ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
Bachelor Thesis Economics & Business
Specialization: Financial Economics

GENDER INVESTMENT GAP: DOES GENDER HAVE A NEGATIVE EFFECT ON STOCK MARKET PARTICIPATION?

Author: Sofia Camilla Weydt
Student number: 558477
Thesis supervisor: Zhu, H
Second reader: Karpati, D
Finish date: 19 July 2024

Abstract

In this thesis I studied the negative relationship between gender and stock market participation, analyzing financial literacy and investor confidence. A survey was conducted with the aim of assessing respondents actual and perceived levels of financial literacy and testing the results against their gender through a series of Probit regressions. It was found that females are significantly less likely to invest in the stock market compared to males, as well as that they exhibit significantly lower levels of basic and advanced financial literacy and are significantly less confident in their financial knowledge than males are. This concludes that the gender investment gap is in fact a phenomenon occurring in real life and is driven by a combination of financial literacy levels and differences in self-confidence.

Table of Contents

1 Introduction

Over the past century, the feminist movement has grown exponentially and made great progress (O'Keefe, 2013). In the same period, private and household investing has seen similar growth, partly due to increased accessibility thanks to digitalization (Liang & Guo, 2015) and the people's increased want for financial independence and security in increasingly uncertain socio-political times (Ubide, 2024). However, the overlap of these two phenomena is a lot smaller than expected, with the gender gap in stock market participation and financial literacy increasing steadily (Muir, 2024). While male representation in the stock market in the United Kingdom has grown to over 1tn GBP in the period from January 2023 to January 2024, that same number for women does not even reach half that amount with 450bn GBP (Muir, 2024). While many ties have been made between gender and financial literacy (Bannier et al., 2019), financial literacy and stock market participation (van Rooij et al., 2011), gender and investor confidence (Yang & Li, 2016), there has been lacking evidence of the effect of gender directly on stock market participation. It is widely acknowledged in both the scientific field and in every-day life that women are generally more risk averse, while males are more likely to be overconfident and take impulsive decisions (Muir, 2024; Zhu et al., 2021). In this paper we are going to dig deeper into the gender investment gap and, on an individual level, the relationship between gender and stock market participation.

Socially, this research contributes to further understanding the issue that is the difference in financial stability between men and women (Hasler & Lusardi, 2017) by analyzing personal investment behaviors. By establishing a relationship between the lower financial literacy of females (Muir, 2024) and their participation in the stock market, government and policy makers are one step closer to tackling the problem at the root and, in the long term, reducing the discrepancy between male and female financial stability (Hasler & Lusardi, 2017). This would also significantly contribute to the ever-growing list of achievements for the feminist movement and bring society one step closer to achieving financial equality between men and women.

Scientifically, this research widely contributes to the field of behavioral finance, giving an insight into the factors underlying female investment activity, and diversifying whether financial literacy is in fact the main underlying factor, or perhaps other aspects such as self confidence might play a significant role. This paper adds to the existing research on the effect of financial literacy and stock market participation (van Rooij et al., 2011), as well as the relationship between gender and financial literacy (Bannier et al., 2019), while filling the gap of lacking research on the effect of gender on stock market participation itself. Specifically, this paper completes the pre-exisiting work of van Rooij et al. (2011), linking financial literacy to stock market participation, by adding a new level to the analysis, which in this case is the gender effect.

Behavioral Finance researchers have been analyzing the gender investment gap for decades now. The common consensus is that there is in fact a gap in financial literacy between men and women (Lanciano et al., 2024; Potrich et al., 2018: Hasler & Lusardi, 2017). Further research has been conducted on more specific markets and asset classes, yet always coming to the same conclusion that women exhibit significantly lower financial literacy levels both in emerging markets such as Brazil (Lanciano et al., 2024), and in more niche asset classes such as cryptocurrencies (Bannier et al., 2019). Similarly, the relationship between gender, investor confidence (or lack thereof), and risk perception has been widely studied, again finding significantly lower investor confidence and stronger financial risk perception in females than in males (Zhu et al., 2021; Yang & Li, 2016). Finally, a significant relationship has been established between degree of financial literacy and stock market participation (van Rooij et al., 2011). Specifically, the finding was that while most people exhibit very basic financial literacy, only a very small proportion has knowledge that goes beyond the most basic concepts, and by extension the low-spread financial literacy significantly decreases probability of stock market participation (van Rooij et al., 2011). The differentiation of basic and advanced financial literacy has been adopted as well in this research, as it was found that the significant effect on stock market participation was among those with low advanced financial literacy (van Rooij et al., 2011).

