of residence).

$$Finknow_i = \beta_0 * experi_i + \beta_1 * studyfield_i + \beta_2 * prct_sec_i + \beta_3 * riskprofile_i + \beta_4 * overc_i + \beta_5 * loss_i + \beta_6 * numeracy_i + \beta_7 * crt_i + \beta_8 * age_i + \beta_9 * gender_i + \beta_{10} * income_i + \beta_{11} * metro_i + \beta_{12} * occupation_i + \epsilon_i; i=1,n$$

In order to simplify this hypothesis interpretation, it is divided into three subsets (H1.1, H1.2, H1.3).

H1.1 Financial Literacy is positively influenced by experience, risk profile and investment importance.

Regardless of the direction of the relationship between experience and financial literacy (it is not tested in this research), there seems to exist a positive relationship between these two variables according to Hilgert et al. (2003) and Moore (2003). In terms of risk profile, a positive relationship (as risk profile becomes less risk averse, financial literacy becomes higher) with financial literacy is observable in Martins et al. (2021). Regarding investment importance (percentage of securities value in total wealth), it is speculated that those with a higher weight of their portfolio in their overall wealth, are more financially literate due to the higher risk associated with this investment intensity.

H1.2 Financial Literacy is positively influenced by cognitive (numeracy and cognitive reflection) and behavioral biases (loss aversion and overconfidence).

Numeracy has a positive relationship with financial literacy as observable in Jappelli and Padula (2013) and Lusardi and Mitchell (2011), thus favoring individuals with higher mathematical skills. Cognitive reflection measured through the “cognitive reflection test” questions is also important to explain financial literacy. According to Munõz-Murillo et al. (2020), those that score highly in this test have higher levels of financial literacy. In terms of behavioral biases, as it is the case of loss aversion and overconfidence, literature suggests that there is a positive relationship between loss aversion (Gerth et al., 2021) and financial literacy, as well with overconfidence (Hayat & Anwar, 2016).

H1.3 Financial Literacy is positively influenced by sociodemographic variables (age – U-shaped inverted pattern, gender, income, study field and area of residence).

Various authors demonstrate that age does not follow a linear pattern when it comes to financial literacy, more precisely it presents a U-shaped inverted pattern (Abreu & Mendes, 2010; Hung et al., 2009; Lusardi & Mitchell, 2011 and Van Rooij et al., 2011). Overall literature considers the man as the most knowledgeable, when compared to female individuals (Abreu & Mendes, 2010; Hung et al., 2009;

Lusardi & Mitchell, 2011; OECD, 2016; Martins et al., 2021; Sebastião et al., 2021; Potrich et al., 2015 and Van Rooij et al., 2011), as well as individuals with higher income (Potrich et al., 2015). Individuals study background also matters when explaining financial literacy since those that come from a socioeconomic area present higher financial literacy (Sebastião et al., 2021; Van Rooij et al., 2021). Lastly, even though this variable is not always tested in literature, in some Portuguese studies, there are regional differences – Sebastião et al. (2021) present a positive relationship between living in metropolitan areas and financial literacy and Abreu and Mendes (2010) demonstrate that individuals that live in Porto metropolitan area demonstrate higher financial literacy – these last authors hypothesize that living in metropolitan areas may help individuals regarding financial literacy due to the availability of information.

H2. Market participation is positively influenced by financial literacy, risk profile, income, whereas age presents a negative relationship. Loss aversion decreases the willingness to participate, whereas overconfidence varies positively.

$$\text{marketpart}_i = \beta_0 + \beta_1 * \text{riskprofile}_i + \beta_2 * \text{finknow}_i + \beta_3 * \text{income}_i + \beta_4 * \text{age}_i + \beta_5 * \text{loss}_i + \beta_6 * \text{overc}_i + \beta_7 * \text{studyfield}_i + \beta_8 * \text{numeracy}_i + \beta_9 * \text{crt}_i + \beta_{10} * \text{gender}_i + \beta_{11} * \text{occupation}_i + \beta_{12} * \text{metro}_i + \varepsilon_i ; i = 1, n$$

Before presenting the literature that backs up the written hypothesis, it should be acknowledged that most of the literature focuses on stock market participation (a form of market participation), and not in overall view of participation in capital markets. Since stock market participation is part of market participation, even if stock ownership has a few singularities, it is expected that the literature results regarding stock ownership do not differ much from market participation (similarly to what happens in Christelis et al., 2010, in which indirect stock participation presents similar results as direct stock holding). Individuals with higher financial literacy are more likely to participate in financial markets - Klapper et al. (2013); Sebastião et al. (2021) - and in stock market (Bucher-Koenen et al., 2021; Lusardi et al., 2009; Van Rooij et al., 2011; Yamori & Ueyama, 2021). In terms of risk profile, it is speculated that individuals who are less risk averse are more likely to participate in the market due to the uncertainty and the risk associated, as it is observable in the stock market (Xia et al., 2014). Individuals with higher income are more likely to participate in the stock market as it is observable in Van Rooij et al. (2011). Regarding the individual's age, younger individuals are more likely to participate in stock the market (Rieger, 2020).

Regarding behavioral biases on market participation, loss aversion and overconfidence appear

to have contradictory directions when it comes to owning securities. Loss aversion has a negative relationship, therefore diminishing the likelihood in participating and favoring less loss averse individuals (Yang, 2019). Whereas, overconfidence displays a positive relationship, increasing the willingness to participate (Xia et al., 2014).

H3. The number of financial products owned is positively influenced by financial literacy, investment importance, income and risk profile and presents a negative relationship with loss aversion and overconfidence

$$\text{portfi} = \beta_0 * \text{finknow}_i + \beta_1 * \text{riskprofile}_i + \beta_2 * \text{prct_sec}_i + \beta_3 * \text{exper}_i + \beta_4 * \text{loss}_i + \beta_5 * \text{overc}_i + \beta_6 * \text{income}_i + \beta_7 * \text{studyfield}_i + \beta_8 * \text{crt}_i + \beta_9 * \text{numeracy}_i + \beta_{10} * \text{gender}_i + \beta_{11} * \text{age}_i + \beta_{12} * \text{metro}_i + \beta_{13} * \text{occupation}_i +$$

$$\epsilon_i ; i=1,n$$

The number of distinct financial products is sometimes used as a proxy for portfolio diversification (e.g Abreu & Mendes, 2010; Nilesh et al., 2021). Abreu and Mendes (2010) demonstrate a positive relationship between the number of different financial products with financial knowledge. Regarding investment importance and risk profile, it is hypothesized that these variables are positively associated with the number of distinct financial products. Those that invest more fiercely with a higher importance of securities in their overall wealth may be more likely to diversify more their portfolio, through a higher number of different products. In terms of risk profile, it is presumed that those that are more risk averse are expected to be more conservative and fearful when it comes to holding financial products. The positive association with income, is observable in Calvet et al. (2007), even though these authors display a positive relationship between income and number of different stocks. The expected negative relationship with loss derives from individuals with higher loss aversion avoiding the negative feeling associated with the possibility of realizing losses, as explained by Rieger (2020). Additionally, there is evidence that overconfident individuals are more likely to hold undiversified portfolios (Pompian, 2006; Shefrin, 2002). Therefore, it is expected a negative relationship between the number of financial products and overconfidence.

H4. The positive gap between hypothetical portfolio and actual portfolio is negatively influenced by financial literacy, experience, investment importance, risk profile, overconfidence and income, and positively influenced by loss aversion

$$pgap_i = \beta_0 * finknow_i + \beta_1 * riskprofile_i + \beta_2 * prct_sec_i + \beta_3 * exper_i + \beta_4 * loss_i + \beta_5 * overc_i + \beta_6 * income_i + \beta_7 * studyfield_i + \beta_8 * crt_i + \beta_9 * numeracy_i + \beta_{10} * gender_i + \beta_{11} * age_i + \beta_{12} * metro_i + \beta_{13} * occupation_i + \epsilon_i ; i = 1, n$$

This hypothesis is not supported by literature, contrarily to others, since it was not found studies that compared individual 's actual behavior and preferences regarding their portfolio composition. The positive side of the gap represents those that have a higher number of financial products in their hypothetical portfolio rather than in their current one. Rationally, this would be the expected result since individuals could display their preferences hypothetically without the real repercussions. Those that displayed a negative gap (hypothetical portfolio < actual behavior) may have done it in order to rebalance their portfolios, but most likely inquiries neglected this question as a result of their mental fatigue. It is expected that individuals with lower financial literacy, low experienced, with low fierceness at investing, more risk averse and loss averse, display a higher gap. It is also hypothesized a negative relationship with overconfidence, since these individuals may already consider their current portfolio optimal, therefore they do not present a high positive gap. These inquiries even if they misplace their 100 000 in their hypothetical portfolio, there are no real consequences, thus they should feel uninhibited at demonstrating their preferences. Lastly, it is also expected a negative association with income, since the individuals with lower incomes do not find money constraints when investing in this hypothetical example, contrarily to what happens in reality. Thus, they should display the highest difference regarding the number of financial products hypothetically.

H5. Investor 's disposition effect is negatively influenced by financial literacy, experience, as well as age and gender and positively influenced by loss aversion and overconfidence.

$$de_i = \beta_0 + \beta_1 * finknow_i + \beta_2 * overc_i + \beta_3 * loss_i + \beta_4 * prct_sec_i + \beta_5 * exper_i + \beta_6 * riskprofile_i + \beta_7 * income_i + \beta_8 * gender_i + \beta_9 * age_i + \beta_{10} * studyfield_i + \beta_{11} * crt_i + \beta_{12} * numeracy_i + \beta_{13} * metro_i + \beta_{14} * occupation_i + \epsilon_i ; i = 1, n$$

Higher levels of financial literacy diminish the likelihood of investors engaging in disposition effect (Baker et al., 2019; Dhar & Zhu, 2006). Experience also presents a negative relationship with

disposition effect, as stated in Feng and Seasholes (2005) and Seru et al. (2010). Older individuals are less likely to present this bias, even with lower cognitive abilities (Korniotis & Kumar, 2011). Rau (2014) affirms that female investors are more likely to display disposition effect as a result of female being more likely to present loss aversion. In terms of behavioral biases, loss aversion is positively related with disposition effect, derived from the negative relationship between loss aversion and realizing capital losses (Rau, 2014). Overconfidence also presents a positive relationship with engaging in disposition effect, derived from the emotional aspects especially the pride felt at retrieving gains and the shame when realizing losses (Chu, 2012; Ho, 2021).

H6. As the definition of financial literacy becomes broader, the differences of results increase.

Derived from the lack of universality when it comes to the definition of financial literacy, the results of different researches may be endangered based on which measure is adopted (Remund, 2010). Martins et al. (2021) demonstrate that as the adopted definition becomes broader, there are more differences in the results.

3.3 Questionnaire design

This questionnaire was constructed and distributed by Comissão do Mercado de Valores Mobiliários, and this research uses 27 out of the original 38 questions. Overall, this questionnaire can be decomposed in 6 categories of questions:

1. Socio-demographic (gender, education, study field, occupation, income, age, and residence)

These sociodemographic questions are standard in all literature in order to characterize the sample, but they are also important at explaining the main concepts as financial literacy, market participation and disposition effect. Throughout all literature, even if these variables are not the main focus at explaining some concept, they are always used as control variables.

2. Experience, risk profile, investment importance and identification of reasons of investing, or contrarily reasons for not investing questions.

Other variables that also seem relevant are experience and risk profile. The identification of reasons for investing or not, serves only for descriptive purposes in order to create a better framework of the sample. In order to measure experience, inquiries are asked when was the first time that they applied money in several financial products and risk profile is retrieved through individuals' self-perception of

their risk attitude.

3. Cognitive Abilities (Cognitive reflection test and numeracy)

Cognitive Reflection Test: In Kahneman (2013), the “bat and ball” problem that in this survey is replaced by a football ball and gloves, explains the impulsivity to follow our intuition and immediate response that came to mind. In this case, the “pop-up” number is 10 cents, however despite the easiness that it may transmit, it is the wrong solution. The individuals that are able to resist the temptation of an easy intuition and answer, answered 5 cents. Additionally, it is explained that these individuals did not question themselves why such an oblivious answer would appear. Thus, those that accepted their intuition, do not come over as active thinkers, being excessively confident in their intuition (Frederick, 2005). The other used question is introduced by Toplak et al. (2014), and it concerns the stock evaluation, in which the intuitive and wrong answer is “I did not lose or gain”, whereas the correct answer is “Lost money”.

Numeracy: Inquiries are asked to write the likelihood of getting tails on the tenth toss, after the first nine tosses were always tail. This question allows to evaluate probability knowledge. This question captures a certain mathematical knowledge, since even those individuals that would have succeed in standard numeracy questions regarding mathematical operations (multiplication, division, etc), may fail if the inquire does not comprehend the concept of independent probability.

4. Self-assessed financial knowledge and actual financial knowledge

Self-assessed financial knowledge: It is asked individuals to self-asses their knowledge concerning financial products and markets. This question is widely available in literature (as it is the case in Allgood & Walstad, 2016; Lusardi & Mitchell, 2011; Sebastião et al.,2021; Van Rooij et al., 2011, among others). Allgood and Walstad (2016), highlight that when evaluating the perceived financial knowledge, it is different to ask inquiries about their financial knowledge and foreseeing their ability – in this questionnaire it is asked for the first option.

Actual financial knowledge: it is measured through the Big Three questions plus three additional questions. The Big three, widely associated with Lusardi and Mitchell (2011), is composed by interest compounding, inflation and risk diversification - the framing of the risk diversification question is not the same as Lusardi and Mitchell (2011), whereas the OECD (2018) version was adopted (QK7_3). The other three questions assess the relationship between return and risk (available in OECD (2018) - QK7_1), structured product rentability (present in Martins et al.,2021; and Sebastião et al., 2021), and

the relationship between fixed interest bond and market interest rates (part of Big Five questions, available in for example in Moore, 2003).

Since financial literacy measurement is not consensual around literature, this research uses three different definitions of financial literacy, therefore allowing a further comparison of the results depending on which measure was used. It starts by the simplest definition, widely known and used in literature, that assesses financial literacy through objective knowledge. It is tested through financial question scores. The second measure comprises self-perceived and actual financial knowledge, similarly to the definition presented in Allgood and Walstad (2016), where individuals were segregated in four groups depending on high or low actual and perceived literacy. For this definition, the actual knowledge is obtained through the objective financial literacy score mentioned in the first definition, whereas the self-perceived component, is gathered through a question that asks inquiries to self-assess their knowledge regarding financial products and markets. The last definition, that features two dimensions of financial literacy, namely financial knowledge, and behavior, is presented as a broader measure, going beyond the notion of knowledge. This definition is presented in Huston (2010), that highlights the importance of knowledge, but also the application part of it. Once again, the financial knowledge component is measured through an objective financial literacy score, contrarily to financial behavior that is assessed based on the different types of products that the individual possesses. Afterward, the two components' scores are added, similarly to Martins et al. (2021).

5. Behavioral biases (loss aversion, overconfidence, and disposition effect)

Overconfidence is the only behavioral bias that does not have a direct question and is computed as the difference between actual and perceived financial literacy similarly to Takeda et al. (2013) and Xia et al. (2014). The loss aversion question, displays several statements in which there is 50% probability of winning a specific value and other 50% of losing a certain amount, taking into consideration that the value to be lost increases gradually, therefore it is assessed indirectly the loss aversion ratio. Disposition effect is measured through individuals' actions in losses and gains domains, similarly to Gabinete de Estudos da CMVM (2019). It is also tested an alternative question for disposition effect without reference point.

6. Market Participation (financial product ownership) and hypothetical portfolio composition.

Regarding the market participation, it is asked individuals which financial product they currently own. This question allows to identify current investors (those that at least invest in one security) and also

determine a simple proxy for portfolio diversification based on the different securities that the inquirer owns (similar concept in Abreu & Mendes, 2010; Nilesch et al., 2021; and Seru et al., 2010). As pointed out by Abreu and Mendes (2010), it is also acknowledged the disadvantages of this rather simplistic proxy for portfolio diversification, namely the inattention of the correlation between products. The hypothetical portfolio question asks inquirers to invest hypothetical 100 000 euros in several products.

4. Data

4.1 Data Collection

This research uses data that has been retrieved through online self-completion questionnaire applied to the Portuguese population by CMVM in 2021 aiming to characterize Portuguese individuals. This research uses “measured financial literacy” through questions that evaluate financial literacy, rather than “manipulated financial literacy” (measures financial literacy through experimental research in order to assess the impact of financial education). This distinction is present in Fernandes et al. (2014). Similarly to Van Rooij et al. (2011) and Lusardi and Mitchell (2011), that stress the need to distinguish the certainty of knowledge or just random guesses, through the inclusion of the option “I do not know”, this is also carefully addressed in certain questions since it affects the results of the survey, hence influencing mensuration of financial insights.

4.2 Data Treatment

The original sample has a total of 1850 inquiries. It should be noted that this questionnaire did not have mandatory answers, conditioning the type of sample since it does not force individuals to answer. If it did oblige inquiries, certain individuals would have disappeared, since they would have dropped the questionnaire. Thus there are individuals that did not answer every question and automatically in some regressions these individuals are not held into account. However, they are not removed out of the sample, since these individuals are most likely the ones with lower levels of knowledge and therefore this research would have been biasing the regression results. Initially the data base was treated in excel, being then exported into the Stata software.

The sample size complies with the average sample size superior at least to 1000 that was advocated in Houston (2010), since it is the mode and median of the 52 studies related to this theme. This section is divided in two subsets, regarding variables creation and treatment and open-answer questions transformation and inconsistencies in answers. The first subset explains how some variables are computed, assessed, and transformed. The second subset explains several alterations that are made in order to create a more cohesive sample, thus altering possible misinterpretations that lead to contrasting answers throughout the questionnaire. Lastly, Table 1 sums up all relevant variables as well as their meaning and range of values

4.2.1 Variables and open-answer questions

- For statistical convenience, **gender** only assumes two possible answers, thus transforming the option “other” as a missing value. Value 1 for Male individuals, and 0 otherwise (5 individuals answered “other” and overall there are 16 missing values.).

- When asking the study field of each inquire, this question that assumed an open-answer question was transformed, during data treatment, into an ordinal variable assuming only three values. The variable **studyfield** assumes the value 3 if the individual belongs to economic/management or related areas, 2 in the case of mathematical/engineering and related areas or 1 if the individual did not belong to either of the two possibilities. This variable treatment is present similarly in Sebastião et al. (2021).

- The variable **metro** presented 1 if the individual lives in Portuguese metropolitan areas, and 0 otherwise. Specifically, the value 1 is attributed if the first two number of the zip code is either 10-19;26-29 (Lisbon Metropolitan Area) or 37; 40-45 (Porto Metropolitan Area).

- Missing Values in financial knowledge questions are computed as wrong answers, therefore they are assigned the value 0, contrarily to those that correctly answer the questions with 1 for each question. The overall score of financial questions is computed into 5 levels, therefore creating **finknow**.

- **Finbehav** is another measure of financial literacy that includes financial behavior and financial knowledge. The financial behavior component is measured through the number of financial products that the individual owns (computed into 5 levels). The financial knowledge component is represented by *finknow*. After the sum of these two dimensions, this score is transformed into a 5 category variable.

- **Perceived** is also an alternative measure of financial literacy, that is present in Allgood and Walstad (2016). These authors create 4 categories in this variable that compares the actual level of financial knowledge and the self-perceived. A high objective financial knowledge occurs if the financial knowledge of the individual is higher than the mean score, and high self-assess score is considered high if it is superior to the average value. Thus four groups are constituted: High *finknow* and *selfassess*, low *finknow* and high *selfassess*, high *finknow* and low *selfassess* and lastly low *finknow* and low *selfassess*.

- Regarding the variable **exper**, it is computed taking into consideration time horizon, but also the number of financial products.

- In the case of the variable that includes investment funds and PPR, since the questionnaire did not tangle these, some individuals state different time periods for the two products, consequently when converging them in terms of experience it is assumed the longest time frame between these two, and it is viewed as one

product.

- Thus, for each financial product a variable is created in which the scale ranges from 0 to 3 (0 – “Never applied”; 1 – “Last year”; 2 – “More than a year but less than 5 years” and 3 – “More than 5 years”). After adding up the time horizon of each financial product, this score is transformed into a 4 category variable, more specifically the number 1 for zero experience, 2 for low ($1 \leq \text{score} \leq 14$), 3 for medium ($14 < \text{score} \leq 28$), and 4 for high ($28 < \text{score} \leq 42$).