However, none of the above-mentioned research analyses the effect of gender directly on stock market participation, despite the lack of female representation in the stock market being acknowledged by many professionals in the field (Muir, 2024) and the common variables among said research papers. This leads us to the construction of our overarching research question, namely:

"Gender Investment Gap: Does gender have a negative effect on stock market participation?"

In order to answer the research question, the analysis is separated into three main null hypotheses. The first null hypothesis is:

$H_0 1$*: "There is no effect of gender on stock market participation, ceteris paribus."*

This hypothesis is our main hypothesis, as it includes our dependent and independent variables as such, and directly answers our research question. This hypothesis is tested against the alternative hypothesis of a significant effect of gender on stock market participation.

The second hypothesis in analysis is:

$H_0 2$*: "There is no gender effect on financial literacy levels, ceteris paribus."*

Due to the aim of our research, we further separate this hypothesis into two sub-hypotheses.

$H_0 2.1$: "*There is no gender effect on basic financial literacy levels, ceteris paribus.*"

$H_0 2.2$: "*There is no gender effect on advanced financial literacy levels, ceteris parib*us."

All these hypotheses are tested against an alternative hypothesis of a significant effect of gender on financial literacy levels. Specific definitions of financial literacy and gender effect will follow in the theoretical framework.

The third and final hypothesis in analysis is:

$H_0 3$*: "Both genders exhibit the same levels of investor overconfidence, ceteris paribus."*

Due to our data set, we can further separate this hypothesis into three sub-hypotheses.

$H_0 3.1$: *"Males exhibit higher levels of overconfidence on basic financial literacy, ceteris paribus.*"

$H_0 3.2$: "*Males exhibit higher levels of overconfidence on advanced financial literacy, ceteris parib*us."

This hypothesis serves the purpose of analyzing whether males are more confident in their financial literacy levels, which might explain a stronger male stock market representation even in cases where financial literacy levels between males and females might not be significantly different. Specific definitions of investor confidence and how it was measured for this research will follow in the theoretical framework. This hypothesis is tested against the alternative hypothesis of a significant effect of gender on investor confidence.

The rest of this research paper is going to be split into 6 sections: Theoretical Framework, where we further analyze the existing research and define the concepts used, Data Analysis, where we describe our data set, Methodology, where we explain the methods used, Results, where we describe our results, Discussion and conclusion, where we draw conclusions from our results and discuss advantages and limitations, and finally the References.

2 Theoretical Framework

The analysis in this research paper is divided into three groups of hypotheses, each aiming to test the relationship between gender and certain financial metrics. These hypotheses were constructed with help of five pre-existing research papers that each touch on one or multiple aspects tested in our hypotheses, and together they aim to answer our leading research question.

2.1 Stock Market Participation

This research paper is based on previous work by multiple professionals in the field, with a strong focus on van Rooij et al.'s (2011) "Financial Literacy and Stock Market participation". The authors of this paper analyze the relationship between financial literacy and stock market participation. Their data set is derived from a survey from the 2005 de Nederlandsche Bank's Household Survey (DHS). The survey in question is conducted annually and covers wealth and saving data, as well as demographics. The panel is run by an institute at Tilburg University called CentERdata, it contains 2028 households and is a representative sample of the Dutch population. Additionally, van Rooij et al. (2011) include self-perceived financial literacy levels in the survey, which acted as inspiration for including investor confidence as a moderator in our research. Van Rooij et al. (2011) find a significant relationship between levels of both basic and advanced financial literacy and stock market participation, establishing that not only most people possess basic financial literacy, but also that of those people only a small minority exhibit advanced financial literacy. This finding is what motivated the separation of basic and financial literacy in this paper. Furthermore, the authors find that of those people that do not possess advanced financial literacy, a significant amount do not participate in the stock market (van Rooij et al., 2011).

In addition to van Rooij et al.'s (2011) research, Lanciano et al.'s (2024) research has a similar starting point, analyzing the relationship between financial resilience and stock market participation, controlling for financial literacy. It is found that the financial literacy component of their survey spills over to other correlated variables such as gender (Lanciano et al., 2024). The overarching finding is that, when adjusting for measurement error, the gender gap in financial resilience completely disappears, and that said measurement error might skew results in analyses about widely discussed financial outcomes, such as stock market participation (Lanciano et al., 2024).

The two abovementioned papers exhibit significant similarities in that both examine the relationship between financial literacy and stock market participation (van Rooij et al., 2011; Lanciano et al., 2024). However, Lanciano et al.'s (2024) research highlights the relevance of measurement error, especially in measurement of financial literacy, and the existence of a gender gap in financial literacy, as well as financial resilience and stock market participation. In contrast, van Rooij et al. (2011) perform a much more detailed

test of financial literacy, which might minimize the measurement error. However, gender is not addressed more than using it as a control variable for one of their regressions. These two research papers help us define and quantify the concepts needed to test against our primary null hypothesis.

For the context of this paper, the dependent variable, stock market participation, will be defined as an individual's investing their own private cash into the stock market, or their having done so in the past (van Rooij et al., 2011). Investments in alternative asset classes such as Cryptocurrencies and real estate will not be taken into consideration. This variable is defined as a simple yes or no answer, accounting for both current and past stock market participation.

2.2 Financial Literacy

Another research that successfully links gender to financial literacy performance is Potrich et al.'s (2018) "How well do women do when it comes to financial literacy? Proposition of an indicator and analysis of gender differences", focusing primarily on emerging markets, in this case Brazil. In this paper the authors establish a significant relationship between gender and financial literacy, with women being significantly underrepresented among those with high financial literacy (Potrich et al., 2018). Nonetheless, it is found that financial literacy levels are low across the board both within the females and the males of the sample of 2485 Brazilians, and that the gender difference in financial literacy is almost exclusively reserved to the highly financially literate population (Potrich et al., 2018). While this result seems to confirm my theory of a gender effect on financial literacy, the context in which said effect is found, meaning only among the highly financially literate minority, does appear surprising.

Furthermore, it is found in Hasler and Lusardi's (2017) "The Gender Gap in financial Literacy: A Global Perspective" that lacking financial literacy is a phenomenon that covers the globe, with only one in three people having sufficient knowledge of finance and economics (Hasler & Lusardi, 2017). Additionally, it is established that the gender differences in financial literacy levels are significant in all countries, irrespective of culture and markets (Hasler and Lusardi). However, the extent and way the financial literacy gender gap presents does vary depending on the country in analysis (Hasler & Lusardi, 2017). This finding does not indicate whether the second hypothesis will be rejected or not.

Nevertheless, Potrich et al.'s (2018) and Hasler and Lusardi's (2017) conclusions are contradicted by Lanciano et al.'s (2024) findings, as Lanciano et al. (2024) sustain that a simple adjustment for measurement error would nullify the gender financial literacy gap, or at least reduce it to insignificant amounts. Based on the three papers, it is unclear whether the financial literacy indicator (Potrich et al., 2018) is immune to the measurement error due to the measurement method (Lanciano et al., 2024). The combination of the three articles mentioned give us the definitions and concepts needed to test our second set of null hypotheses.

When researching financial literacy differences among gender in the context of alternative asset classes, Bannier et al's (2019) „The gender gap in 'Bitcoin literacy" confirms that the phenomenon of reduced financial literacy in women (Lanciano et al, 2024) not only holds for emerging markets (Potrich et al, 2018), but also for untraditional asset classes. The authors analyze the levels of bitcoin literacy, defined as the knowledge and understanding of Bitcoin (Bannier et al, 2019) across a nationally representative US sample. It is found that socio-demographic factors do not exhibit a significant effect on bitcoin literacy (Bannier et al., 2019), but that around 40% of the gender gap in Bitcoin literacy can be explained by measures of actual and perceived literacy, which resembles the 7-layer scale of self-perceived basic and advanced financial literacy (van Rooij et al., 2011).