- When questioning about the type of products that the individual currently owns, those that at least hold either, savings and treasury certificates/treasury bonds, stocks, commercial paper/bonds, sustainable investment funds, investment funds (including PPR), complex financial products, crowdfunding investments, and bitcoin, ICO and other digital coins investments are considered investors. Thus, the variable **marketpart** assumes only two values: 1 if the individual is an investor, and 0 if not. For those that state other securities in the option “other”, specifically financial insurances (including unit-linked and capitalization) and futures, the value 1 is also assigned in these cases.

- Regarding the overconfidence biases, it is subtracted *selfassess* from *finknow*. Thus, the maximum value is 4 and the minimum is -4. This range is converted into a 5 category variable denominated as **overc**.

- The variable **loss** has four values. Being 1 the lower level of loss aversion where the individual presents the lowest loss aversion ratio, and 4 the highest where the individual only accepts losses if gains are 10 times greater.

- Relatively to the question of **numeracy**, those that fail at delivering a number but explain that the probability of happening is exactly the same as the previous coin flipping experiments are considered correct. Several inquiries fail at answering this question not because of lack of numeracy, but fail at interpreting it (assumed the overall probability of having 10 flipping experiments with the same side)

- For the assessment of the cognitive reflection test question (**crt**) that is related to a bat and ball and stock evaluation problem, those that write an interval of possible numbers as an answer or those that reply integer numbers without stating the unit of measure (e.g. between “0.10 and 0” or stating “10” or “5”, without specifying if it is cents or euros), are classified as wrong answers.

- The variable **de**, that stands for disposition effect, assumes two values: 1 if this bias is present, and 0 otherwise. In order for it to be present, the individual when facing a drop in the price of a financial product that he/she owns compared to the buying price does not sell but when facing a rise the individual

immediately sells or partially sells it. It is added another possible choice of “Selling partially” derived from the vastness of individuals that answered it in the open-answer of “other”. The logic behind **de1** is the same, but the questions that compose this variable do not feature a reference point.

- The variable **portf** measures the level of portfolio composition/diversification of the individual considering the number of financial products.

- The variable **pgap** measures the positive difference between individuals’ preferences and actual behavior (only measures how far the hypothetical portfolio has more products than the current portfolio). The products presented in the hypothetical portfolio and in the section of current products displays a different division and structure, therefore in order to correctly compare these and subtract them, there must be a similar base. Gold is present in the hypothetical portfolio, but not in the current products, so this product is ignored. Thus, in specific categories of the hypothetical portfolio it is combined two products of the actual portfolio (it assumes the value 1 if it currently has one of the two products that correspond to one category in the hypothetical portfolio).

- Lastly, whenever there is the possibility of an open-answer or the option “other”, any answer that is not enlightening or it is an outlier, it is ignored, and in the cases where multiple options are written by the individual it is assumed the first one mentioned. Inclusively, if the individual chooses a category and also answers the open answer option, it is only considered the category chosen.

Table 1: Summary description of variables

Variables	Description
age	Age
agesq	Squared term of Age
crt	Cognitive reflection test score (varies between 0 and 2)
de	=1 if the individual engages in disposition effect
de1	alternative measure of disposition effect, =1 if the individual engages in disposition effect
exper	Experience score (varies between 1 and 4, 1 for zero experience, 2 for low, 3 for medium and 4 for high)
finbehav	Score of the combination of financial knowledge and financial behavior (varies between 1 and 5) [Financial Literacy Measure]
finknow	Financial Knowledge level (ranges from 1 to 5)) [Financial Literacy Measure]
gender	=1 for male individuals and 0 for female
income	=1 if income < 500€; 2 if between 501€ and 1000€; 3 if between 1001€ and 1500€; 4 if between 1501€ and 2000€; 5 if between 2001€ and 2500€; 6 if between 2501€ and 3500€; 7 if between 3501€ and 5000€ and 8 if income is > 5001€
loss	Loss aversion score (varies between 1 and 4, being 1 the individuals with lower levels of loss aversion)

marketpart	= 1 if the individual participates in the market
metro	=1 for individuals that reside in metropolitan areas and 0 for non-metropolitan residence
numeracy	=1 if correctly answers the numeracy question
occupation	= 1 if individual is self-employed, 2 if employee, 3 if unemployed, 4 if retired, 5 if student and 6 if the individual has a different occupation
overc	= 1 if very underconfident, 2 if underconfident, 3 if neutral, 4 if overconfident and 5 if very overconfident
perceived	Segregation in four groups: 1 if high finknow and selfassess, 2 if low finknow and high selfassess, 3 if high finknow and low selfassess and lastly if low finknow and low self-assess.
pgap	Positive Level of gap between preferences (hypothetical portfolio) and actual behavior (varies between 1 and 5, being 1 very low, 2 low, 3 medium, 4 high and 5 very high)
portf	Portfolio composition score (varies between 1 and 5, being 1 very low, 2 low, 3 medium, 4 high and 5 very high)
prct_sec	= 1 if the individual does not have securities, 2 if the percentage of securities in total wealth is between 0 and 25%, 3 if it is between 26% and 50%, 4 if it is between 51% and 75% and 5 if the percentage is more than 76%
riskprofile	= 1 if individual considers being high risk averse, 2 if risk averse, 3 if neutral to risk, 4 if risk taker and 5 if high risk taker
selfassess	Financial Knowledge self-assessment score (varies between 1 and 5, being 1 very low, 2 low, 3 medium, 4 high and 5 very high)
studyfield	= 1 if individual has a non-economic or mathematical area, 2 if mathematics, engineering, and related areas and 3 if the individual has economics, management and related areas

Note: Table 1 summarizes all variables that are used in this research.

4.2.2 Inconsistencies

- For those that answered, “other products” but their written answer could be embedded in one of the options available, this written answer was eliminated and substituted by the ownership of that product in the correct category – this happened mainly with Exchange-traded funds (ETF) that were incorporated in the option “Complex financial products”.

- Those that stated currently having a specific product, but in the question of experience answered, “I have never applied” in the same product, their answer is altered into having applied money in “The last year”.

- Regarding the weight of the portfolio of securities in the total wealth (*prct_sec*) , those that responded one of the options with a specific percentage, even though they do not own securities at the moment, their answer is changed into the option “ I do not have securities”. Similarly, those that do have securities, but state that they do not have a portfolio, their answer is substituted by a missing answer, since the former answer did not portray a true image.

4.3 Descriptive Analysis

In the following section, a descriptive analysis is going to be conducted, in order to provide a first overview and picture of the data gathered and its distribution. As mentioned by Coutinho (2013), different measures and choices regarding descriptive statistics must be held into account depending on which type of data, and consequently variable, we encounter.

Table 2 displays summary statistics of the variables *age*. In terms of age, the average is approximately 32 years old, with a considerable high standard deviation, therefore demonstrating the high variability of ages between the individuals. The minimum age is 18 years old and the maximum reaches 77 years old, however as observed when comparing the mean and the median (median is considerably lower when compared to the mean), there is a positive skewness that can be explained with the pronounced proportion of younger individuals.

Table 2: Summary statistics of age

	N	Mean	Std. Dev.	min	p25	Median	p75	max
age	1837	31.718	13.865	18	21	25	43	77

Note: This table presents the summary statistics of the independent variable age. "Freq." Stands for frequency, "Cum" for Cumulative, "std.dev." for standard deviation, p25 and p75 for percentile 25 and 75, and "min" for minimum and "max" for maximum

In Table 3, there is no substantial difference in terms of gender participation, since the framework is roughly equally divided, however there are more female (52.60%) than male participants (46.40%). Most individuals are located in upper education, displaying a negative skewness. Having a pos-graduation, masters, MBA or PhD is the highest category of education presented, nonetheless it is the second most category that more individuals have chosen, nearly 35%. Almost the majority of the inquiries present economic, management or related areas as their study field (49.35%), followed by 37.89% who have not studied mathematical or economic areas – only 12.76% of the individuals display mathematical/engineering education. Most of them, are students (48.06%) or active workers (39.99% are employees and 4.42% self-employed), and for those that responded "other", the most common answer was working-student followed by researcher/grantee. Most individuals do not live in metropolitan areas (60.05%). Lastly, in terms of socio-demographical characteristics, the most chosen category refers to the lowest monthly income level (less than 500€), inclusively only nearly 30% of respondents have at least or more than 1501€ monthly. This derives in part from the individual's occupation (Carmér's $V = 0.2922$), for example the category with more individuals is "Less than 500€" which is in part a result of the number of students contained in the sample.

In terms of risk profile, the majority of the individuals consider themselves as risk averse and very risk averse (61.39%), being “risk averse” the most chosen category (46.96%). The extreme categories display different risk propensity, afterward only 1.17% assume being high risk takers, contrarily to the frequency of very risk averse that is considerably larger (14.43%).

Table 3: Summary Statistics regarding sociodemographic (gender, education, occupation, income, metro), risk profile and study field

Variable	Freq.	Percent	Cum.
gender			
Female = 0	983	53.60	53.60
Male = 1	851	46.40	100.00
Total	1834	100.00	
education			
No primary education = 1	2	0.11	0.11
Primary education = 2	1	0.05	0.16
Middle school education = 3	4	0.22	0.38
High School education = 4	126	6.82	7.20
Attending college = 5	699	37.82	45.02
College degree = 6	368	19.91	64.94
Pos-graduation, Masters, MBA or PhD = 7	648	35.06	100.00
Total	1848	100.00	
studyfield			
Non-economic or mathematical = 1	701	37.89	37.89
Mathematics, Engineering and related = 2	236	12.76	50.65
Economics, Management and related = 3	913	49.35	100.00
Total	1850	100.00	
occupation			
Self-employed = 1	81	4.42	4.42
Employee = 2	733	39.99	44.41
Unemployed = 3	50	2.73	47.14
Retired = 4	18	0.98	48.12
Student = 5	881	48.06	96.18
Other = 6	70	3.82	100.00
Total	1833	100.00	
income			
< 500€ = 1	498	28.92	28.92
Between 501€ and 1000€ = 2	390	22.65	51.57

Between 1001€ and 1500€ = 3	325	18.87	70.44
Between 1501€ and 2000€ = 4	195	11.32	81.77
Between 2001€ and 2500€ = 5	158	9.18	90.94
Between 2501€ and 3500€ = 6	87	5.05	95.99
Between 3501€ and 5000€ = 7	34	1.97	97.97
> 5001€ = 8	35	2.03	100.00
Total	1722	100.00	
<hr/>			
metro			
Non-metropolitan residence = 0	1111	60.05	60.05
Residence in metropolitan areas = 1	739	39.95	100.00
Total	1850	100.00	
<hr/>			
riskprofile			
Very risk averse = 1	185	14.43	14.43
Risk Averse = 2	602	46.96	61.39
Neutral to risk = 3	279	21.76	83.15
Risk taker = 4	201	15.68	98.83
High risk taker = 5	15	1.17	100.00
Total	1282	100.00	

Note: Table 3 presents the summary statistics of sociodemographic variables as it is the case of gender, education, occupation, income and metro, as well as risk profile and study field. "Freq. " stands for frequency, "Cum." for cumulative

Table 4 starts by exhibiting the differences within financial product ownership. Overall, most of the individuals of this sample have demand deposits (83.41%). After this first place, the other top five financial products are insurances (almost half of the sample owns it – 45.78%); followed by investment funds including PPR with 23.78%, housing credit with 18.43% and stocks with 18.27% . Conversely, there are financial products that less than 5% of the inquires hold, specifically complex products (2.22%), crowdfunding investments (2.27%), pension funds (2.59%), and commercial paper/bonds (3.78%). The ownership of these types of products is influenced by several factors, for example there is an association between owning a pension fund and age (Kruskall Wallis = 22.371; $p = 0.0001$), therefore the ownership of specific products may be sensible to the individual´s and sample characteristics. Another example may be regarding the proportion of housing credit, even though it features the fourth place, the ownership of this product is linked with the individual´s age (Kruskall Wallis = 252.940; $p = 0.0001$) – this sample can be considered young, thus if the sample assumed older values of age, most probably, there would be an even higher propensity to own housing credit. Experience demonstrates that this sample is low experiences, since the grand majority (82.79%) is located in the "low" category, followed by the medium category with only 10.49%. Very few inquiries are considered highly experienced (0.49%).

In this sample, an individual is considered an investor if the inquire currently owns at least one security. That way, almost half of the sample is considered as current investors (45.14%) and the majority of these consider as the main reason for investing a higher yield when compared to bank deposits (59.62%), followed by a smaller share of investors who pointed out “realize capital gains” as their main reason (15.43%). Interestingly, the influence of external individuals or parties does not appear as a decisive reason for individuals having securities, whether they are institutional sources (6.29%) or closed relatives (5.14%). For those that do not invest, they considered as not having enough money as the central cause (56.85%), followed by not having enough knowledge (31.18%). Lastly, when asking individuals what made them or what would make them invest, the option with higher acceptance is having more money, higher securities ´rentability and more information regarding securities. Overall, all facts presented appear relevant since all of them were chosen by almost all individuals, inclusively “Having more money” was selected by 95.98% of the individuals.

Table 4: Summary Statistics regarding investors, product ownership, reasons to invest and not to invest, which facts made them start investing or would make them invest and experience level

Variables	Freq.	Percent	Cum.
Type of products currently held			
prod_demanddep	1543	83.41	
prod_structdep	120	6.49	
prod_treasury	298	16.11	
prod_stocks	338	18.27	
prod_bonds	70	3.78	
prod_sustinv	96	5.19	
prod_funds	440	23.78	
prod_pension	48	2.59	
prod_complex	41	2.22	
prod_insurance	847	45.78	
prod_crowdfund	42	2.27	
prod_bitcoins	190	10.27	
prod_housecredit	341	18.43	
prod_othcredit	191	10.32	
marketpart			
Does not have securities = 0	1015	54.86	54.86
Has one or more securities = 1	835	45.14	100.00
Total	1850	100.00	
prct_assets			

Does not have securities = 1	1015	62.97	62.97
Between 0% and 25% = 2	374	23.20	86.17
Between 26% and 50% = 3	105	6.51	92.68
Between 51% and 75% = 4	57	3.54	96.22
More than 76% = 5	61	3.78	100.00
Total	1612	100.00	

reason_inv*			
Higher yield than bank deposits = 1	313	59.62	59.62
Advice from bank account manager = 2	33	6.29	65.90
Tax reasons = 3	17	3.24	69.14
Friends/acquaintances/family members also invest = 4	27	5.14	74.29
Realize capital gains = 5	81	15.43	89.71
Enjoying risk = 6	19	3.62	93.33
Habit = 7	15	2.86	96.19
Other = 8	20	3.81	100.00
Total	525	100.00	

reasonnot_inv**			
I do not have sufficient knowledge = 1	289	31.18	31.18
too risky = 2	57	6.15	37.32
costs are high = 3	20	2.16	39.48
I lost money on past investments = 4	6	0.65	40.13
I do not have enough money = 5	527	56.85	96.98
other = 6	28	3.02	100.00
Total	927	100.00	

What made or would make individuals invest			
fact_money	787	95.98	
fact_info	460	87.62	
fact_trust	335	82.72	
fact_stab	272	77.49	
fact_rentab	424	87.78	
fact_lowint	362	84.78	
fact_do notinv	95	52.20	

exper			
Zero = 1	117	6.32	6.32
Low = 2	1530	82.70	89.03
Medium = 3	194	10.49	99.51
High = 4	9	0.49	100.00
Total	1850	100.00	

Note: Table 4 presents the summary statistics of the type of products held, the market participation, and facts and reasons that made them or not invest, as well as the level of experience. *Since there is no obligatory in answering, it is ignored the option of " I do not have a portfolio of securities", since the answer in this question is only valid if the individual is an investor or owns securities. **Those that have securities but also answered why they do not have and those that currently do not have securities, but responded the reason they have, they are ignored due to the overlapping of conflicting answers.

In table 5, the average correct response is not considered relatively high, with an average of 3.7 out of 6, assuming all correct responses. Nonetheless, this average hides a certain divergence, immediately observable through the high standard deviation, between individual's knowledge of financial concepts, as it is going to be addressed hereinafter. Regarding cognitive characteristics that are assessed through a cognitive reflection test score (the score ranges from 0 – no correct answers – up to 2 with the two questions correctly answered), the average score is 0.78, demonstrating that in average not even one question is successfully answered. The numeracy question appears to divide inquiries, since only half of them (50.7%) correctly answer this question.

Table 5: Summary Statistics regarding financial literacy, cognitive reflection test and numeracy

	N	Mean	Std. Dev.	min	p25	Median	p75	max
Financial literacy	1850	3.368	2.084	0	2	4	5	6
crt	1850	.783	0.806	0	0	1	1	2
numeracy	1850	.507	0.500	0	0	1	1	1

Note: Table 5 displays the summary statistics of the level of financial knowledge, the cognitive reflection test and numeracy. "Freq." stands for frequency, "Cum" for cumulative, "std.dev." for standard deviation, p25 and p75 for percentil 25 and 75, and "min" for minimum and "max" for maximum

As observable in table 6, not all concepts questioned appear to present the same difficulty The concept that displays the higher proportion of correct responses is regarding the relationship between risk and return with the majority of individuals acing it – nearly 75.41% of the individuals answered correctly. Other concepts as inflation and the relationship with risk and diversification also appear to be known by the individuals, with 66.16% and 63.62% respectively. All other concepts present lower correct response rates being understood by less than 60% of the sample. However, there is a question that clearly diverges from this scenario. When individuals were asked about the relationship between fixed interest bond and market interest rates, only 29.24% managed to correctly answer it, demonstrating clearly that this question is not on the same level of difficulty for the inquiries. In terms of the subjective financial literacy that is assessed through the self-perceived financial literacy that each individual states, most individuals do not self-assess their knowledge towards the "Very knowledgeable" category, and as a result the majority of individuals situate between little knowledgeable (37.46%) and moderately knowledgeable (31.40%). Finally, when questioning individuals about their knowledge when compared to the Portuguese average, most individuals consider that at least their knowledge is equal to the Portuguese average. The category with more individuals is "Above average" with 36.97% which could

translate either in individuals perceiving the national average as low and/or their knowledge being high.

The first question presented that relates to a ball and glove problem, despite its popularity and common use, it has lower correct responses when compared to the stock evaluation problem (30% vs 48.27%). Lastly the majority of the individuals have low levels of experience (almost 83%), and not even 1% of the individuals has high levels of experience, therefore demonstrating a sample with overall low levels of experience.

Table 6: Frequency regarding the level of objective financial literacy and self-perceived, cognitive reflection test, and experience

	Variables	Freq.	Percent	Cum.
Financial Literacy				
	interest_dummy	963	52.05	
	inflation_dummy	1224	66.16	
	riskreturn_dummy	1395	75.41	
	riskdivers_dummy	1177	63.62	
	structures_dummy	930	50.27	
	fixinterest_dummy	541	29.24	
Cognitive Reflection Test				
	crt_ball	555	30.00	
	crt_stocks	893	48.27	
	Crt score = 0	846	45.73	45.73
	Crt score = 1	560	30.27	76.00
	Crt score = 2	444	24.00	100.00
selfassess				
	Not knowledgeable = 1	191	12.57	12.57
	Little knowledgeable = 2	569	37.46	50.03
	Moderately knowledgeable = 3	477	31.40	81.44
	Knowledgeable = 4	226	14.88	96.31
	Very knowledgeable = 5	56	3.69	100.00
	Total	1519	100.00	

Note: Table 6 displays the distribution regarding the level of financial literacy, self-assessment and the cognitive reflection test. "Freq." stands for frequency and "Cum" for Cummulative

Table 7 displays the differences when comparing this research result's with other results with a Portuguese sample (all other studies are representative of the population). This sample presents more optimistic results regarding concepts as compounding interest and relationship between fixed interest bond and market interest rates, however in concepts as inflation, relationship between risk and return and risk and diversification this sample underperforms. Since this research does not guarantee representativeness in terms of population, and is therefore biased in the type of inquiries, this comparison enlightens possible differences within both samples. Thus, the interpretation and results

must be handled carefully.

Table 7: Summary table regarding the percentage of Portuguese individuals who know financial questions across different studies.

Questions	This research	PNFF 2015*	OECD 2015**	PNFF 2020***
Compounding interest	52,05	39,5	41	31
Inflation	66,16	87	87(a)	74,4
Relationship between risk and return	75,41	81	82	71,6
Relationship between risk and diversification	63,62	72,4	73	45,1
Structured product	50,27	-	-	-
Relationship between fixed interest bond and market interest rates	29,24	6,5	-	-

Note: Table 7 displays the percentage of correctness responses regarding other studies with Portuguese individuals. Not all studies embody all question therefore, those that did not study specific concepts are represented with a "-". * Plano Nacional de Formação Financeira (2016). ** OECD (2016). *** Plano Nacional de Formação Financeira (2021). (a) the inflation question in OECD (2016) is not worded similarly.

4.4. Internal Consistency analysis

An internal consistency analysis measures how several variables, that represent questions, are intertwined in order to assess a certain construct. Therefore, when computing a score based on several questions, these shall be related, as referred in Bryman and Bell (2007). There are various ways to perform an internal consistency analysis depending on the type of questions/answers presented in the questionnaire (Coutinho., 2013). Cronbach´s alpha is the most common used method, according to Hill and Hill (2002) and it is advised for Linkert or rating scales, and dichotomous and short answers. According to Bryman and Bell (2007), it is considered a good internal consistency when Cronbach´s alpha reaches 0.80 (since it varies from 0 to 1, being 1 the highest internal reliability). It should be noted that the number of questions used influences the value of this indicator (more items increase the alpha value, as pointed out in Hill & Hill, 2002) and the reference value changes depending on the author adopted. In this research, the Cronbach's alpha using all variables (except for sociodemographic variables) is 0.8147, presenting a good internal consistency according to Bryman and Bell (2007) and Hill and Hill (2002).

Alpha Cronbach	Number of items.
0.8147	16

4.5. Multicollinearity analysis

Multicollinearity is observed, that means that two variables have precisely a zero (perfect multicollinearity) or almost zero (imperfect multicollinearity), linear combination, according to Daniels and Minot (2018). As stated in Gujarati and Porter (2009), multicollinearity (or imperfect multicollinearity, since perfect multicollinearity is rare) is associated with OLS estimators presenting wide variances/covariances, despite the estimators remaining BLUE. As a consequence, the interval of confidence also presents larger intervals which results in a more likely approval of the null hypothesis. The coefficients tend to be non-statistically significant, even though the overall goodness of fit (R²) may be high. Finally, the standard errors and the coefficients can sensitively vary as a result of minor data alterations. In order to test multicollinearity, first it is constructed a correlation matrix (as advised in Daniels & Minot, 2018). At first instance, this matrix allows us to directly observe if there are two variables that measure the same construct, as explained in , which results in high levels of correlation between them. Since this matrix only allows for a correlation measurement between two variables, the inclusion of Variance Inflation Factor (VIF) , displays the correlation between one variable and all others. Despite this last method not having a consensual value, there is a generic rule of thumb, namely values of VIF that are superior to 10, multicollinearity problems may be present (Daniels & Minot, 2018; Gujarati & Porter, 2009).

In table 8, it is constructed a Pearson correlation matrix, in order to assess if there are bilateral high correlation problems. Daniels and Minot (2018) affirm that if in this correlation matrix, if there are correlations superior to 0.8, multicollinearity issues may appear. Since the variable *agesq*, results from the square of the variable *age*, these would have high correlation, thus *agesq* is not present in the matrix since it does not provide relevant additional information. Correlations between *finbehav* and *finknow* and *prct_sec* and *marketpart*, present dangerously high correlations, but they do not cause issues since *finknow* and *finbehav* are used as alternative measures of financial literacy and except for the first level of *prct_sec* all other levels assume that the individual participates in the market, thus these variables are not used simultaneously since they do not provide additional data. In this matrix, except for these variables, there are no bilateral coefficient correlation with values near the threshold presented by Daniels and Minot (2018).

However, table A.1 (present in the appendix section) presents a different scenario and shows the difference of assessing multicollinearity only through bilateral correlations and when assessing it with the interaction with the other predictor variables. Table A.1 presents two different possibilities: the first

one comprises all predictors of model 1 including *education* and the second column removes *education*. The first column displays a possible problematic issue, derived from the large mean VIF. When evaluating its components, immediately the variable *education* stands out. It displays unsettling values for its categories, that in far exceed the threshold of 10, as previously stated. Thus, when ignoring this variable and moving on to the second column, the mean VIF significantly reduces. Interestingly, even though *education* with another variable did not display dangerously high correlations in the pearson correlation matrix, when computing the variance inflation factor, *education* presents itself as a problematic variable. Thus removing *education* removes possible multicollinearity problems.

The other models are also tested in terms of VIF, but since all models use the same independent variables there is no need to present the result 's table since the VIF will not change.

Table 8: Pearson Correlation Matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) finknow	1.000										
(2) perceived	-0.395***	1.000									
(3) finbehav	0.904***	-0.380***	1.000								
(4) de	0.085***	-0.045*	0.081***	1.000							
(5) de1	0.023	0.041	0.031	0.197***	1.000						
(6) loss	-0.139***	0.161***	-0.117***	0.033	0.091***	1.000					
(7) overc	-0.647***	0.001	-0.533***	-0.087***	-0.086***	-0.023	1.000				
(8) exper	0.255***	-0.148***	0.360***	0.009	0.017	-0.037	-0.081***	1.000			
(9) crt	0.583***	-0.234***	0.530***	0.031	0.015	-0.074**	-0.231***	0.163***	1.000		
(10) portf	0.274***	-0.287***	0.486***	0.002	-0.048*	-0.123***	-0.063**	0.391***	0.180***	1.000	
(11) marketpart	0.293***	-0.248***	0.468***	0.046*	-0.036	-0.107***	-0.092***	0.349***	0.190***	0.557***	1.000
(12) prct_sec	0.294***	-0.317***	0.425***	-0.015	-0.085***	-0.110***	-0.034	0.284***	0.241***	0.593***	0.795***
(13) numeracy	0.513***	-0.162***	0.471***	-0.023	-0.102***	-0.001	-0.161***	0.128***	0.560***	0.134***	0.150***
(14) riskprofile	0.164***	-0.275***	0.163***	-0.016	-0.160***	-0.366***	0.093***	0.068**	0.092***	0.201***	0.183***
(15) gender	0.326***	-0.314***	0.314***	0.011	-0.074***	-0.195***	-0.112***	0.144***	0.271***	0.225***	0.239***
(16) education	0.213***	-0.087***	0.310***	-0.023	0.070**	-0.052*	-0.115***	0.292***	0.138***	0.259***	0.318***
(17) studyfield	0.281***	-0.386***	0.218***	0.035	-0.007	-0.124***	-0.053**	0.022	0.198***	0.064***	0.005
(18) occupation	-0.139***	0.006	-0.273***	0.003	-0.052*	0.017	0.100***	-0.294***	-0.032	-0.237***	-0.322***
(19) age	0.168***	0.011	0.302***	-0.003	0.096***	-0.002	-0.140***	0.398***	0.053**	0.279***	0.332***
(20) metro	0.085***	-0.101***	0.091***	-0.023	-0.027	-0.053*	-0.013	0.087***	0.061***	0.107***	0.080***
(21) income	0.189***	-0.093***	0.316***	-0.009	0.049*	-0.013	-0.089***	0.358***	0.111***	0.344***	0.353***

Variables	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
(1) finknow										
(2) perceived										
(3) finbehav										
(4) de										
(5) de1										
(6) loss										
(7) overc										
(8) exper										
(9) crt										
(10) portf										

(11) marketpart										
(12) prct_sec	1.000									
(13) numeracy	0.199***	1.000								
(14) riskprofile	0.253***	0.072***	1.000							
(15) gender	0.294***	0.218***	0.300***	1.000						
(16) education	0.204***	0.120***	-0.092***	0.061***	1.000					
(17) studyfield	0.070***	0.172***	0.195***	0.135***	-0.082***	1.000				
(18) occupation	-0.179***	-0.024	0.085***	-0.076***	-0.534***	0.177***	1.000			
(19) age	0.187***	0.005	-0.163***	0.114***	0.547***	-0.233***	-0.656***	1.000		
(20) metro	0.103***	0.047**	0.031	0.090***	0.115***	0.030	-0.072***	0.043*	1.000	
(21) income	0.248***	0.063***	-0.049*	0.133***	0.454***	-0.081***	-0.460***	0.583***	0.089***	1.000

Note: Table 8 displays the Pearson correlation matrix for all variables present in this research, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

5. Empirical Results

In this section, several hypotheses are going to be tested, in order to meet the previous presented goals. The default interval of confidence that is used is 95% through all tested hypothesis even though it is also presented 90% and 99%. All regressions use robust standard deviations in order to mitigate possible heteroskedasticity.

5.1 Financial Literacy

Before analyzing which factors affect and dictate different financial literacy levels, table 9 demonstrates the financial literacy level distribution. The most common level is “Very High” with 36.49% of the individuals correctly scoring at least 5 questions. Despite this positive scenario, the second financial literacy level with the most individuals is actually the lowest level (“Very low”), as a consequence of the high number of inquiries that do not answer any question correctly. This upsetting result may derive from non-answered questions, and since the questions were not of obligatory nature, it is not possible to distinguish between not knowing and not wanting to answer. The majority of the individuals (69.84%) is able to correctly respond at least half of the presented questions.

Table 9: Frequency distribution of *finknow*

	Freq.	Percent	Cum.
Very low	427	23.08	23.08
Low	131	7.08	30.16
Medium	259	14.00	44.16
High	358	19.35	63.51
Very high	675	36.49	100.00
Total	1850	100.00	

Note: Table 9 displays the frequency table for the *finknow* variable. “Freq.” stands for frequency and “Cum” for Cumulative

Table 11 displays the marginal effects of an ordered probit model with *finknow* as dependent variable. It should be noted that *finknow* is the first measure of financial literacy that is going to be used, being the rest of the measures tested in the last hypotheses (H6). This variable varies between 1 and 5, being five the highest level. Initially it is presented a reduced model (present in table 10) that becomes broader as new dimensions are added (table 11), allowing a comparison of how the marginal effects of the initial variables fluctuate with the addition of newer variables. Lastly, it is also tested an interaction term between the variables *gender* and *overc*, in order to disentangle the individual effect of each in financial literacy levels. In order to make the reading lighter, the results of the interaction term are discussed in this section, however its marginal effects table is available in the appendix area (table A.2).

Firstly, it is presented a reduced model with the variables that are considered, according to literature, promising at explaining different levels of financial literacy. Indeed, all presented variables seem strongly related to financial literacy, except for numeracy, which is not statistically significant for any level of financial literacy. This lack of statistical significance is explained by the addition of the study field variable, that captures the numeracy effect (numeracy is statistically significant if *studyfield* is not added to the model). Experience portrays a positive relationship with financial literacy, similarly to Hilgert et al. (2003) and Moore (2003). Moreover an additional level in the inquiries' experience, on average increases the likelihood of the individual top-scoring (presenting a very high financial literacy level) by 16.11% and decreases the likelihood of having any level inferior to "Very High", when holding constant other variables. The investment importance in overall wealth is also relevant since as the individual increases one level in the investment weight (and as a result becomes an investor), the individual on average is less likely to present a very low level (by 1.35%) and more likely to display a very high level (by 8.44%). This may indicate that individual's financial literacy increases as a result of the learning process associated with the time length of owning financial products and intensity at investing. There is also other possible explanation: individuals in order to dive in into financial products and certain amounts at investing feel the need to reach higher levels of financial literacy in order to do so. The direction of the relationship between these variables is not studied, thus this question remains unanswered – this endogeneity problem, especially when it comes to experience is widely pointed out in literature. It is also observable that the individual academic background also matters since individuals from mathematics/engineering and economic/management areas are more likely to present higher levels of financial literacy when compared with individuals from non-mathematical and non-economic areas (Sebastião et al., 2021, and Van Rooij et al., 2011, display the same connection). However, even though mathematical/engineering background displays interesting results, those that come from an economic/management area have upper advantage (on average, these individuals increase their likelihood of having very high levels of financial literacy by 28.30%, *ceteris paribus*). This positive result for individuals with mathematical/engineering study field may be explained by the link between numeracy or mathematical capacity with financial literacy (as pointed out in the literature), but also by the fact that some of the engineering curriculum plans are composed by financial management units. Lastly, behavioral, and cognitive aspects are also determinants when explaining financial literacy. In terms of behavioral biases, overconfidence presents a negative relationship with financial literacy, therefore demonstrating that additional points in the overconfidence scale, increases the probability of the individual scoring very low, low and medium levels and decreases the likelihood of top-scoring

substantially (for example, one additional unit of overconfidence, on average, makes the individual 33.40% less likely to present a very high level of financial literacy, *ceteris paribus*). Even at a different direction, Thaler and Sunstein (2008) and Takeda et al. (2013) also display a negative association with these two variables. Loss aversion also displays a negative relationship, therefore demonstrating that a unit increase in loss aversion, decreases the likelihood of scoring highly (this finding contradicts the positive relationship presented by Gerth et al., 2021). Relatively to cognitive aspects, cognitive reflection presents a positive relationship, thus demonstrating its positive impact on individuals' financial literacy levels. More specifically, an additional point in the cognitive reflection test, on average increases the likelihood displaying very high level of financial literacy by 9.27%, *ceteris paribus* (in accordance with Muñoz-Murillo et al., 2020).

Table 10: Marginal effects of ordered probit with *finknow* as dependent variable

	Very Low	Low	Medium	High	Very High
exper	-0.0258*** (0.0045)	-0.0279*** (0.0054)	-0.0641*** (0.0105)	-0.0434*** (0.0073)	0.1611*** (0.0252)
studyfield					
Mathematics, Engineering	-0.0216*** (0.0060)	-0.0280*** (0.0082)	-0.0575*** (0.0168)	-0.0109* (0.0058)	0.1180*** (0.0344)
Economics, Management	-0.0412*** (0.0049)	-0.0539*** (0.0074)	-0.1274*** (0.0134)	-0.0605*** (0.0070)	0.2830*** (0.0232)
prct_sec	-0.0135*** (0.0022)	-0.0146*** (0.0026)	-0.0336*** (0.0050)	-0.0227*** (0.0034)	0.0844*** (0.0117)
riskprofile	-0.0041** (0.0020)	-0.0044** (0.0021)	-0.0101** (0.0049)	-0.0069** (0.0033)	0.0255** (0.0121)
overc	0.0534*** (0.0046)	0.0578*** (0.0062)	0.1328*** (0.0098)	0.0900*** (0.0072)	-0.3340*** (0.0138)
loss	0.0052** (0.0020)	0.0056** (0.0023)	0.0129*** (0.0049)	0.0087** (0.0034)	-0.0324*** (0.0125)
numeracy	-0.0061 (0.0037)	-0.0066 (0.0040)	-0.0151* (0.0090)	-0.0102* (0.0062)	0.0380* (0.0227)
crt	-0.0148*** (0.0025)	-0.0160*** (0.0028)	-0.0369*** (0.0057)	-0.0250*** (0.0042)	0.0927*** (0.0136)
Observations	1,002	1,002	1,002	1,002	1,002

Notes: Table 10 demonstrates the marginal effect of ordered probit with *finknow* as dependent variable. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

As mentioned, table 11 presents a broader regression that includes sociodemographic variables to the previous regression in table 10. With the inclusion of new variables, there are not much fluctuations,

except for some magnitudes as it is expected, except for the risk profile of the individual that becomes non-statistically significant. In terms of sociodemographic variables that so far have not been introduced, individual's age is not relevant in explaining any level of financial literacy. This lack of relevance from age at explaining financial literacy levels contradicts several research (Abreu & Mendes, 2010; Hung et al., 2009; Van Rooij et al., 2011) and the thought of a non-linear pattern introduced by Lusardi and Mitchell (2011). Whereas gender matters since men are more likely to present a very high level of financial literacy when compared to female individuals. Inclusively, being male increases on average 13.61% of displaying a very high level. This male advantage is also present in Abreu and Mendes (2010); Hung et al. (2009); Lusardi and Mitchell (2011); OECD (2016); Martins et al. (2021); Sebastião et al. (2021); Potrich et al. (2015); Van Rooij et al. (2011). Income also appears as a relevant factor, favoring higher levels of income - one additional level of income, on average, makes the individual 1.81% more likely to top-score, *ceteris paribus*. This indicates that those that have lower levels of income are more likely to have lower levels of financial literacy (similarly to Abreu & Mendes, 2010); Disney & Gathergood, 2013; Hung et al., 2009; Sebastião et al., 2021). Lastly, living in a metropolitan area positively affects financial literacy, since these individuals are on average 5.87% more likely to score the highest (Sebastião et al., 2021 display the same conclusion).

Table 11: Marginal effects of ordered probit with *finknow* as dependent variable

	Very Low	Low	Medium	High	Very High
exper	-0.0161*** (0.0046)	-0.0175*** (0.0053)	-0.0379*** (0.0107)	-0.0283*** (0.0080)	0.0998*** (0.0277)
studyfield					
Mathematics, Engineering	-0.0195*** (0.0065)	-0.0238*** (0.0085)	-0.0471*** (0.0171)	-0.0106* (0.0060)	0.1010*** (0.0364)
Economics, Management	-0.0430*** (0.0051)	-0.0540*** (0.0076)	-0.1234*** (0.0134)	-0.0683*** (0.0076)	0.2887*** (0.0238)
prct_sec	-0.0108*** (0.0020)	-0.0118*** (0.0023)	-0.0255*** (0.0043)	-0.0190*** (0.0032)	0.0672*** (0.0108)
riskprofile	-0.0021 (0.0019)	-0.0022 (0.0021)	-0.0048 (0.0046)	-0.0036 (0.0034)	0.0127 (0.0120)
overc	0.0505*** (0.0044)	0.0550*** (0.0061)	0.1191*** (0.0095)	0.0889*** (0.0072)	-0.3135*** (0.0140)
loss	0.0039** (0.0020)	0.0043* (0.0022)	0.0093** (0.0046)	0.0069** (0.0035)	-0.0244** (0.0121)
numeracy	-0.0052 (0.0038)	-0.0056 (0.0041)	-0.0122 (0.0087)	-0.0091 (0.0065)	0.0321 (0.0230)
crt	-0.0138***	-0.0150***	-0.0325***	-0.0243***	0.0856***

	(0.0025)	(0.0027)	(0.0055)	(0.0044)	(0.0136)
age	-0.0003	-0.0003	-0.0007	-0.0005	0.0017
	(0.0002)	(0.0002)	(0.0005)	(0.0004)	(0.0014)
gender	-0.0219***	-0.0239***	-0.0517***	-0.0386***	0.1361***
	(0.0038)	(0.0046)	(0.0084)	(0.0060)	(0.0205)
income	-0.0029**	-0.0032**	-0.0069**	-0.0051**	0.0181**
	(0.0013)	(0.0014)	(0.0030)	(0.0022)	(0.0078)
metro	-0.0093***	-0.0101***	-0.0219***	-0.0163***	0.0575***
	(0.0033)	(0.0036)	(0.0076)	(0.0057)	(0.0197)
occupation					
self-employee	-0.0071	-0.0079	-0.0174	-0.0144	0.0467
	(0.0087)	(0.0097)	(0.0222)	(0.0194)	(0.0599)
Employee	0.0000	0.0000	0.0000	0.0000	-0.0001
	(0.0049)	(0.0053)	(0.0115)	(0.0087)	(0.0303)
Unemployed	-0.0025	-0.0028	-0.0061	-0.0047	0.0161
	(0.0083)	(0.0092)	(0.0202)	(0.0161)	(0.0537)
Retired	0.0312	0.0328	0.0623*	0.0304***	-0.1568*
	(0.0214)	(0.0216)	(0.0355)	(0.0097)	(0.0868)
Other	-0.0001	-0.0001	-0.0002	-0.0001	0.0005
	(0.0070)	(0.0076)	(0.0165)	(0.0124)	(0.0435)
Observations	934	934	934	934	934

Notes: Table 11 demonstrates the marginal effect of ordered probit with *finknow* as dependent variable in a broader model. The variable *occupation* uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