For the context of this paper, financial literacy is defined as the level of financial knowledge an individual displays, such as the understanding of concepts like inflation, time value of money, the difference between stocks and bonds etc. (van Rooij et al., 2011). It is divided into two layers, basic financial literacy and advanced financial literacy (van Rooij et al., 2011), and is measured following the same index as is used in van Rooij et al.'s (2011) research. Basic financial literacy is defined as having a grasp of basic financial concepts such as inflation and time value of money, while advanced financial literacy is defined as understanding the functionings of the stock market, bonds, stocks and savings accounts, as well as risk diversification (van Rooij et al., 2011). The index used is composed of a scale of 7 layers, ranging from 1 (very low) to 7 (very high), with the in-between categories being 2 (low), 3 (moderate), 4 (average), 5 (advanced) and 6 (high).

2.3 Investor Confidence

Our fourth set of hypotheses is centered around gender and investor confidence. Yang and Li's (2016) "Ambiguity risk: An experimental study of overconfidence, gender and trading activity." touches exactly on these topics. The authors find that males exhibit significantly higher degrees of investor overconfidence, and that they trade significantly more than females do (Yang & Li, 2016). This result suggests that both the primary null hypothesis of no gender effect on stock market participation, as well as the third set of null hypotheses of no gender effect on investor overconfidence will be rejected (Yang & Li, 2016).

Similarly, Zhu et al.'s (2021) "Interaction and decomposition of gender difference in financial risk perception." find a significant relationship between financial risk perception amongst men and women, with the latter exhibiting higher levels of risk aversity. Additionally, the authors link the difference in risk perception to potential overconfidence on the male's side, specifically focusing on the realm of investments (Zhu et al., 2021). However, this paper also establishes the relevance of other environmental factors that have a joint effect with gender on risk perception, such as culture and age (Wang et al., 2021). This cultural

factor might explain the previously mentioned surprising finding of the gender difference in financial literacy in Brazil (Potrich et al, 2018). However, this goes against Bannier et al.'s (2019) finding that socio-demographic factors do not explain variations in Bitcoin literacy across genders, although that might be related to Cryptocurrencies being an alternative asset class, and the sample being only limited to the United States of America (Bannier et al, 2019).

The two aforementioned papers establish a significant correlation between gender and financial risk perception (Yang & Li, 2016; Zhu et al., 2021), where the latter focus on risk perception and risk aversity patterns, which results in the differences in investor confidence highlighted by the former. In this research paper I aim to test for a gender effect on investor confidence through our fourth set of hypotheses, based on the assumption that women have a stronger perception of risk, and therefore are less confident in their investments (Yang & Li, 2016). Additionally, Yang and Li (2016) find a significant relationship between male and female investor confidence and the reduced trading activity of women. This finding represents another plausible reason for the lack of female representation in the stock market (Muir, 2024) and therefore adds another layer to the answer to our leading research question, suggesting the third hypothesis might be rejected.

Although the inclusion of investor confidence might seem unrelated to our other 2 hypotheses, the choice of including this third set of hypotheses was indeed very intentional. That is, although van Rooij et al.'s (2011) findings did not particularly focus on gender, once having established a relationship between financial literacy and stock market participation (van Rooij et al, 2011), it appears logical that a person with lesser knowledge of the mechanisms of the world of finance is less confident in their financial decision making, perceive risk more strongly (Zhu et al., 2021), and hence have a smaller likelihood of investing their own money in the stock market (Yang & Li, 2016). Having seen that women exhibit higher risk aversity (Zhu et al., 2021), and are generally less confident in their investing abilities (Yang & Li, 2016), it is logical to include investor confidence as a hypothesis in testing the gender effect on stock market participation.

For the context of this paper, overconfidence will be defined as having a perception of one's own level of financial literacy higher than one's actual financial literacy. This will be tested based on the results of the self-assessed basic and advanced financial literacy scales (van Rooij et al, 2011) in comparison to one's results on the basic and advanced financial literacy assessment (van Rooij et al., 2011). The aim of the third set of hypotheses is to see whether there is a significant difference in overconfident individuals between the group of male and female respondents.