5.1.1 Interaction term between gender and overconfidence

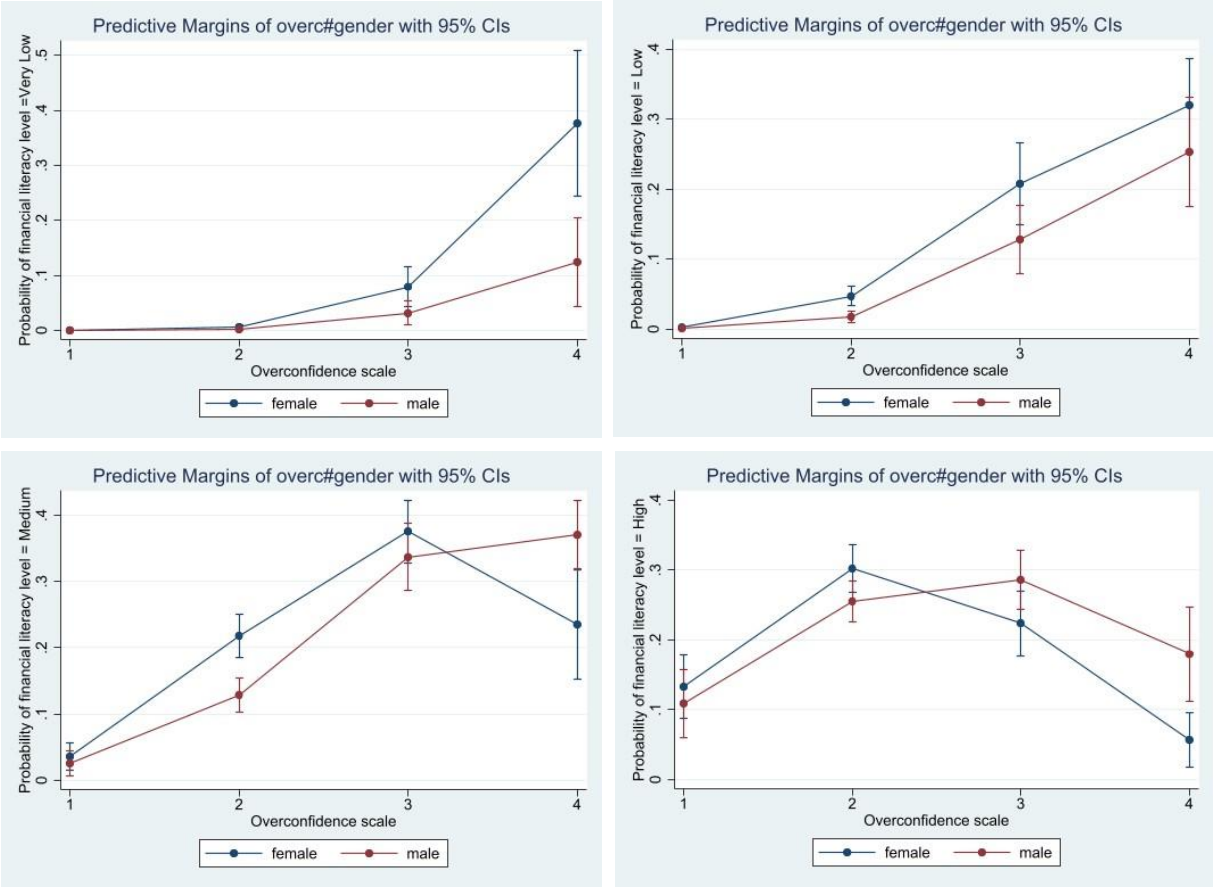
Finally, Lusardi (2012) states that men have better correct response rates in financial literacy and numeracy questions, and women opt for the "I don't know" option more frequently. It is hypothesized that female's lower level of financial literacy may be a representation of lower knowledge or lower confidence. Moreover, Bucher-Koenen et al. (2021) believe that specifically 38% of the divergence of financial literacy between genders may be justified by lack of confidence from women. Hence, in order to assess whether the lower levels of financial literacy among women may derive from lower levels of confidence when answering, it is included an interaction term between *gender* and *overc*.

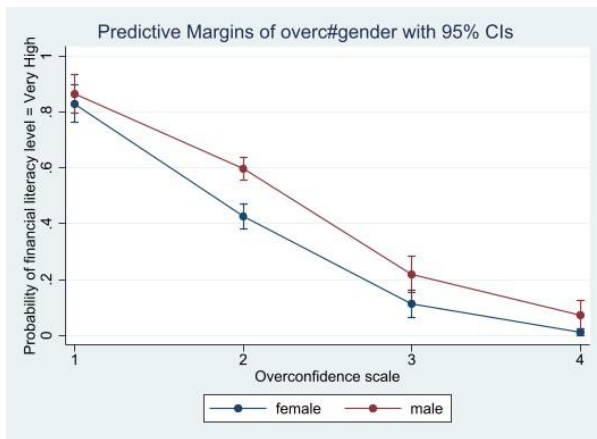
Figure 1 displays the marginal effects of the interaction term regarding all possible outcomes of the dependent variable. From the tables' observation, it is not possible to gather information regarding the statistical significance, thus table A.2 in appendix dives thoroughly into this issue. The likelihood of presenting a very low and low level of financial literacy is not statistically significant for individuals that are very underconfident, thus the probability of displaying these levels when both genders are very

underconfident shall not be interpreted. Another scenario that is also not interpreted regards the probability of presenting a very high level of financial literacy in which both genders are overconfident (since being female and overconfident is not statistically significant, it is not possible to compare it with male results).

Overall, women are more likely to display very low or low levels of financial literacy, even if both genders are overconfident (in the case of the “low” level and both genders are overconfident, the confidence intervals cross). In the medium and high levels, it is observable a shift between the leading gender – with lower levels of overconfidence (as it is the case of being very underconfident and underconfident) female individuals are more likely to position in this level, however once both genders are neutral in terms of overconfidence, male individuals present superior probability. Regarding the highest level of financial literacy, in general there are no substantial differences between both genders (men still display upper advantage), especially when both genders are very underconfident. Even if both genders are underconfident or neutral, male individuals appear to be more likely to present a very high level of financial literacy.

Figure 1: Marginal effects of interaction term gender and overc





Notes: Figure 1 displays the likelihood of achieving each level of *finknow* depending on the combination of gender and *overc*. It is a representation of the marginal effects of the several combinations between these two variables.

Overall, the first subset of hypothesis 1 (H1.1) fails at rejecting the positive association between experience and investment importance, however the positive relationship with individual's risk profile is not observable, since this variable is not statistically significant at explaining financial literacy. The following subset (H1.2) composed by cognitive and behavioral aspects, rejects the positive association with both behavioral biases (loss aversion and overconfidence), since these two factors negatively affect financial literacy. Whereas, cognitive aspects, as it is the case of numeracy and cognitive reflection test, increase the likelihood of reaching higher levels of financial literacy, as hypothesized. Lastly, sociodemographic factors are also determinant and the third subset (H1.3) fails at rejecting the positive relationship with gender, income, study field and residence. The hypothesized U-shaped inverted pattern observable in age, is rejected since age is not statistically significant.

5.2 Market Participation

Worldwide it is debated the low rates of market participation by individuals, thus this hypothesis intends to gather further information in order to comprehend what may justify this lack of willingness to participate. Table 12 displays three models, it starts from the simplest model (model 1), where it is included financial literacy (under the form of financial knowledge), risk profile, income, and age. Later on it is added both behavioral biases (loss aversion and overconfidence) for the second model. The third model includes the rest of cognitive and sociodemographic variables.

The first model shows that additional levels in the individual's risk profile, on average increases by 10.18% the likelihood of participating in the market, keeping other variables constant. Financial literacy, and income also demonstrate being positively associated, therefore, higher levels of these

variables translate into a higher willingness to market participation. This positive relationship between financial literacy and market participation is also observable in other authors (Klapper et al.2013; Sebastião et al., 2021; Yamori & Ueyama, 2021), as well as the positive association with income observable in Xia et al. (2014). An additional year in individual’s age on average increases the likelihood of participating in the market by 0.73%, *ceteris paribus*. With the inclusion of behavioral biases (model 2), as it is the case of loss aversion and overconfidence, there are no substantial differences. Loss aversion is not capable of explaining market participation (as opposed to Yang, 2019), contrarily to overconfidence that displays a positive relationship (as observable in Xia et al., 2014). Thus, an additional level in the overconfidence scale, on average, increases the likelihood of participating by 9.35%, *ceteris paribus*. The positive direction of *overc* indicates that as the gap between self-assessed knowledge and actual knowledge increases, the individual increases its willingness to participate. Thus, the perception of financial knowledge is as important when explaining the likelihood of participating in the market. Lastly, with the inclusion of other sociodemographic, an additional level in financial literacy increases its positive impact in market participation. Curiously, individuals with economic/management background are less likely to participate in the market, when compared to individuals with non-economic and mathematic background. Retired and employees are also more likely to participate when compared to students.

Table 12: Marginal effects of probit with *marketpart* as dependent variable

	M1	M2	M3
riskprofile	0.1018*** (0.0132)	0.0938*** (0.0156)	0.0883*** (0.0159)
finknow	0.0837*** (0.0104)	0.1256*** (0.0146)	0.1435*** (0.0177)
income	0.0544*** (0.0084)	0.0492*** (0.0090)	0.0378*** (0.0090)
age	0.0073*** (0.0011)	0.0077*** (0.0012)	0.0033** (0.0016)
loss		-0.0001 (0.0166)	-0.0055 (0.0163)
overc		0.0935*** (0.0254)	0.1189*** (0.0257)
studyfield			
Mathematics, Engineering			-0.0356 (0.0445)
Economics, Management			-0.1327*** (0.0328)

numeracy	0.0494		
	(0.0315)		
crt	0.0255		
	(0.0203)		
gender	0.0267		
	(0.0284)		
occupation			
self-employee	0.1717**		
	(0.0821)		
Employee	0.1721***		
	(0.0429)		
Unemployed	-0.0459		
	(0.0814)		
Retired	0.4200***		
	(0.1158)		
Other	0.0159		
	(0.0720)		
metro	-0.0224		
	(0.0267)		
Observations	1,208	1,073	1,057

Notes: Table 12 demonstrates the marginal effect of probit with *marketpart* as dependent variable. The variable *occupation* uses “ Student” as reference, and the variable *studyfield* uses “non-mathematic and non-economic background” as reference. Model 1 is the simplest model, model 2 includes loss aversion and overconfidence, and model 3 adds the rest of sociodemographic variables. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Overall, these results fail to reject the positive relationship with financial literacy, risk profile, income, and overconfidence, However, the negative association between age or loss aversion and market participation is rejected, since age presents a positive relationship and loss aversion is not statistically significant.

5.3 Portfolio composition

As highlighted in Abreu and Mendes (2010), using the number of financial products as a measure of diversification is a rather simplistic proxy, since it does not consider the correlation between the different held products, and as a result it neglects an important part of the diversification concept. It should be acknowledged that this proxy may simplify this concept and not be able to portray the true level of diversification. As observable in table 13, portfolio composition is transformed into five different levels (very low, low, medium, high and very high). The number of different financial products that presents more supporters is only one financial product, therefore the grand majority of the individuals not even hold one tenth of the listed financial products. This poor result is observable throughout the

distribution table that displays a highly positive skewness – few individuals have high number of different products, inclusively the average number is 1.74. Moreover, there is no individual that matches the fifth and highest level of portfolio composition.

Table 13: Frequency distribution of the number of different financial product

	Freq.	Percent	Cum.
0	197	10.65	10.65
1	841	45.46	56.11
2	399	21.57	77.68
3	212	11.46	89.14
4	114	6.16	95.30
5	46	2.49	97.78
6	29	1.57	99.35
7	7	0.38	99.73
8	3	0.16	99.89
9	2	0.11	100.00
Total	1850	100.00	

Note: Table 13 displays the frequency table. “Freq.” stands for frequency and “Cum” for Cummulative.

The following table (table 14) presents the marginal effects for *portf* – dependent variable that measures portfolio composition in regard to the number of distinct financial products. The reduced model displays a positive relationship between *portf* and financial literacy, risk profile, investment importance and experience. Thus additional levels in each one of these variables, increases the likelihood of individuals presenting medium levels of portfolio composition. Contrarily to what is hypothesized in regard to behavioral biases, neither loss aversion nor overconfidence are statistically significant at explaining different levels of portfolio composition. There is no variable capable of explaining the highest level of *portf*, since none of them display statistical significance.

Table 14: Marginal effects of probit with *portf* as dependent variable – reduced model

	very low	low	medium	high
finknow	-0.0723*** (0.0139)	0.0548*** (0.0106)	0.0150*** (0.00359)	0.00245* (0.00145)
riskprofile	-0.0307*** (0.0109)	0.0233*** (0.00839)	0.00636** (0.00248)	0.00104 (0.000634)
prct_sec	-0.116*** (0.00792)	0.0879*** (0.00646)	0.0240*** (0.00392)	0.00393* (0.00215)
exper	-0.196*** (0.0220)	0.149*** (0.0177)	0.0407*** (0.00771)	0.00667* (0.00353)
loss	0.00952 (0.0117)	-0.00722 (0.00893)	-0.00197 (0.00239)	-0.000323 (0.000470)
overc	-0.0101	0.00770	0.00210	0.000345

	(0.0209)	(0.0158)	(0.00428)	(0.000807)
Observations	1,002	1,002	1,002	1,002

Notes: Table 14 demonstrates the marginal effect of ordered probit with *portf* as dependent variable. The fifth category is not present derived from lack of individuals who match the criteria. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

With the addition of new variables in table 15, mainly cognitive and sociodemographic, the results do not change substantially. An additional point in financial literacy, on average decreases the likelihood of presenting a very low level of portfolio composition (by 5.91% more precisely) and increases the probability of displaying a medium level (by 1.36%), *ceteris paribus*. This positive impact of financial literacy in portfolio diversification is also pointed out in Abreu and Mendes (2010). The same positive impact appears in risk profile, since an additional level (towards being risk taker), on average makes the individuals 4.57% less likely to have very low levels of portfolio composition and 1.05% more likely to have medium level. The investment importance and experience also positively influence portfolio composition, as well as income when assessing the medium level of portfolio composition. An additional level in income, on average makes the individual 0.687% more likely at presenting a medium level when it comes to the number of financial products.

Table 15: Marginal effects of probit with *portf* as dependent variable

	very low	low	medium	high
finknow	-0.0591*** (0.0158)	0.0434*** (0.0115)	0.0136*** (0.00420)	0.00213* (0.00127)
riskprofile	-0.0457*** (0.0107)	0.0336*** (0.00803)	0.0105*** (0.00293)	0.00164* (0.000885)
prct_sec	-0.112*** (0.00820)	0.0826*** (0.00667)	0.0258*** (0.00412)	0.00405* (0.00212)
exper	-0.105*** (0.0268)	0.0770*** (0.0199)	0.0240*** (0.00718)	0.00377* (0.00209)
loss	0.00157 (0.0108)	-0.00115 (0.00794)	-0.000360 (0.00247)	-5.65e-05 (0.000394)
overc	-0.000223 (0.0201)	0.000164 (0.0147)	5.13e-05 (0.00460)	8.04e-06 (0.000723)
income	-0.0299*** (0.00745)	0.0220*** (0.00552)	0.00687*** (0.00193)	0.00108* (0.000641)
studyfield				
Mathematics, Engineering	-0.0509 (0.0343)	0.0368 (0.0247)	0.0121 (0.00846)	0.00197 (0.00173)
Economics, Management	-0.00622 (0.0245)	0.00463 (0.0183)	0.00138 (0.00543)	0.000204 (0.000811)
crt	0.0202	-0.0149	-0.00465	-0.000729

	(0.0165)	(0.0121)	(0.00387)	(0.000748)
numeracy	0.0309	-0.0227	-0.00709	-0.00111
	(0.0263)	(0.0193)	(0.00603)	(0.00119)
gender	-0.0257	0.0189	0.00590	0.000925
	(0.0213)	(0.0157)	(0.00492)	(0.000916)
age	0.000119	-8.77e-05	-2.74e-05	-4.30e-06
	(0.00120)	(0.000880)	(0.000275)	(4.35e-05)
metro	-0.0120	0.00880	0.00275	0.000431
	(0.0198)	(0.0145)	(0.00456)	(0.000711)
occupation				
self-employee	-0.0336	0.0267	0.00614	0.000754
	(0.0531)	(0.0419)	(0.00987)	(0.00143)
Employee	-0.0904***	0.0698***	0.0181***	0.00252
	(0.0317)	(0.0256)	(0.00605)	(0.00157)
Unemployed	-0.137**	0.103**	0.0292*	0.00452
	(0.0607)	(0.0437)	(0.0154)	(0.00358)
Retired	-0.153*	0.114*	0.0334	0.00535
	(0.0849)	(0.0607)	(0.0212)	(0.00475)
Other	-0.0841	0.0651	0.0167	0.00229
	(0.0598)	(0.0450)	(0.0132)	(0.00234)
Observations	934	934	934	934

Notes: Table 15 demonstrates the marginal effect of ordered probit with *portf* as dependent variable. The variable *occupation* uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Overall, the positive association between the number of distinct financial products and financial literacy, investment importance and risk profile is not rejected, contrarily to the negative hypothesized relationship with loss aversion and overconfidence that is rejected, since these variables are not statistically significant.

5.4 Gap between hypothetical and current portfolio

As evidenced in table 16, when individuals are asked to spread 100 00 € in a hypothetical portfolio, the financial product that presents the highest average is deposits with an average of roughly 37 thousand euros, representing more than a third of the proposed initial value. The financial products that have the least adherence in terms of amount is gold, funds and bitcoin and digital coins. Despite bitcoin not presenting extremely high values in terms of average, it has one of the highest standard deviations, displaying a high volatility of presented amounts. When assessing the gap between hypothetical behavior and actual behavior, it would be expected that the majority of the individuals would have a positive gap (hypothetical behavior>actual). However, the high number of individuals that have

much more products currently than in the hypothetical portfolio, may be explained by individual 's want to drop certain products in their hypothetical portfolio, whereas they cannot do that currently in their actual portfolio. But it may derive from tiredness when answering. This question was the last question of the questionnaire, in which it asked individuals to divide 100 000€ through multiple products, and therefore this question required much more mental strength (as a result of having to decide which products to choose, which amounts for each of them and still remembering and accounting that the sum of these products should be 100 000€). This need of more focus and energy when answering, and since it is already the last question after the completion of the whole questionnaire, individuals may have neglected it and simplified the task in terms of number of products to make the summation simpler. This hypothesis seems the most probable scenario since the actual number of individuals who displayed the negative gap and responded other questions is low.

Therefore, in the following paragraphs, it is going to be tested the positive side of this gap. In table 17, the category of *pgap* in which individuals are more located is the third category, with a “medium” gap. Overall there are a relatively substantial number of individuals who present a great difference between behavior and preferences, demonstrating a higher number of products (or portfolio diversification), when compared to what they currently hold.

Table 16: Summary Statistics of the components of the hypothetical portfolio

Variable	Obs	Mean	Std. Dev.	Min	Max	Perc.
hypoth dep	1034	37 180.765	30 509.799	0	100000	50.59%
hypoth bonds	893	16 026.607	17 618.316	0	100000	37.62%
hypoth sustfund	876	12 816.22	14 542.988	0	100000	37.14%
hypoth othfund	753	9 004.807	13 866.723	0	100000	24.54%
hypoth 15stock	910	17 068.143	15 891.262	0	100000	41.78%
hypoth knownstock	833	11 845.654	13 156.551	0	100000	35.68%
hypoth gold	768	8 660.247	12 529.636	0	100000	27.41%
hypoth bitcoin	811	10 729.211	17 224.774	0	100000	29.41%

Notes: Table 16 demonstrates the summary statistics of the several products presented in the hypothetical portfolio. Std. dev stands for standard deviation, min for minimum, max for maximum and perc. for the proportion of individuals who allocated money (higher than 0 €) to each category. The column “ Perc.” percentages when summed exceed 100% derived from individuals' possibility to choose multiple products to allocate money.

Table 17: Summary Statistics of *pgap*

<i>pgap</i>	Freq.	Percent	Cum.
1	160	18.58	18.58
2	167	19.40	37.98
3	291	33.80	71.78
4	121	14.05	85.83
5	122	14.17	100.00
Total	861	100.00	

Notes: Table 17 demonstrates the frequency distribution of the *portf* variable. “Freq.” stands for frequency and “Cum”. for cumulative.

Table 18 demonstrates that financial literacy is not able to explain the positive gap observed between preferences and actual behavior, as well as the individual’s risk profile and loss aversion that appear non-statistically significant for any level of the dependent variable. Whereas the investment importance, experience and income appear to decrease the likelihood of high gap between preferences and actual behavior, and increase the likelihood of very few differences Thus, individuals that have low investment importance, experience and income are more likely to diversify more their hypothetical portfolio when compared to their actual current portfolio.

Table 18: Marginal effects of ordered probit with *pgap* as dependent variable – reduced model

	very low	low	medium	high	very high
finknow	0.0136 (0.00975)	0.00615 (0.00440)	-0.00100 (0.000984)	-0.00563 (0.00404)	-0.0131 (0.00946)
riskprofile	-0.0132 (0.0112)	-0.00597 (0.00505)	0.000975 (0.00112)	0.00547 (0.00462)	0.0128 (0.0108)
prct_sec	0.0610*** (0.00920)	0.0275*** (0.00478)	-0.00449 (0.00320)	-0.0252*** (0.00417)	-0.0589*** (0.00995)
exper	0.0962*** (0.0288)	0.0434*** (0.0135)	-0.00708 (0.00543)	-0.0397*** (0.0122)	-0.0927*** (0.0284)
loss	0.0140 (0.0116)	0.00633 (0.00522)	-0.00103 (0.00113)	-0.00580 (0.00484)	-0.0135 (0.0111)
income	0.0285*** (0.00604)	0.0129*** (0.00283)	-0.00210 (0.00157)	-0.0118*** (0.00269)	-0.0275*** (0.00587)
Observations	673	673	673	673	673

Notes: Table 18 demonstrates the marginal effect of ordered probit with *pgap* as dependent variable – reduced model. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 19 introduces a broader model, including sociodemographic and cognitive variables that are not present in the reduced model. With this addition, there is a substant change in the marginal effects of income that are no longer statistically significant at explaining any levels of the positive gap. An increase in the investment importance level, on average makes the individual 5.83% more likely to display a very low gap, and by 2.57% at presenting a low gap. It also decreases the likelihood of having a very high difference (by 5.67%), keeping all other variables constant. Experience is also important at explaining this yield between preferences and behavior since an additional level in the individual’s experience, on average increases the likelihood of having lower gap levels and decreases the probability of displaying high gaps. Regarding the new variables, overconfidence is not relevant at explaining the

positive gap and it is highlighted the negative relationship with age and gender, thus men and younger individuals are less likely to present substantial differences. Not many variables are capable at explaining this positive gap, possibly as a result of individuals not engaging thoroughly in this question, therefore it would be interesting exploring the difference between individuals' preferences and actual behavior.

Table 19: Marginal effects of ordered probit with *pgap* as dependent variable

	very low	low	medium	high	very high
finknow	0.00820 (0.0139)	0.00362 (0.00617)	-0.000536 (0.000995)	-0.00331 (0.00563)	-0.00797 (0.0136)
riskprofile	-0.0116 (0.0119)	-0.00511 (0.00527)	0.000757 (0.00100)	0.00467 (0.00480)	0.0113 (0.0116)
prct_sec	0.0583*** (0.00946)	0.0257*** (0.00469)	-0.00381 (0.00298)	-0.0235*** (0.00418)	0.0567*** (0.0101)
exper	0.0665** (0.0288)	0.0294** (0.0131)	-0.00435 (0.00382)	-0.0268** (0.0118)	-0.0647** (0.0285)
loss	0.0165 (0.0117)	0.00729 (0.00515)	-0.00108 (0.00117)	-0.00666 (0.00479)	-0.0161 (0.0113)
income	0.0104 (0.00717)	0.00461 (0.00319)	-0.000682 (0.000696)	-0.00421 (0.00292)	-0.0102 (0.00702)
overc	0.00307 (0.0193)	0.00136 (0.00853)	-0.000201 (0.00128)	-0.00124 (0.00779)	-0.00299 (0.0188)
studyfield					
Mathematics, Engineering	-0.0138 (0.0330)	-0.00591 (0.0143)	0.00125 (0.00292)	0.00557 (0.0133)	0.0129 (0.0313)
Economics, Management	-0.0185 (0.0262)	-0.00803 (0.0113)	0.00147 (0.00267)	0.00746 (0.0106)	0.0176 (0.0244)
crt	0.0202 (0.0146)	0.00891 (0.00648)	-0.00132 (0.00142)	-0.00814 (0.00593)	-0.0196 (0.0142)
numeracy	0.0133 (0.0246)	0.00587 (0.0109)	-0.000869 (0.00174)	-0.00536 (0.00995)	-0.0129 (0.0240)
gender	0.0449** (0.0206)	0.0198** (0.00902)	-0.00293 (0.00275)	-0.0181** (0.00833)	-0.0437** (0.0198)
age	0.00312** (0.00132)	0.00138** (0.000590)	-0.000204 (0.000182)	-0.00126** (0.000552)	0.00303** (0.00128)
metro	-0.0131 (0.0194)	-0.00578 (0.00851)	0.000855 (0.00146)	0.00528 (0.00782)	0.0127 (0.0188)
occupation					
self-employee	0.0621 (0.0694)	0.0264 (0.0254)	-0.00833 (0.0166)	-0.0258 (0.0280)	-0.0545 (0.0507)
Employee	0.0384 (0.0281)	0.0177 (0.0133)	-0.00327 (0.00372)	-0.0162 (0.0123)	-0.0366 (0.0261)
Unemployed	-0.0502	-0.0312	-0.0103	0.0216*	0.0701

	(0.0311)	(0.0207)	(0.0114)	(0.0127)	(0.0495)
Retired	-0.0852*	-0.0604	-0.0351	0.0343**	0.146
	(0.0511)	(0.0449)	(0.0465)	(0.0156)	(0.126)
Other	0.104*	0.0385**	-0.0211	-0.0415*	-0.0798**
	(0.0593)	(0.0171)	(0.0203)	(0.0220)	(0.0349)
Observations	656	656	656	656	656

Notes: Table 19 demonstrates the marginal effect of ordered probit with *overc* as dependent variable, but for a broader model. The variable *occupation* uses “Student” as reference, and the variable *studyfield* uses “non-mathematic and non-economic background” as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Overall, it is rejected the negative relationship with financial literacy, risk profile and income since these variables are not capable at explaining any relationship with *pgap*. It is also rejected the positive relationship of loss aversion and overconfidence since neither of these variables is statistically significant at explaining the positive gap. However, the negative association hypothesized with experience and investment importance is not rejected.

5.5 Disposition effect among investors

In order to study disposition effect, it is asked individuals what they would do depending on facing losses or gains, in regard to a reference point. This reference point is considered a key aspect when explaining disposition effect. However, and as previously pointed out by Da Costa et al. (2008), if this reference point no longer exists, then investors will see price changes, instead of gains and losses – this shift affects the existence of disposition effect. In order to assess if the removal of the reference point and substitution by price changes influences investor’s behavior, it is also questioned investor’s actions with price cuts and rises. The results are present at the end of this sub-section and the respective marginal effect table is displayed in the appendix area (table A.13).

5.5.1 Disposition effect with reference point

Table 20 demonstrates that if investors faced losses at the end of one year, the grand majority (75.96%) would have kept the investment for one more year. However, if facing the opposite scenario, 59.39% of the individuals would have still kept the investment for one more year. As evidenced in the second table only 26.75% of the investors engage in disposition effect.

Table 20: Frequency regarding the questions of *de* (as it is the case of *losses* and *gains*)

losses	gains		
	0	1	Total
0	309	168	477
	49.20	26.75	75.96
1	64	87	151
	10.19	13.85	24.04
Total	373	255	628
	59.39	40.61	100.00

de	Freq.	Percent	Cum.
0	460	73.25	73.25
1	168	26.75	100.00
Total	628	100.00	

Notes: Table 20 displays the cross tabulation between the variables that computed *de*. In each combination the second line represents the percentage by column.

In table 21, it is presented five probit models with disposition effect as dependent variable. The first model only assesses financial literacy, overconfidence, and loss aversion, the second model adds the investment importance, experience, and risk profile and the third model increments an income variable, and the fourth age and gender. The last model (M5) adds other cognitive and sociodemographic variables.

M1 does not present any relevant variable at explaining investor's engagement in disposition effect, thus financial, overconfidence, and loss aversion so far are not capable at explaining this financial behavior (contrarily to Baker et al., 2019; Feng & Seasholes, 2005; and Rau, 2014). The second, third and fourth model maintain this lack of statistical significance by financial literacy, overconfidence, and loss aversion, and additionally experience, investor's risk profile and income are also non-statistically significant. The investment importance (*prct_sec*) appeared relevant in the second model, displaying a negative relationship, therefore an additional level in the investment importance, on average, makes the investor less likely to engage in disposition effect by 4.34%, *ceteris paribus*. However, with the addition of income in the third model (M3), this negative relationship is no longer statistically significant since it was capturing part of the income effect. The third model also demonstrates how the statistical significance of overconfidence shifts once income is added – overconfidence becomes statistically significant with a negative relationship, and as new sociodemographic variables are also added, the effect becomes larger. Therefore, and as observable in M5, an additional level in the overconfidence scale, on average makes the investor 12.02% less likely to display disposition effect, *ceteris paribus*. Thus, overconfidence in part acts as a deterrent from disposition effect, as opposed to what the positive

relationship present in Chu (2012) and Ho(2021). Age and gender are not statistically significant, as observable in the fourth model. Regarding the rest of the added variables, the only relevant determinant is numeracy that presents an interesting role. Those that correctly respond the numeracy question, on average, decrease their probability of engaging in disposition effect by 15.87%, keeping other variables constant.

Table 21: Marginal effects of probit with de as dependent variable

	M1	M2	M3	M4	M5
finknow	-0.0275 (0.0244)	-0.0239 (0.0321)	-0.0292 (0.0328)	-0.0220 (0.0341)	-0.0089 (0.0369)
overc	-0.0651* (0.0340)	-0.0813* (0.0421)	-0.0868** (0.0427)	-0.0852** (0.0432)	-0.1202*** (0.0448)
loss	0.0137 (0.0214)	0.0141 (0.0253)	0.0098 (0.0257)	0.0095 (0.0260)	0.0010 (0.0255)
prct_sec		-0.0434** (0.0221)	-0.0405* (0.0227)	-0.0385* (0.0231)	-0.0288 (0.0230)
exper		-0.0247 (0.0446)	-0.0092 (0.0484)	0.0050 (0.0513)	0.0106 (0.0495)
riskprofile		-0.0216 (0.0246)	-0.0284 (0.0252)	-0.0235 (0.0266)	-0.0287 (0.0266)
income			-0.0082 (0.0121)	-0.0028 (0.0138)	-0.0005 (0.0136)
age				-0.0014 (0.0020)	-0.0015 (0.0023)
gender				-0.0576 (0.0500)	-0.0314 (0.0508)
studyfield					
Mathematics, Engineering					-0.0830 (0.0619)
Economics, Management					0.0621 (0.0521)
crt					-0.0559* (0.0334)
numeracy					-0.1577*** (0.0523)
metro					0.0239 (0.0424)
occupation					
self-employee					-0.0479 (0.0937)
Employee					-0.0147 (0.0602)
Unemployed					0.1093

					(0.1479)
Retired					-0.0712
					(0.1518)
Other					0.1833
					(0.1294)
Observations	562	429	418	414	410

Notes: Table 21 demonstrates the marginal effect of *de* with *de* as dependent variable. The variable *occupation* uses “Student” as reference, the variable *studyfield* uses “non-mathematic and non-economic background” as reference. The first model is a reduced one composed only by financial literacy and behavioral biases, whereas the following models include other variables. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Overall, the negative relationship between disposition effect and financial literacy, experience, age, gender and loss aversion is rejected, since neither of these variables is statistically significant. Overconfidence is statistically significant, but presents a negative relationship, contrarily to what is hypothesized, therefore the positive relationship with overconfidence is also rejected. To a better understanding of this bias, new variable must be added, since the opted ones are not enough to explain the occurrence of it. It may indicate widespread irrational behavior.

5.5.2 Disposition effect without reference point

Table 22 allows comparison between both measures of disposition effect and there are actually more individuals presenting disposition effect in the measure without reference point, than in the measure that incorporates it. There are more investors who did not engage in disposition effect in *de*, but afterward did in *de1*, rather than those that already engaged in *de* but when faced with *de1* did not display disposition effect – even though the difference is very minimal. Thus, this contradicts Da Costa et al. (2008) statement. This almost unnoticeable difference between disposition effect measures may result from the proximity in which these questions were presented, therefore when facing *de1* (the first presented measure was *de*), investors may have answered the same due to the similarity between questions.

Table 22: Frequencies regarding disposition effect with (*de*) and without reference point (*de1*)

de	de1		
	0	1	Total
0	362	97	459
	57.74	15.47	73.21
1	95	73	168
	15.15	11.64	26.79
Total	457	170	627
	72.89	27.11	100.00

Notes: Table 22 displays the cross tabulation between the variables that computed *de*. In each combination the second line

represents the percentage.

Table A.13 displays the use of *deI* as dependent variable. Financial literacy is once again not relevant at explaining this measure of disposition effect, as well as experience, investment importance gender, age and loss aversion (as observable in *de*). With this new dependent variable, overconfidence that appeared with a negative relationship with *de*, is no longer statistically significant. In terms of risk profile, it presents a negative relationship thus and additional level in the investor's risk profile, on average, decreases the likelihood of engaging in disposition effect (*deI*) by 7.10%, *ceteris paribus*. Additionally, having an economic/management background when compared with investors from non-mathematic and economic areas increases the probability of the investor presenting disposition effect by 15.31%, *ceteris paribus*. Retired investors, when compared to students, are less likely to engage in this bias (by 17.35%), keeping other variables constant. When comparing both measures, the main differences are the drop of overconfidence with a negative relationship with *de* to non-statistically significant with *deI*. However, this alternative measure presents other relevant variables that are not statistically significant with the default measure - investors who have an economic/management background are more likely to engage in disposition effect and retired individuals (when compared to students) are less likely to.

5.6 Robustness tests: different financial literacy measures

In order to make the comparison easier, between the different measures of financial literacy, this section directly compares the results from the most complete model of each hypothesis to understand how these results alter based on different measures. The tables with marginal effects of each hypothesis tested are in the appendix section. Firstly, it is introduced the frequencies of these new measures of financial literacy in order to illustrate how the individuals are distributed.

According to table 23, there are few individuals that meet the requirements of the highest level of *finbehav* (only 0.32%). Most individuals locate at the third category, therefore presenting a medium level. In terms of the perceived variable, the category with the highest number of individuals is High/High with 35.42% of the individuals, being the lowest category High/Low.

Table 23: Frequencies regarding alternative measures of financial literacy – *finbehav* and *perceived*

<i>finbehav</i>	Freq.	Percent	Cum.
Very Low	316	17.08	17.08
Low	413	22.32	39.41

Medium	712	38.49	77.89
High	403	21.78	99.68
Very High	6	0.32	100.00
Total	1850	100.00	

Perceived	Freq.	Percent	Cum.
High/High	656	35.42	35.42
Low/High	436	23.54	58.96
High/Low	379	20.46	79.43
Low/Low	381	20.57	100.00
Total	1852	100.00	

Notes: Table 23 displays the frequencies of the other two measures that are also used to assess financial literacy (*finbehav* and *perceived*). “Freq.” stands for frequency and “Cum.” for cumulative.

5.6.1 Financial Literacy

Dependent variable: *finbehav*

When comparing the most used worldwide measure (*finknow*) with a measure that includes financial knowledge but also behavior (*finbehav*), there are differences within the results. In table A.3, all effects that are statistically significant decrease their effect and their likelihood of predicting a level of financial literacy when using *finbehav* comparatively to *finknow*. This is observable in almost every variable. Other differences present in this new model are the fact that loss aversion and living in a metropolitan area are no longer relevant at explaining the financial literacy levels (as opposed to what happens with *finknow*). Additionally, *riskprofile* presents a positive relationship, demonstrating that an additional level in risk, on average increases the likelihood of having high and very high financial literacy by 2.57% and 0.20% respectively, ceteris paribus (when measuring financial literacy with only financial knowledge component, this variable is not statistically significant). Age is also relevant to this dependent variable (contrarily to what happens when studying *finknow*), demonstrating that an additional year in individual’s age decreases the likelihood of lower levels of financial literacy. As observable in table A.4, male individuals are more likely to display higher levels of financial literacy, and even if both genders are overconfident male individuals are still more likely (it cannot be compared the likelihood of displaying very high levels of financial literacy if both genders are overconfident due to lack of statistical significance).

Dependent variable: *perceived*

In table A.5, contrarily to *finbehav*, the use of *perceived* as a measure of financial literacy does not display lower magnitudes, on the contrary sometimes the marginal effects in specific variables are greater than the ones observable with the default measure of financial literacy. Similarly to *finknow* and *finbehav*, experience also displays a positive association, in this case with high financial knowledge and

high self-assess. This indicates that more experienced individuals not only are more likely to have higher financial knowledge, but also a higher vision of self-knowledge, that is also observable through the negative effect of belonging in the group with high financial knowledge but low self-assess. For most of the variables the categories High/High and High/Low display opposite signs, demonstrating that the perception of individuals financial knowledge is negatively affected by experience, investment importance, having an economic/management background, overconfidence, cognitive reflection test and gender. Similarly to *finbehav*, numeracy and loss aversion are also not statistically significant at explaining financial literacy, contrarily to what happens with the use of *finknow* (loss displays a negative relationship). Overconfidence displays a positive relationship with H/H category, contrarily to *finknow* and *finbehav* that present a negative association. In terms of sociodemographic factors, the main differences are the lack of statistical significance from *metro*, demonstrating that living in a metropolitan area is not relevant for this measure of financial literacy (similarly to *finbehav* and contrarily to *finknow*). Regarding the interaction term of overconfidence and gender (see table A.6), when both genders are very underconfident, female individuals are slightly more likely to be ranked in the High/High category, however as the overconfidence scale increases the gender gap rises, favoring men. In terms of the H/L category, except for when both genders are very underconfident, women are always more likely to locate in this category, therefore more likely to underestimate their knowledge.

5.6.2 Market participation

Independent variable: *finbehav*

As observable in table A.7, and similarly to the model with *finknow*, financial literacy, risk profile, income and overconfidence display a positive relationship therefore increasing the likelihood of participating those that present additional levels of these variables. The high magnitude of *finbehav* is expected since the definition of *finbehav* comprises a component of financial products ownership. It also does not display a statistically significant relationship between market participation and loss aversion, but in this model, age is also not statistically significant (in the model with *finknow*, age displayed a positive relationship contrarily to what is hypothesized).

Independent variable: perceived

In table A.7, perceived is an independent variable, and therefore it is used the last category (Low/Low) as the base reference level. For simplicity, it is displayed only the complete model. When compared with the default measure of financial literacy, overconfidence is no longer statistically

significant (contrarily to when using *finknow* and *finbehav*). Additionally, age is statistically significant, presenting a positive relationship – this did not happen when using *finbehav*. Interestingly, having a High/High or Low/High knowledge, as opposed to having Low/Low increases the probability of participating in the market, although the same does not happen if the individual has high knowledge but perceives himself as having low knowledge (High/Low), demonstrating the importance of individuals perception of their own abilities. The variable *overc* that is statistically significant with the use of *finknow* and *finbehav*, whereas when using *perceived* there is no longer statistically significant.

5.6.3 Portfolio Composition

Independent variable: *finbehav*

In table A.8, and similarly to the model with *finknow*, financial literacy and investment importance are statistically significant, however income and risk profile are not statistically significant (contrarily to what is observed in the model with *finknow*). The positive relationship between this measure of financial literacy and the level of portfolio composition was already expected of being positive, since this measure of financial literacy considers the number of different financial products (financial behavior). Loss aversion remains non-statistically significant, whereas overconfidence displays a positive relationship, therefore an additional level in the overconfidence scale makes the individual more likely by 1.12% of displaying a medium level of portfolio composition.

Independent variable: *perceived*

As observed in table A.9, this financial literacy measure presents similar results to *finknow*, except for overconfidence that appears more relevant in this model with a negative relationship, therefore an additional level in the overconfidence scale decreases the likelihood of presenting higher levels of portfolio composition. In the model with the default measure of financial literacy, overconfidence is not statistically significant. Financial literacy (through *perceived*) is positively linked with the number of financial products, however only if the individual displays simultaneously high financial knowledge and high perceived (H/H). These individuals on average increase the likelihood of having a medium level of diversification by 2.04%. But those that have high financial knowledge, but low perceived knowledge (H/L) do not display statistically significant relationship with any level of the dependent variable. Even though the use of *finknow* leads to the same positive relationship, this alternative measure of financial literacy portrays the importance of individuals self-perception. As with the default measure, financial literacy, risk profile, investment importance and income display a positive relationship with portfolio

composition, and loss aversion and is not statistically significant.

5.6.4 Gap between hypothetical and actual portfolio

Independent variable: *finbehav*

In table A.10 it is observable that the financial literacy measure is able to explain this gap between preferences and actual behavior (contrarily to the use of *finknow*), displaying a negative relationship. Thus, an additional level in financial literacy (*finbehav*) increases the likelihood of lower positive differences by 4.50% (of displaying very low) and 1.98% (of displaying low). It also decreases the probability of presenting much more financial products in the hypothetical portfolio than in the current one (on average decreases the willingness to present a high gap by 1.80% and a very high gap by 4.37%). Experience only presents the lowest level of gap statistically significant in which an additional level in experience increases the likelihood of displaying a very low positive gap, however this variable is not statistically significant for the rest of the dependent categories. Similarly, to *finknow*, income, loss aversion, overconfidence and risk profile are not relevant at explaining the positive gap, and experience and investment importance also display a negative association with the dependent variable. The magnitudes of the marginal effects of the statistically significant variables do not change much.

Independent variable: perceived

Table A.11 demonstrates that this measure provides similar results to the default financial literacy measure. Thus, similarly financial literacy is not statistically significant, neither is risk profile, income, loss aversion and overconfidence. Experience and investment importance also display negative relationship, as observed with the use of other financial literacy measures. The magnitudes are very similar to *finknow* and *finbehav*.

5.6.5 Disposition effect

5.6.5.1 Disposition effect with reference point

Independent variable: *finbehav*

Table A.12 starts by pointing out that financial literacy is not relevant at explaining the individual's propensity to engage in this bias. Other variables that are also not statistically significant at explaining it are experience, loss, gender, and age (as observable in *finknow*). Overconfidence presents a negative relationship, also present in the model with the default measure of financial literacy. Even though numeracy is not part of the hypothesis, it also presents a negative association with this independent

variable. Similarly to *finknow*, this hypothesis is fully rejected since none of the variables is statistically significant, except for overconfidence that presents a negative relationship (the opposite sign of what is hypothesized).

Independent variable: *perceived*

In table A.12, regarding financial literacy, neither of the combinations with actual and perceived financial knowledge are statistically significant (similarly to *finknow* and *finbehav*). Loss, experience, age and gender are also not relevant at explaining disposition effect. As observable with other measures of financial literacy, numeracy presents a negative relationship, therefore those that correctly respond the numeracy question are less likely to present this bias. The conclusions are exactly the same with the other alternative measures of financial literacy. This measure presents higher magnitudes of effect when compared with the other two measures.

5.6.5.2 Disposition effect without reference point

Independent variable: *finknow*, *finbehav* and *perceived*

In table A.13, the only difference from the three models with different measures of financial literacy occurs in the occupation variable in which being retired (compared to being student) presents a negative relationship in the models with *finknow* and *perceived*, whereas in the model with *finbehav* this category is not statistically significant.

In sum, some variables shift from being statistically significant depending on the measure adopted, thus there are slight changes regarding the rejection of the hypothesis, however these conclusions do not vary greatly. The determinants of financial literacy vary slightly depending on the measure adopted (as it is the case risk profile, loss aversion, age, income and living in a metropolitan area). Regarding market participation, the negative association with age and loss aversion is rejected in any financial literacy definition, and the positive relationship with overconfidence is rejected in the perceived measure. For the third model that captures the portfolio composition, the differences revolve around income, risk profile and overconfidence. When assessing the positive gap between the hypothetical portfolio and actual portfolio, the effect of financial literacy and experience varies depending on the measure chosen. In terms of disposition effect, either measuring with reference point or not, there are no substantial changes regardless of the financial literacy measure adopted. Moreover most effects that are statistically significant decrease their effect and their likelihood of predicting the studied dependent variables when using *finbehav* comparatively to *finknow*. The use of *perceived* does not

display such substantial differences in terms of magnitudes of effects. Therefore, the use of the standardized measure (*finknow*) amplifies considerably and maybe overestimates the effects. Thus, as hypothesized as the definition includes different dimensions of financial literacy, the results increase. This is observable especially from *finknow* to *finbehav* in which in this latter definition it is included a behavioral component. However, the measure of *perceived* appears as broader when compared to *finknow*, (it includes an additional dimension of financial knowledge), the difference in magnitudes and statistical significance between these two variables does not appear as substantial. It may be explained by the fact that *finknow* and *perceived* only assess financial knowledge even if the latter measure also includes self-assessment.

Table 24: Summary results regarding alternative measures of financial literacy

	finknow	finbehav	perceived
H1.1	exper (+), riskprofile (n.s.s), prct_sec (+)	exper (+), riskprofile (+), prct_sec (+)	exper (+), riskprofile (+), prct_sec (+)
H1.2	numeracy (n.s.s), crt (+), loss (-), overc (-)	numeracy (n.s.s), crt (+), loss (n.s.s), overc (-)	numeracy (n.s.s), crt (+), loss (n.s.s), overc (+)
H1.3	age (n.s.s), gender(+), income (+), studyfield (+), metro (+)	age (+), gender(+), income (+), studyfield (+), metro (n.s.s)	age (n.s.s), gender(+), income (n.s.s), studyfield (+), metro (n.s.s)
H2	finknow (+), riskprofile (+), income (+), age (+), loss (n.s.s), overc(+)	finbehav (+), riskprofile (+), income (+), age (n.s.s), loss (n.s.s), overc(+)	perceived (+), riskprofile (+), income (+), age (+), loss (n.s.s), overc (n.s.s)
H3	finknow (+), prct_sec (+), income (+), riskprofile (+), loss (n.n.s) and overc (n.s.s)	finbehav (+), prct_sec (+), income (n.s.s), riskprofile (n.s.s), loss (n.n.s) and overc (+)	perceived (+), prct_sec (+),income (+), riskprofile (+), loss (n.s.s) and overc (-)
H4	Finknow (n.s.s), exper (-), prct_sec (-),risk (n.s.s), income (n.s.s), and. Loss (n.s.s) and overc (n.s.s)	Finbehav (-), exper (n.s.s), prct_sec (-),risk (n.s.s), income (n.s.s), and. Loss (n.s.s) and overc (n.s.s)	perceived (n.s.s), exper (-), prct_sec (-),risk (n.s.s), income (n.s.s), and. Loss (n.s.s) and overc (n.s.s)
H5 - de	finknowledge (n.s.s), exper (n.s.s), age (n.s.s), gender (n.s.s), loss (n.s.s) and overc(-)	finbehav (n.s.s), exper (n.s.s), age (n.s.s), gender (n.s.s), loss (n.s.s) and overc(-)	perceived (n.s.s), exper (n.s.s), age (n.s.s), gender (n.s.s), loss (n.s.s) and overc(-)
H5 - de1	finknowledge (n.s.s), exper (n.s.s), age (n.s.s), gender (n.s.s), loss (n.s.s) and overc(n.n.s)	finknowledge (n.s.s), exper (n.s.s), age (n.s.s), gender (n.s.s), loss (n.s.s) and overc(n.s.s)	finknowledge (n.s.s), exper (n.s.s), age (n.s.s), gender (n.s.s), loss (n.s.s) and overc(n.s.s)

Notes: Table 24 displays an overview of how the hypothesis hold and change depending on the measure adopted. "n.s.s" stands for non-statistically significance.

6. Conclusion

The lack of financial literacy is a worldwide disseminated issue, with an on-going debate on how governments can boost this vital concept. However, and despite many financial education programs and measures taken, it is still a hard and unsolved topic. Additionally, it is also observable a certain stiffness when it comes to market participation, specifically in the stock market, as it has been widely documented in literature. Not only individuals are reluctant in participating in the capital market, once they do it, they are still propense at displaying sub-optimal behaviors, as it is the case of the disposition effect. Therefore, this dissertation intends to, at first instance, explain financial literacy. Afterward it aims a better understandment of individuals' willingness to participate in capital markets and once they do it, their propensity to engage in disposition effect. In order to comprehend market participation and investor's disposition effect it is studied the impact of financial literacy and behavioral biases (as it is the case of loss aversion and overconfidence) in these financial behaviors. Throughout this research it is also discussed sub-themes of these three main topics. Regarding the first topic that concerns financial literacy, it is also examined this concept through alternative measures (tested through H6). In terms of market participation, it is also assessed the portfolio composition or diversification of the individuals (H3), as well as the positive gap between a hypothetical portfolio and their actual portfolio (H4).

As hypothesized, financial literacy is positively explained by experience and investment importance, thus an additional level in experience and investment importance, makes the individual more likely of achieving higher levels of financial literacy. Contrarily to what is hypothesized, the individual's risk profile is not able to explain its association with financial literacy. In terms of cognitive and behavioral aspects, cognitive reflection is also positively associated with financial literacy, whereas loss aversion and overconfidence present a negative relationship, therefore an additional level in these variables decreases the likelihood of the individual being financially literate (the positive relationship with loss aversion and overconfidence is rejected). In the case of numeracy, this variable is not statistically significant because *studyfield* incorporates its effect, nonetheless mathematical capacities are important and present a positive impact in financial literacy levels. As said, displaying mathematic/engineering background as well as economic/management is positively associate with financial literacy levels. Being male, wealthier and living in metropolitan areas is associated with higher levels of financial literacy. Whereas, and contrarily to the vast literature, age is not statistically significant. This upper advantage in financial literacy for male individuals, in part results from differences of overconfidence levels.

In terms of market participation, an additional level in individual's risk profile increases the likelihood of participating the market. This positive relationship is also observable with financial literacy,

income and age. Loss aversion is not statistically significant at explaining market participation, whereas overconfidence displays a positive association. Thus, not only the actual financial knowledge matters, but also the perception that the individual has of its own knowledge. As the difference between self-assessed and actual knowledge increases (overconfidence), the individual increases its willingness to participate. Therefore, individuals that are less risk averse, present higher financial literacy, income, age and overconfidence are the most likely individuals to participate in the market. Related with market participation, two subsets of topics are also studied: portfolio composition (H3) and also the gap between hypothetical and actual portfolio (H4). When explaining the portfolio composition (some authors use this composition as a proxy for diversification), additional levels in financial literacy, risk profile, investment importance and experience increase the likelihood of displaying higher levels of portfolio composition (and therefore more financial products/higher portfolio diversification). However, the hypothesized negative association with loss aversion and overconfidence is not observable due to lack of statistical significance. Lastly, and in regard to the positive gap between preferences and behavior, experience, investment importance and income are able to explain this dependent variable. More precisely, these variables display a negative association (as hypothesized), thus individuals with higher experience, those that are wealthier and invest more aggressively are less likely to present a hypothetical portfolio much more diversified than their current one. Financial literacy, risk profile, loss aversion and overconfidence, are not capable at explaining this phenomenon.

Regarding disposition effect among investors, the only hypothesized variable that appears to explain disposition effect is overconfidence in which, an additional level in the overconfidence scale, makes the investor less likely to engage in this bias. Therefore, overconfidence acts as a deterrent. However, none of the rest hypothesized variables (financial literacy, experience, age, loss aversion and gender) are statistically significant. Disposition effect determinants remain fairly unknown in this research, which may indicate that this bias is widespread and does not provide a cohesive understanding, however new variables must be added in order to provide a better understanding. Contrarily to Da Costa et al. (2008) results that disposition effect decreases if the reference point is removed, that is not observable in this data, but it may also derive from the proximity of questions either temporal or in terms of content. The main results remain unaltered, except for overconfidence that is no longer able to explain disposition effect with an alternative measure.

Lastly, and since financial literacy does not provide a universal definition or measure, it is assessed how the results change based on which measure is adopted. When comparing the three measures of financial literacy, there are some alterations in terms of variables being statistically

significant, and as a result there are changes in terms of hypothesis rejection. Overall, the use of the default measure (*finknow*) displays the highest magnitudes of effects, followed by the *perceived* measure. Thus, the standard definition, composed only by financial knowledge may overestimate financial literacy when compared to a broader measure, *finbehav*, that is also composed by financial behavior. Indeed, with the addition of another dimension to the financial literacy definition the difference of results rises, as hypothesized. However, even though *perceived* is more complex when compared to *finknow*, the lack of difference in terms of results may derive from the fact that, ultimately, it is only the financial knowledge component that is being assessed in both these measures.

6.1 Limitations and avenues for further research

One of the first limitations that this research faces is related with the construction of the questionnaire. Since this secondary data did not impose mandatory requirements in terms of answers, individuals that would have already given up answering (normally associated with individuals with lower levels of knowledge), persisted doing it. This creates a further challenge of not being able to distinguish missing data as wrong answers and the individual not knowing the correct solution or whether these individuals consciously do not feel at answering it. Additionally, there are a lot of inconsistencies in individual's questionnaire interpretation, which required data manipulation in order to provide a more cohesive data. There is also some difficulty in explaining concepts as disposition effect or even the gap between preferences and actual behavior, which indicates the need for additional variables in order to better comprehend these sub-optimal behaviors. Lastly, this research results must be handled carefully regarding its generalization to the Portuguese population. This sample is not representative of the Portuguese population as it is easily observable in the descriptive statistics. Moreover, its mean of distribution biases the kind of individuals that complete the questionnaire – online distributed questionnaires target a specific type of inquiries.

In terms of further research, it would be advised to replicate this research to a representative sample in order for a better overview of the Portuguese population. Even though this research features other financial literacy measures other than financial knowledge, the addition of the “financial attitudes” component in the broader financial literacy measure would be appealing, therefore complying with a broader vision of financial literacy as supported by OECD. Additionally, the inclusion of other explanatory variables as frequency in trading or other behavioral biases may also enrichen the results. Regarding disposition effect, new variables are needed when explaining this concept, since the more traditional variables were not able to explain this financial behavior. Even though this research does not elaborate

further in the importance of the reference point, it would be curious understanding how the inclusion of reference point changes the dynamic of disposition effect.

References

- Abreu, M., & Mendes, V. (2010). Financial literacy and portfolio diversification. *Quantitative Finance*, 10(5), 515–528. <https://doi.org/10.1080/14697680902878105>
- Agarwal, S., & Mazumder, B. (2013). Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics*, 5(1), 193-207. <https://doi.org/10.1257/app.5.1.193>.
- Allgood, S., & Walstad, W. B. (2016). The effects of perceived and actual financial literacy on financial behaviors. *Economic Inquiry*, 54(1), 675–697. <https://doi.org/10.1111/ecin.12255>
- Banco de Portugal (2011). *Inquérito à Literacia Financeira da população portuguesa. Edição: Departamento de Supervisão Bancária*. <https://www.bportugal.pt/sites/default/files/anexos/documentos-relacionados/relatorioinqueritoliteraciafinanceira.pdf>
- Bannier, C. E., & Neubert, M. (2016). Gender differences in financial risk taking: The role of financial literacy and risk tolerance. *Economics Letters*, 145, 130-135. <https://doi.org/10.1016/j.econlet.2016.05.033>
- Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The quarterly journal of economics*, 116(1), 261-292. <https://doi.org/10.1162/003355301556400>
- Baker, H. K., Kumar, S., Goyal, N., & Gaur, V. (2019). How financial literacy and demographic variables relate to behavioral biases. *Managerial Finance*, 45(1), 124–146. <https://doi.org/10.1108/mf-01-2018-0003>
- Bryman, A., & Bell, E. (2007). *Research Designs. In Business research methods*. Oxford University Press.
- Bucher-Koenen, T., Alessie, R., Lusardi, A., & Van Rooij, M. (2021). Fearless Woman: Financial Literacy and Stock Market Participation. *SSRN Electronic Journal*, March. <https://doi.org/10.2139/ssrn.3805715>
- Burton, E., & Shah, S. (2013). *Behavioral finance: understanding the social, cognitive, and economic debates* (Vol. 854). John Wiley & Sons.
- Calcagno, R., & Monticone, C. (2015). Financial literacy and the demand for financial advice. *Journal of Banking & Finance*, 50, 363–380. <https://doi.org/10.1016/j.jbankfin.2014.03.013>
- Calvet, L. E., Campbell, J. Y., & Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5), 707-747. <https://doi.org/10.1086/524204>
- Christelis, D., Jappelli, T., & Padula, M. (2010). Cognitive abilities and portfolio choice. *European Economic Review*, 54(1), 18-38. <https://doi.org/10.1016/j.euroecorev.2009.04.001>

Coutinho, C. P. (2013). *Metodologia de Investigação em Ciências Sociais e Humanas*. Almedina.

Da Costa Jr, N., Mineto, C., & Da Silva, S. (2008). Disposition effect and gender. *Applied Economics Letters*, 15(6), 411-416. <https://doi.org/10.1080/13504850600706560>

Daniels, L., & Minot, N. (2019). *An Introduction to Statistics and Data Analysis Using Stata®: From Research Design to Final Report*. Sage Publications.

Dhar, R., & Zhu, N. (2006). Up close and personal: Investor sophistication and the disposition effect. *Management Science*, 52(5), 726–740. <https://doi.org/10.1287/mnsc.1040.0473>

Disney, R., & Gathergood, J. (2013). Financial literacy and consumer credit portfolios. *Journal of Banking & Finance*, 37(7), 2246–2254. <https://doi.org/10.1016/j.jbankfin.2013.01.013>

Feng, L., & Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance*, 9(3), 305–351. <https://doi.org/10.1007/s10679-005-2262-0>

Fernandes, D., Lynch, J. G., & Netemeyer, R. G. (2014). Financial Literacy, Financial Education, and Downstream Financial Behaviors. *Management Science*, 60(8), 1861–1883. <https://doi.org/10.1287/mnsc.2013.1849>

Fogel, S. O. C., & Berry, T. (2010). The disposition effect and individual investor decisions: the roles of regret and counterfactual alternatives. In *Handbook of behavioral finance*. Edward Elgar Publishing. https://doi.org/10.1207/s15427579jpfm0702_5

Gabinete de Estudos da CMVM. (2019). Resultados do Inquérito Online ao Investidor 2018. *CMVM*. Retrieved from https://www.cmvm.pt/pt/EstatisticasEstudosEPublicacoes/Estudos/Documents/ResultadosInqueritoOnlinePerfilInvestidor_2019.pdf

Gaudecker, H. M. V. (2015). How does household portfolio diversification vary with financial literacy and financial advice?. *The Journal of Finance*, 70(2), 489-507. <https://doi.org/10.1111/jofi.12231>

Georgarakos, D., & Inderst, R. (2011). Financial Advice and Stock Market Participation. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1641302>

Gerth, F., Lopez, K., Reddy, K., Ramiah, V., Wallace, D., Muschert, G., Frino, A., & Jooste, L. (2021). The Behavioural Aspects of Financial Literacy. *Journal of Risk and Financial Management*, 14(9), 395. <https://doi.org/10.3390/jrfm14090395>

Gujarati, D. N., & Porter, D. (2009). *Basic Econometrics* Mc Graw-Hill International Edition.

Hayat, A. and Anwar, M. (2016), Impact of behavioral biases on investment decision; moderating role of financial literacy. *SSRN Electronic Journal*, pp. 1-14. <https://doi.org/10.2139/ssrn.2842502>

Hilgert, M. A., Hogarth, J. M., & Beverly, S. G. (2003). Household financial management: The connection between knowledge and behavior. *Federal Reserve Bulletin*, 309-322

Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance*, 56(4), 1533–1597. <https://doi.org/10.1111/0022-1082.00379>

Hsu, J. W. (2011). Aging and Strategic Learning: The Impact of Spousal Incentives on Financial Literacy. *SSRN Electronic Journal*. Published. <https://doi.org/10.2139/ssrn.1969411>

Huston, S. J. (2010). Measuring Financial Literacy. *Journal of Consumer Affairs*, 44(2), 296–316. <https://doi.org/10.1111/j.1745-6606.2010.01170.x>

Jappelli, T., & Padula, M. (2013). Investment in financial literacy and saving decisions. *Journal of Banking & Finance*, 37(8), 2779–2792. <https://doi.org/10.1016/j.jbankfin.2013.03.019>

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decisions under Risk," *Econometrica*, 47(2), 263- 291

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127). https://doi.org/10.1142/9789814417358_0006

Klapper, L., Lusardi, A., & Panos, G. A. (2013). Financial literacy and its consequences: Evidence from Russia during the financial crisis. *Journal of Banking & Finance*, 37(10), 3904– 3923. <https://doi.org/10.1016/j.jbankfin.2013.07.014>

Korniotis, G. M., & Kumar, A. (2011). Do older investors make better investment decisions?. *The Review of Economics and Statistics*, 93(1), 244-265. https://doi.org/10.1162/REST_a_00053

Leal, C.C. and Oliveira, B. (2021). Choice Architecture: Nudging for Sustainable Behavior. In *Sustainable Management for Managers and Engineers* (eds C.F. Machado and J.P. Davim). <https://doi.org/10.1002/9781119804345.ch1>

Leal, C.C., Armada, M.R. (2005). *Fundamentos Psicológicos e Comportamentais da Actuação do Investidor* [Unpublished Manuscript]. Department of Management, University of Minho

Lusardi, A. (2012). *Numeracy, financial literacy, and financial decision-making* (No. w17821). National Bureau of Economic Research.

Lusardi, A., & Mitchell, O. S. (2011). Financial Literacy Around the World: An Overview. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1810551>

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., & Oggero, N. (2017). Millennials and financial literacy: A global perspective. *Global Financial Literacy Excellence Center*.

Martins, P., Gonçalves, T., Correia, L., Gouveia, S., & Lourenço-Gomes, L. (2021). Diferentes medidas de literacia financeira: o caso dos estudantes universitários portugueses. *Cadernos Do Mercado de Valores Mobiliários*, 69, 9-25.

<https://www.cmvm.pt/pt/EstatisticasEstudosEPublicacoes/CadernosDoMercadoDeValoresMobiliarios/Documents/CMVM-CADERNOS-N69.pdf>

Moore, D. L. (2003). *Survey of financial literacy in Washington State: Knowledge, behavior, attitudes, and experiences*. Washington State Department of Financial Institutions. <https://doi.org/10.13140/2.1.4729.4722>

Muñoz-Murillo, M., Álvarez-Franco, P. B., & Restrepo-Tobón, D. A. (2020). The role of cognitive abilities on financial literacy: New experimental evidence. *Journal of Behavioral and Experimental Economics*, *84*, 101482. <https://doi.org/10.1016/j.socec.2019.101482>

Nilesh, K., Cruz, M., Leitão, M., Jivan, R., & Abreu, M. (2021). As mulheres no mercado financeiro português. *Cadernos Do Mercado de Valores Mobiliários*, *69*, 40-74. <https://www.cmvm.pt/pt/EstatisticasEstudosEPublicacoes/CadernosDoMercadoDeValoresMobiliarios/Documents/CMVM-CADERNOS-N69.pdf>

Odean, T. (2019). Are investors reluctant to realize their losses? *Choices, Values, and Frames*, *LIII(5)*, 371–392. <https://doi.org/10.1017/CBO9780511803475.022>

OECD (2016), “OECD/INFE International Survey of Adult Financial Literacy Competencies”, OECD, Paris, www.oecd.org/finance/OECD-INFE-International-Survey-of-Adult-Financial-Literacy-Competencies.pdf

OECD. (2019). *PISA 2021 financial literacy analytical and assessment framework*. OECD Publishing. <https://www.oecd.org/pisa/sitedocument/PISA-2021-Financial-Literacy-Framework.pdf>

OECD (2020), *PISA 2018 Results (Volume IV): Are Students Smart about Money?*, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/48ebd1ba-en>.

Özen, E., & Ersoy, G. (2019). The impact of financial literacy on cognitive biases of individual investors. *Contemporary Studies in Economic and Financial Analysis*, *101*, 77–95. <https://doi.org/10.1108/S1569-375920190000101007>

Plano Nacional de Formação Financeira. (2020). Portugal em 7o lugar no inquérito internacional à literacia financeira dos adultos 2020 da OCDE/INFE. Todos Contam. <https://www.todoscontam.pt/pt-pt/noticias/portugal-em-7o-lugar-no-inquerito-internacional-literacia-financeira-dos-adultos-2020-da>

Plano Nacional de Formação Financeira. (2021). *Relatório do 3.º Inquérito à Literacia Financeira da população portuguesa*. <https://www.todoscontam.pt/sites/default/files/2021-06/relatorio3inqlf.pdf>

Pompian, M. M. (2006). *Behavioral Finance and Wealth Management: How to Build Optimal Portfolios That Account for Investor Biases (Wiley Finance)* (1st ed.). Wiley.

Potrich, A. C. G., Vieira, K. M., & Kirch, G. (2015). Determinants of financial literacy: Analysis of the influence of socioeconomic and demographic variables. *Revista Contabilidade e Finanças*, *26(69)*, 362–377. <https://doi.org/10.1590/1808-057x201501040>

Rau, H. A. (2014). The disposition effect and loss aversion: Do gender differences matter?. *Economics Letters*, 123(1), 33-36. <https://doi.org/10.1016/j.econlet.2014.01.020>

Remund, D. L. (2010). Financial literacy explicated: The case for a clearer definition in an increasingly complex economy. *Journal of Consumer Affairs*, 44(2), 276–295. <https://doi.org/10.1111/j.1745-6606.2010.01169.x>

Rieger, M. O. (2020). Uncertainty avoidance, loss aversion and stock market participation. *Global Finance Journal*, 100598. <https://doi.org/10.1016/j.gfj.2020.100598>

Rosa, V. (2021). O PISA e a literacia financeira: os resultados de Portugal. *Revista Portuguesa de Investigação Educacional*, (21), 1-20.

Schmeiser, M. D., & Seligman, J. S. (2013). Using the Right Yardstick: Assessing Financial Literacy Measures by Way of Financial Well-Being. *Journal of Consumer Affairs*, 47(2), 243–262. <https://doi.org/10.1111/joca.12010>

Sebastião, H., Silva, N., Torres, P., & Godinho, P. (2021). Determinantes do enviesamento de perceção da literacia financeira. *Cadernos Do Mercado de Valores Mobiliários*, 69, 26-39. <https://www.cmvm.pt/pt/EstatisticasEstudosEPublicacoes/CadernosDoMercadoDeValoresMobiliarios/Documents/CMVM-CADERNOS-N69.pdf>

Seru, A., Shumway, T., & Stoffman, N. (2010). Learning by trading. *The Review of Financial Studies*, 23(2), 705-739. <https://doi.org/10.1093/rfs/hhp060>

Shefrin, H. (2002). *Beyond Greed and Fear: Understanding Behavioral Finance and the Psychology of Investing* (2nd ed.). Oxford University Press.

Stolper, O. (2018). It takes two to Tango: Households' response to financial advice and the role of financial literacy. *Journal of Banking & Finance*, 92, 295–310. <https://doi.org/10.1016/j.jbankfin.2017.04.014>

Takeda, K., Takemura, T., & Kozu, T. (2013). Investment Literacy and Individual Investor Biases: Survey Evidence in the Japanese Stock Market. *The Review of Socionetwork Strategies*, 7(1), 31–42. <https://doi.org/10.1007/s12626-012-0031-z>

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness* (1st ed.). Yale University Press. <https://doi.org/10.1007/s10602-008-9056-2>

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>

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, 20(2), 147-168. <https://doi.org/10.1080/13546783.2013.844729>

Whitehead, M., Jones, R., Howell, R., Lilley, R., & Pykett, J. (2014). Assessing the global impact of the behavioural sciences on public policy: Nudging all over the world. *Economic and Social Research Council*. <https://changingbehaviours.files.wordpress.com/2014/09/nudgedesignfinal.pdf>

Xia, T., Wang, Z., & Li, K. (2014). Financial literacy overconfidence and stock market participation. *Social indicators research*, 119(3), 1233-1245. <https://doi.org/10.1007/s11205-013-0555-9>

Yamori, N., & Ueyama, H. (2021). Financial Literacy and Low Stock Market Participation of Japanese Households. *Finance Research Letters*, March, 102074. <https://doi.org/10.1016/j.frl.2021.102074>

Yang, L. (2019). Loss aversion in financial markets. *Journal of Mechanism and Institution Design*, 4(1), 119-137.

Yao, R., Sharpe, D. L., & Wang, F. (2011). Decomposing the age effect on risk tolerance. *The Journal of Socio-Economics*, 40(6), 879-887. <https://doi.org/10.1016/j.socec.2011.08.023>

Appendix

Table A.1: Variance Inflation Factor

	VIF		VIF
exper	1.429	exper	1.409
studyfield		studyfield	
Mathematics, Engineering	1.339	Mathematics, Engineering	1.331
Economics, Management	1.473	Economics, Management	1.448
prct sec	1.264	prct sec	1.259
riskprofile	1.386	riskprofile	1.379
overc	1.096	overc	1.093
loss	1.202	loss	1.197
numeracy	1.145	numeracy	1.123
crt	1.223	crt	1.201
age	3.065	age	2.849
gender	1.266	gender	1.254
income	1.956	income	1.862
metro	1.079	metro	1.05
education		occupation	
Middle school education	2.057	Self-employee	1.338
High School education	58.277	Employee	2.119
Attending college	223.862	Unemployed	1.115
College degree	162.297	Retired	1.322
Pos-graduation, Masters, MBA or PhD	225.72	Other	1.106
occupation			
Self-employee	1.415		
Employee	2.45		
Unemployed	1.146		
Retired	1.347		
Other	1.138		
Mean VIF	30.375	Mean VIF	1.414

Notes: Table A.1 demonstrates the Variance Inflation Factor for two models. The first model includes all variables from a standard hypothesis, including education, whereas the second variable removes a possible source of multicollinearity (education).

Table A.2: Marginal effects of the interaction term between *gender* and *overc* with *finknow* as dependent variable for each outcome

Predictive margins

Expression : Pr(finknow==1), predict(outcome(#1))

	Margin	Delta-method			[95%Conf.	Interval]
		Std.Err.	z	P>z		
overc#gender						
1#female	0.000	0.000	1.240	0.214	-0.000	0.000
1#male	0.000	0.000	1.050	0.294	-0.000	0.000
2#female	0.006	0.002	3.140	0.002	0.002	0.010
2#male	0.001	0.001	2.310	0.021	0.000	0.002
3#female	0.079	0.019	4.210	0.000	0.042	0.116
3#male	0.031	0.011	2.810	0.005	0.009	0.053
4#female	0.376	0.068	5.550	0.000	0.243	0.509
4#male	0.124	0.041	3.020	0.002	0.044	0.205

Expression : Pr(finknow==2), predict(outcome(#2))

	Delta-method					
	Margin	Std.Err.	z	P>z	[95%Conf.	Interval]
overc#gender						
1#female	0.002	0.001	1.930	0.054	-0.000	0.004
1#male	0.001	0.001	1.560	0.120	-0.000	0.003
2#female	0.047	0.007	6.550	0.000	0.033	0.061
2#male	0.018	0.004	4.440	0.000	0.010	0.025
3#female	0.208	0.030	6.950	0.000	0.149	0.267
3#male	0.128	0.025	5.180	0.000	0.080	0.177
4#female	0.320	0.034	9.450	0.000	0.254	0.387
4#male	0.253	0.040	6.340	0.000	0.175	0.331

Expression : Pr(finknow==3), predict(outcome(#3))

	Delta-method					
	Margin	Std.Err.	z	P>z	[95%Conf.	Interval]
overc#gender						
1#female	0.036	0.011	3.370	0.001	0.015	0.057
1#male	0.026	0.010	2.600	0.009	0.006	0.045
2#female	0.218	0.017	13.010	0.000	0.185	0.251
2#male	0.129	0.013	9.720	0.000	0.103	0.155
3#female	0.375	0.024	15.600	0.000	0.328	0.423
3#male	0.337	0.026	13.030	0.000	0.286	0.388
4#female	0.235	0.043	5.530	0.000	0.152	0.319
4#male	0.370	0.027	13.940	0.000	0.318	0.422

Expression : Pr(finknow==4), predict(outcome(#4))

	Delta-method					
	Margin	Std.Err.	z	P>z	[95%Conf.	Interval]
overc#gender						
1#female	0.133	0.023	5.680	0.000	0.087	0.178
1#male	0.109	0.025	4.370	0.000	0.060	0.158
2#female	0.302	0.018	17.040	0.000	0.267	0.337
2#male	0.255	0.015	16.950	0.000	0.226	0.285
3#female	0.224	0.023	9.530	0.000	0.178	0.270
3#male	0.286	0.022	13.210	0.000	0.244	0.328
4#female	0.057	0.020	2.880	0.004	0.018	0.096
4#male	0.180	0.034	5.260	0.000	0.113	0.246

Expression : Pr(finknow==5), predict(outcome(#5))

	Delta-method					
	Margin	Std.Err.	z	P>z	[95%Conf.	Interval]
overc#gender						
1#female	0.829	0.034	24.490	0.000	0.763	0.896
1#male	0.864	0.035	24.710	0.000	0.795	0.932
2#female	0.426	0.023	18.350	0.000	0.381	0.472
2#male	0.597	0.021	28.030	0.000	0.555	0.638
3#female	0.114	0.025	4.610	0.000	0.065	0.162
3#male	0.218	0.034	6.450	0.000	0.151	0.284
4#female	0.011	0.006	1.850	0.064	-0.001	0.023
4#male	0.073	0.026	2.790	0.005	0.022	0.124

Notes: Table A.2 demonstrates the marginal effects of the interaction term of *gender* and *overc*, for each possible outcome in financial knowledge levels.

Table A.3: Marginal effects of ordered probit with *finbehav* as dependent variable

	Very Low	Low	Medium	High	Very High
exper	-0.0255*** (0.0054)	-0.0789*** (0.0161)	-0.0302*** (0.0065)	0.1248*** (0.0235)	0.0098*** (0.0038)
studyfield					
Mathematics, Engineering	-0.0205*** (0.0070)	-0.0592*** (0.0212)	0.0063 (0.0044)	0.0705*** (0.0260)	0.0030 (0.0020)
Economics, Management	-0.0366*** (0.0056)	-0.1219*** (0.0151)	-0.0193*** (0.0065)	0.1668*** (0.0182)	0.0110*** (0.0037)
prct_sec	-0.0177*** (0.0026)	-0.0549*** (0.0063)	-0.0210*** (0.0034)	0.0868*** (0.0091)	0.0068*** (0.0023)
riskprofile	-0.0053*** (0.0020)	-0.0163*** (0.0059)	-0.0062** (0.0025)	0.0257*** (0.0093)	0.0020** (0.0010)
overc	0.0399*** (0.0043)	0.1237*** (0.0078)	0.0474*** (0.0079)	-0.1957*** (0.0123)	-0.0154*** (0.0050)
loss	0.0019 (0.0019)	0.0058 (0.0060)	0.0022 (0.0023)	-0.0091 (0.0095)	-0.0007 (0.0008)
numeracy	-0.0066* (0.0038)	-0.0204* (0.0116)	-0.0078* (0.0047)	0.0323* (0.0185)	0.0025 (0.0017)
crt	-0.0091*** (0.0024)	-0.0282*** (0.0071)	-0.0108*** (0.0034)	0.0447*** (0.0113)	0.0035** (0.0014)
age	-0.0004** (0.0002)	-0.0013** (0.0006)	-0.0005** (0.0002)	0.0021** (0.0010)	0.0002* (0.0001)
gender	-0.0122*** (0.0036)	-0.0377*** (0.0107)	-0.0144*** (0.0045)	0.0595*** (0.0166)	0.0047** (0.0021)
income	-0.0050*** (0.0014)	-0.0156*** (0.0041)	-0.0060*** (0.0017)	0.0247*** (0.0064)	0.0019** (0.0008)
metro	-0.0035 (0.0031)	-0.0108 (0.0097)	-0.0041 (0.0038)	0.0171 (0.0153)	0.0013 (0.0013)
occupation					
self-employee	-0.0223*** (0.0063)	-0.0758*** (0.0234)	-0.0444* (0.0235)	0.1340*** (0.0486)	0.0085* (0.0046)
Employee	-0.0188*** (0.0046)	-0.0618*** (0.0147)	-0.0303*** (0.0098)	0.1048*** (0.0267)	0.0060*** (0.0023)
Unemployed	-0.0159** (0.0074)	-0.0510** (0.0257)	-0.0215 (0.0177)	0.0840* (0.0470)	0.0044 (0.0035)
Retired	0.0178 (0.0206)	0.0458 (0.0478)	-0.0027 (0.0113)	-0.0590 (0.0559)	-0.0019 (0.0018)
Other	-0.0146* (0.0078)	-0.0465* (0.0263)	-0.0184 (0.0168)	0.0757 (0.0472)	0.0039 (0.0036)
Observations	934	934	934	934	934

Notes: Table A.3 demonstrates the marginal effect of ordered probit with *finbehav* as dependent variable. Variable *occupation* uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as

income	0.0070	-0.0001	-0.0019	-0.0050
	(0.0090)	(0.0001)	(0.0024)	(0.0065)
metro	0.0338	-0.0005	-0.0091	-0.0242
	(0.0237)	(0.0004)	(0.0064)	(0.0170)
occupation				
self-employee	0.1057	-0.0031	-0.0330	-0.0696*
	(0.0665)	(0.0031)	(0.0234)	(0.0405)
Employee	0.0212	-0.0003	-0.0057	-0.0153
	(0.0354)	(0.0005)	(0.0096)	(0.0254)
Unemployed	0.1270*	-0.0042	-0.0410	-0.0818*
	(0.0740)	(0.0041)	(0.0276)	(0.0428)
Retired	-0.0966	-0.0010	0.0192	0.0784
	(0.1125)	(0.0032)	(0.0162)	(0.0997)
Other	-0.0197	0.0001	0.0049	0.0148
	(0.0541)	(0.0002)	(0.0129)	(0.0411)
Observations	934	934	934	934

Notes: Table A.5 demonstrates the marginal effect of ordered probit with *perceived* as dependent variable. Variable occupation uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Marginal effects of the interaction term between *gender* and *overc* with *perceived* as dependent variable for each outcome

Predictive margins Number of obs = 934
Model VCE : Robust
Expression : Pr(perceived==1), predict(outcome(#1))

	Delta-method					
	Margin	Std.Err.	z	P>z	[95%Conf.	Interval]
overc#gender						
1#female	0.338	0.028	11.930	0.000	0.282	0.393
1#male	0.337	0.027	12.700	0.000	0.285	0.389
2#female	0.428	0.028	15.460	0.000	0.373	0.482
2#male	0.710	0.026	27.480	0.000	0.660	0.761
3#female	0.377	0.046	8.150	0.000	0.287	0.468
3#male	0.515	0.051	10.040	0.000	0.415	0.616
4#female	0.388	0.065	5.970	0.000	0.260	0.515
4#male	0.549	0.065	8.490	0.000	0.422	0.676

Expression : Pr(perceived==3), predict(outcome(#3))

	Delta-method					
	Margin	Std.Err.	z	P>z	[95%Conf.	Interval]
overc#gender						
1#female	0.283	0.018	15.940	0.000	0.248	0.318
1#male	0.283	0.018	15.950	0.000	0.248	0.318
2#female	0.267	0.018	14.990	0.000	0.232	0.302
2#male	0.159	0.015	10.690	0.000	0.129	0.188
3#female	0.277	0.019	14.450	0.000	0.240	0.315
3#male	0.241	0.021	11.650	0.000	0.201	0.282

4#female	0.276	0.022	12.790	0.000	0.233	0.318
4#male	0.229	0.027	8.380	0.000	0.176	0.283

Notes: Table A.6 displays the marginal effects of the interaction term of *gender* and *overc*, for each possible outcome of these variables for the High/High and Low/High levels of perceived.

Table A.7: Marginal effects of ordered probit with *marketpart* as dependent variable [use of *perceived* and *finbehav* as independent variables]

	marketpart		marketpart
riskprofile	0.0902***	riskprofile	0.0666***
	(0.0158)		(0.0153)
perceived		finbehav	0.3005***
H/H	0.290***		(0.0183)
	(0.0426)	income	0.0234***
L/H	0.144**		(0.0085)
	(0.0686)	age	0.0012
H/L	0.0852*		(0.0015)
	(0.0502)	loss	-0.0117
income	0.0417***		(0.0153)
	(0.00904)	overc	0.1273***
age	0.00366**		(0.0208)
	(0.00161)	studyfield	
loss	-0.00754	Mathematics, Engineering	-0.0540
	(0.0164)		(0.0404)
overc	-0.0212	Economics, Management	-0.1569***
	(0.0281)		(0.0288)
studyfield		numeracy	0.0322
Mathematics, Engineering	-0.0255		(0.0289)
	(0.0448)	crt	0.0159
Economics, Management	-0.117***		(0.0181)
	(0.0328)	gender	0.0315
numeracy	0.0563*		(0.0259)
	(0.0316)	occupation	
crt	0.0342*	self-employee	0.0774
	(0.0200)		(0.0727)
gender	0.0291	Employee	0.0917**
	(0.0285)		(0.0388)
occupation		Unemployed	-0.0740
self-employee	0.170**		(0.0711)
	(0.0835)	Retired	0.3480***
Employee	0.180***		(0.1097)
	(0.0422)	Other	-0.0360
Unemployed	-0.0479		(0.0622)
	(0.0802)	metro	-0.0226

Retired	0.418*** (0.122)	Observations	1,057
Other	0.0184 (0.0727)		
metro	-0.0135 (0.0267)		
Observations	1,057		

Notes: Table A.7 demonstrates the marginal effect of probit with *marketpart* as dependent variable. The base category for perceived is Low/Low. Variable occupation uses “ Student” as reference, and the variable *studyfield* uses “non-mathematic and non-economic background” as reference . In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.8: Marginal effects of ordered probit with *portf* as dependent variables [use of *finbehav* as independent variable]

	very low	low	medium	high
finbehav	-0.255*** (0.0133)	0.172*** (0.0160)	0.0752*** (0.0110)	0.00795** (0.00360)
riskprofile	-0.00880 (0.00840)	0.00593 (0.00568)	0.00260 (0.00251)	0.000275 (0.000284)
prct_sec	-0.0458*** (0.00662)	0.0308*** (0.00522)	0.0135*** (0.00254)	0.00143** (0.000678)
exper	-0.0206 (0.0184)	0.0139 (0.0126)	0.00607 (0.00536)	0.000642 (0.000640)
loss	-0.00225 (0.00664)	0.00152 (0.00447)	0.000665 (0.00196)	7.03e-05 (0.000207)
overc	-0.0379*** (0.0128)	0.0255*** (0.00917)	0.0112*** (0.00373)	0.00118* (0.000662)
income	-0.00832* (0.00446)	0.00561* (0.00307)	0.00246* (0.00133)	0.000260 (0.000177)
studyfield				
Mathematics, Engineering	-0.0333 (0.0278)	0.0211 (0.0178)	0.0110 (0.00912)	0.00123 (0.00110)
Economics, Management	0.0442*** (0.0167)	-0.0295** (0.0115)	-0.0132** (0.00530)	-0.00141* (0.000763)
crt	0.0359*** (0.0114)	-0.0242*** (0.00802)	-0.0106*** (0.00358)	-0.00112* (0.000592)
numeracy	0.0347** (0.0164)	-0.0234** (0.0112)	-0.0103** (0.00497)	-0.00108 (0.000726)
gender	0.0150 (0.0136)	-0.0101 (0.00927)	-0.00444 (0.00397)	-0.000469 (0.000470)
age	-0.000405 (0.000737)	0.000273 (0.000495)	0.000119 (0.000220)	1.26e-05 (2.33e-05)
metro	-0.0181 (0.0125)	0.0122 (0.00845)	0.00536 (0.00380)	0.000566 (0.000470)
occupation				

self-employee	0.0298 (0.0298)	-0.0217 (0.0222)	-0.00731 (0.00718)	-0.000757 (0.000788)
Employee	-0.0176 (0.0193)	0.0122 (0.0137)	0.00487 (0.00509)	0.000541 (0.000624)
Unemployed	-0.0137 (0.0328)	0.00953 (0.0227)	0.00377 (0.00914)	0.000417 (0.00101)
Retired	-0.108 (0.0695)	0.0708 (0.0451)	0.0335 (0.0224)	0.00396 (0.00305)
Other	-0.0326 (0.0354)	0.0223 (0.0240)	0.00927 (0.0104)	0.00104 (0.00128)
Observations	934	934	934	934

Notes: Table A.8 demonstrates the marginal effect of ordered probit with *portf* as dependent variable. The base category for perceived is Low/Low. Variable occupation uses “ Student” as reference, and the variable *studyfield* uses “non-mathematic and non-economic background” as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.9: Marginal effects of ordered probit with *portf* as dependent variables [use of *perceived* as independent variable]

	very low	low	medium	high
perceived				
H/H	-0.104*** (0.0386)	0.0809** (0.0316)	0.0204*** (0.00710)	0.00269 (0.00165)
L/H	-0.147** (0.0736)	0.112** (0.0527)	0.0309 (0.0196)	0.00451 (0.00352)
H/L	0.0215 (0.0435)	-0.0178 (0.0360)	-0.00332 (0.00690)	-0.000328 (0.000677)
riskprofile				
	-0.0439*** (0.0107)	0.0324*** (0.00807)	0.00990*** (0.00283)	0.00160* (0.000887)
prct_sec				
	-0.112*** (0.00818)	0.0829*** (0.00666)	0.0254*** (0.00400)	0.00410* (0.00221)
exper				
	-0.109*** (0.0267)	0.0804*** (0.0200)	0.0246*** (0.00714)	0.00398* (0.00220)
loss				
	0.00611 (0.0107)	-0.00451 (0.00789)	-0.00138 (0.00239)	-0.000223 (0.000428)
overc				
	0.0851*** (0.0286)	-0.0628*** (0.0213)	-0.0192** (0.00759)	-0.00311** (0.00149)
income				
	-0.0335*** (0.00716)	0.0247*** (0.00534)	0.00757*** (0.00189)	0.00122* (0.000724)
studyfield				
Mathematics, Engineering	-0.0477 (0.0342)	0.0348 (0.0249)	0.0111 (0.00820)	0.00181 (0.00169)
Economics, Management	-0.0106 (0.0239)	0.00793 (0.0179)	0.00234 (0.00520)	0.000352 (0.000825)
crt				
	0.0135 (0.0164)	-0.00998 (0.0121)	-0.00305 (0.00373)	-0.000494 (0.000700)
numeracy				
	0.0275	-0.0203	-0.00622	-0.00101

	(0.0265)	(0.0195)	(0.00596)	(0.00120)
gender	-0.0239	0.0176	0.00539	0.000873
	(0.0212)	(0.0157)	(0.00480)	(0.000931)
age	-8.36e-05	6.17e-05	1.89e-05	3.05e-06
	(0.00116)	(0.000853)	(0.000261)	(4.19e-05)
metro	-0.0107	0.00786	0.00240	0.000389
	(0.0196)	(0.0145)	(0.00444)	(0.000712)
occupation				
self-employee	-0.0344	0.0274	0.00628	0.000785
	(0.0524)	(0.0413)	(0.00971)	(0.00147)
Employee	-0.0900***	0.0695***	0.0179***	0.00254
	(0.0310)	(0.0250)	(0.00596)	(0.00161)
Unemployed	-0.135**	0.102**	0.0286*	0.00448
	(0.0582)	(0.0421)	(0.0146)	(0.00354)
Retired	-0.145*	0.109*	0.0312	0.00500
	(0.0846)	(0.0610)	(0.0207)	(0.00458)
Other	-0.0824	0.0639	0.0162	0.00226
	(0.0583)	(0.0441)	(0.0127)	(0.00231)
Observations	934	934	934	934

Notes: Table A.9 demonstrates the marginal effect of ordered probit with *portf* as dependent variable. The base category for perceived is Low/Low. Variable occupation uses “ Student” as reference, and the variable *studyfield* uses “non-mathematic and non-economic background” as reference . In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.10: Marginal effects of ordered probit with *pgap* as dependent variables [use of *finbehav* as independent variable]

	very low	low	medium	high	very high
finbehav	0.0533**	0.0235**	-0.00351	-0.0214**	-0.0518**
	(0.0209)	(0.00943)	(0.00303)	(0.00867)	(0.0205)
riskprofile	-0.0134	-0.00588	0.000880	0.00537	0.0130
	(0.0119)	(0.00525)	(0.00107)	(0.00478)	(0.0115)
prct_sec	0.0506***	0.0223***	-0.00334	-0.0204***	-0.0492***
	(0.00968)	(0.00463)	(0.00259)	(0.00412)	(0.0101)
exper	0.0601**	0.0264**	-0.00396	-0.0242**	-0.0584**
	(0.0285)	(0.0129)	(0.00351)	(0.0116)	(0.0282)
loss	0.0173	0.00760	-0.00114	-0.00695	-0.0168
	(0.0117)	(0.00515)	(0.00120)	(0.00479)	(0.0113)
income	0.00992	0.00437	-0.000654	-0.00399	-0.00964
	(0.00703)	(0.00313)	(0.000665)	(0.00285)	(0.00689)
overc	0.00388	0.00171	-0.000256	-0.00156	-0.00378
	(0.0154)	(0.00678)	(0.00105)	(0.00619)	(0.0149)
studyfield					
Mathematics, Engineering	-0.0152	-0.00636	0.00161	0.00610	0.0139
	(0.0328)	(0.0139)	(0.00335)	(0.0131)	(0.0304)
Economics, Management	-0.0252	-0.0109	0.00210	0.0102	0.0239
	(0.0244)	(0.0104)	(0.00287)	(0.00989)	(0.0224)

crt	0.0164 (0.0147)	0.00723 (0.00649)	-0.00108 (0.00129)	-0.00661 (0.00592)	-0.0160 (0.0143)
numeracy	0.0107 (0.0245)	0.00470 (0.0108)	-0.000703 (0.00170)	-0.00429 (0.00986)	-0.0104 (0.0239)
gender	0.0388* (0.0207)	0.0171* (0.00904)	-0.00256 (0.00248)	-0.0156* (0.00832)	-0.0377* (0.0200)
age	0.00304** (0.00132)	0.00134** (0.000585)	-0.000200 (0.000178)	-0.00122** (0.000546)	-0.00295** (0.00127)
metro	-0.0175 (0.0195)	-0.00772 (0.00854)	0.00116 (0.00160)	0.00705 (0.00785)	0.0170 (0.0189)
occupation					
self-employee	0.0644 (0.0677)	0.0270 (0.0244)	-0.00906 (0.0167)	-0.0265 (0.0272)	-0.0558 (0.0487)
Employee	0.0369 (0.0280)	0.0170 (0.0132)	-0.00312 (0.00363)	-0.0155 (0.0122)	-0.0352 (0.0261)
Unemployed	-0.0517* (0.0311)	-0.0321 (0.0207)	-0.0106 (0.0116)	0.0221* (0.0126)	0.0724 (0.0498)
Retired	-0.0796 (0.0512)	-0.0549 (0.0428)	-0.0290 (0.0400)	0.0326* (0.0170)	0.131 (0.116)
Other	0.103* (0.0590)	0.0380** (0.0170)	-0.0210 (0.0201)	-0.0410* (0.0219)	-0.0790** (0.0349)
Observations	656	656	656	656	656

Notes: Table A.10 demonstrates the marginal effect of ordered probit with *pgap* as dependent variable. The base category for perceived is Low/Low. Variable occupation uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.11: Marginal effects of ordered probit with *pgap* as dependent variables [use of *perceived* as independent variable]

	very low	low	medium	high	very high
perceived					
H/H	0.00777 (0.0316)	0.00336 (0.0139)	-0.000671 (0.00253)	-0.00314 (0.0129)	-0.00732 (0.0302)
L/H	-0.0240 (0.0495)	-0.0116 (0.0252)	0.000161 (0.00350)	0.00985 (0.0204)	0.0256 (0.0560)
H/L	-0.0101 (0.0381)	-0.00463 (0.0174)	0.000444 (0.00190)	0.00411 (0.0155)	0.0102 (0.0384)
riskprofile	-0.0115 (0.0118)	-0.00507 (0.00523)	0.000756 (0.001000)	0.00463 (0.00477)	0.0112 (0.0115)
prct_sec	0.0585*** (0.00943)	0.0259*** (0.00471)	-0.00386 (0.00299)	-0.0236*** (0.00417)	0.0569*** (0.0101)
exper	0.0666** (0.0285)	0.0294** (0.0129)	-0.00439 (0.00383)	-0.0269** (0.0116)	-0.0648** (0.0282)
loss	0.0166	0.00734	-0.00109	-0.00670	-0.0162

	(0.0117)	(0.00516)	(0.00118)	(0.00480)	(0.0113)
income	0.0102	0.00453	-0.000675	-0.00413	-0.00997
	(0.00717)	(0.00319)	(0.000689)	(0.00291)	(0.00702)
overc	-0.00391	-0.00173	0.000258	0.00158	0.00380
	(0.0235)	(0.0104)	(0.00155)	(0.00946)	(0.0229)
studyfield					
Mathematics, Engineering	-0.0112	-0.00481	0.000985	0.00451	0.0105
	(0.0329)	(0.0143)	(0.00280)	(0.0133)	(0.0313)
Economics, Management	-0.0157	-0.00685	0.00122	0.00634	0.0150
	(0.0256)	(0.0111)	(0.00246)	(0.0104)	(0.0240)
crt	0.0204	0.00902	-0.00135	-0.00824	-0.0199
	(0.0147)	(0.00651)	(0.00144)	(0.00594)	(0.0143)
numeracy	0.0138	0.00612	-0.000912	-0.00558	-0.0135
	(0.0246)	(0.0109)	(0.00176)	(0.00991)	(0.0239)
gender	0.0457**	0.0202**	-0.00301	-0.0184**	-0.0444**
	(0.0206)	(0.00908)	(0.00280)	(0.00837)	(0.0199)
age	0.00320**	0.00141**	-0.000211	-0.00129**	0.00311**
	(0.00132)	(0.000588)	(0.000185)	(0.000549)	(0.00128)
metro	-0.0126	-0.00558	0.000833	0.00510	0.0123
	(0.0194)	(0.00851)	(0.00146)	(0.00780)	(0.0188)
occupation					
self-employee	0.0653	0.0275	-0.00917	-0.0270	-0.0567
	(0.0705)	(0.0254)	(0.0176)	(0.0283)	(0.0505)
Employee	0.0381	0.0176	-0.00323	-0.0161	-0.0364
	(0.0281)	(0.0133)	(0.00371)	(0.0123)	(0.0262)
Unemployed	-0.0500	-0.0311	-0.0102	0.0215*	0.0697
	(0.0317)	(0.0212)	(0.0116)	(0.0130)	(0.0506)
Retired	-0.0856*	-0.0608	-0.0355	0.0345**	0.148
	(0.0519)	(0.0459)	(0.0478)	(0.0157)	(0.129)
Other	0.104*	0.0385**	-0.0212	-0.0414*	-0.0798**
	(0.0601)	(0.0173)	(0.0206)	(0.0223)	(0.0353)
Observations	656	656	656	656	656

Notes: Table A.11 demonstrates the marginal effect of ordered probit with *pgap* as dependent variable. The base category for perceived is Low/Low. Variable occupation uses “ Student” as reference, and the variable *studyfield* uses “non-mathematic and non-economic background” as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.12: Marginal effects of probit with *de* as dependent variables [use of *perceived* and *finbehav* separately as independent variable]

	M1	M2
finbehav	0.0289	
	(0.0427)	
perceived		
H/H		0.0824
		(0.0812)

L/H		0.2352	(0.1594)
H/L		0.0126	(0.0870)
overc	-0.1063**	-0.1677***	(0.0421) (0.0546)
loss	0.0007	-0.0012	(0.0254) (0.0252)
prct_sec	-0.0296	-0.0289	(0.0231) (0.0228)
exper	0.0026	0.0103	(0.0494) (0.0496)
riskprofile	-0.0315	-0.0331	(0.0267) (0.0265)
income	-0.0028	-0.0008	(0.0138) (0.0138)
studyfield			
Mathematics, Engineering	-0.0900	-0.0915	(0.0631) (0.0627)
Economics, Management	0.0491	0.0456	(0.0518) (0.0527)
crt	-0.0603*	-0.0566*	(0.0326) (0.0335)
numeracy	-0.1610***	-0.1649***	(0.0522) (0.0522)
age	-0.0015	-0.0015	(0.0023) (0.0023)
gender	-0.0366	-0.0434	(0.0496) (0.0503)
metro	0.0234	0.0210	(0.0422) (0.0419)
occupation			
self-employee	-0.0558	-0.0447	(0.0932) (0.0936)
Employee	-0.0198	-0.0037	(0.0607) (0.0602)
Unemployed	0.0946	0.1206	(0.1466) (0.1485)
Retired	-0.0705	-0.0502	(0.1538) (0.1555)
Other	0.1690	0.1755	(0.1336) (0.1280)
Observations	410	410	

Notes: Table A.12 demonstrates the marginal effect of probit with *de* as dependent variable, with the inclusion of *finbehav* (M1) and *perceived* (M2). The base category for perceived is Low/Low. Variable occupation uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as reference. In order to mitigate possible

heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A.13: Marginal effects of probit with *deI* as dependent variables [use of *finknow*, *perceived* and *finbehav* separately as independent variable]

	M1	M2	M3
finknow	-0.0348 (0.0361)		
finbehav		-0.0096 (0.0406)	
perceived			
H/H			-0.0183 (0.0899)
L/H			0.2245 (0.1659)
H/L			0.0416 (0.1016)
overc	-0.0231 (0.0409)	-0.0284 (0.0389)	-0.0367 (0.0488)
loss	-0.0034 (0.0244)	-0.0038 (0.0240)	-0.0058 (0.0241)
prct_sec	-0.0236 (0.0217)	-0.0208 (0.0215)	-0.0216 (0.0213)
exper	-0.0194 (0.0497)	-0.0047 (0.0494)	-0.0088 (0.0486)
riskprofile	-0.0723*** (0.0243)	-0.0663*** (0.0243)	-0.0710*** (0.0237)
income	-0.0027 (0.0130)	0.0003 (0.0127)	-0.0013 (0.0131)
studyfield			
Mathematics, Engineering	-0.0023 (0.0561)	0.0019 (0.0551)	-0.0026 (0.0537)
Economics, Management	0.1411*** (0.0467)	0.1536*** (0.0453)	0.1554*** (0.0468)
crt	-0.0142 (0.0335)	-0.0065 (0.0330)	-0.0005 (0.0345)
numeracy	-0.1044** (0.0491)	-0.1027** (0.0491)	-0.1017** (0.0493)
age	0.0031 (0.0022)	0.0030 (0.0022)	0.0028 (0.0022)
gender	-0.0546 (0.0470)	-0.0413 (0.0475)	-0.0452 (0.0464)
metro	-0.0152 (0.0408)	-0.0072 (0.0410)	-0.0162 (0.0406)

occupation			
self-employee	-0.0804 (0.0785)	-0.0779 (0.0771)	-0.0806 (0.0770)
Employee	0.0384 (0.0576)	0.0376 (0.0567)	0.0467 (0.0568)
Unemployed	0.3205** (0.1622)	0.3322** (0.1628)	0.3467** (0.1629)
Retired	-0.1675** (0.0792)	-0.1623* (0.0832)	-0.1609** (0.0784)
Other	0.1788 (0.1252)	0.1787 (0.1226)	0.1741 (0.1176)
Observations	415	415	415

Notes: Table A.13 demonstrates the marginal effect of probit with *deI* as dependent variable, with the inclusion of *finlow* (M1), *finbehav* (M2) and *perceived* (M3). The base category for perceived is Low/Low. Variable occupation uses "Student" as reference, and the variable *studyfield* uses "non-mathematic and non-economic background" as reference. In order to mitigate possible heteroskedasticity it is used robust standard errors. Robust Standard errors are in parentheses, *** p<0.01, ** p<0.05, * p<0.1