Finally, Hasler and Lusardi (2017), as well as Lanciano et al. (2024) and Potrich et al. (2018) establish a significant relationship between gender and levels of financial literacy. The overarching finding is that financial literacy is lacking in all markets (Hasler and Lusardi, 2017; Potrich et al., 2018). However, the specifics of the observed gender gap in financial literacy vary depending on the research, with Hasler and Lusardi (2017) establishing a global trend, with variations depending on cultural factors and market situations, which is backed by Poitrich et al.'s (2018) findings on financial literacy gender gap in Brazil. All three of the aforementioned papers focus on gender and financial literacy, while our leading research by van Rooij et al. (2011) only briefly touches on the gender effect on financial literacy, the combination of these four articles tying gender to financial literacy, and ultimately financial literacy to stock market participation, suggest that, thanks to our three sets of hypotheses, our research question will be answered by the end of this report.

3 Data Analysis

For this research, we have 16 variables in analysis, ranging from basic demographics to scales of financial literacy levels and confidence levels. The dataset used consists of 196 respondents and is retrieved from the survey "Financial Literacy and Stock Market Participation" designed to replicate the financial literacy assessment questions and scales used by van Rooij et al. (2011), retrieved from the 2005 de Nederlandsche Bank's Household Survey (DHS). The survey had the aim of assessing respondents' levels of actual and self-perceived financial literacy, using questions of numeracy, time value of money, risk differentiation, inflation and others (van Rooij et al., 2011). It was run on Erasmus University Qualtrics and was adapted to include demographics and metrics that would be useful to this analysis, as well as removing data that would be irrelevant for the means of this paper. See appendix C for the exact questions and answer options of the survey.

The dependent variable in question is stock market participation (stockmarket). This is a binary variable that takes value 1 when the individual has invested in the stock market in the past or in the present, and 0 when the individual has never invested in the stock market.

To measure financial literacy there are four categorical variables, all following the same scale of 1 to 7, with 1 as the lowest and 7 as the highest, which was adapted from the scale used in our inspirational paper (van Rooij et al., 2011). Firstly, the self-assessed level of basic financial literacy (sbl1), indicating the level of basic financial knowledge the individual reports once having answered the basic financial literacy questions. Secondly, the actual basic financial literacy (actual_basic), which is the level of basic financial literacy an individual presents based on their answers in the basic financial literacy assessment. An individual will have level 7 "very high" basic financial literacy if they get all 5 basic financial literacy questions right, level 1 "very low" if they get none right, level 2 "low" if they get only one right, level 2 "moderate" if they get two right, level 5 "advanced" if they get 3 right, sand level 6 "high" if they get 4 correct. Thirdly, the self-assessed advanced financial literacy (sal1), as stated by the respondent after completing the advanced financial literacy assessment, and finally, actual advanced financial literacy (actual_advanced), as the result of their advanced financial literacy test results. An individual will be a level 7 "very high" if they had all 11 questions right, level 6 "high" if they had 9 or 10 questions right, level 5 "advanced" if they had 7 or 8 questions right, level 4 "average" if they had exactly 6 questions right, level 3 "moderate" if they had 4 or 5 questions right, level 2 "low" if they had 3 or 2 questions right, and 1 "very low" if they had one or no correct answers.

To measure investor confidence, we compare each respondent's self-assessed basic and advanced financial literacy level and compare it to their actual performance in the financial literacy assessment. We hence

construct 6 dummy variables: under confident on basic knowledge (underconfident_basic), taking 1 when an individual's self-assessed basic financial literacy is lower than their actual performance and 0 otherwise, under confident on advanced knowledge (underconfident_advanced), taking 1 when an individual's self-assessed advanced financial literacy is lower than their performance in the assessment and 0 otherwise, confident on basic knowledge (confident_basic), taking 1 when an individual's self-assessed and actual basic financial literacy coincide and 0 otherwise, confident on advanced knowledge (confident_advanced), taking 1 when an individual's self-assessed and actual advanced financial literacy coincide and 0 otherwise, and finally, over confident on basic knowledge (overconfident_basic) and over confident on advanced knowledge (overconfident_advanced), taking 1 when an individual's self-assessed financial literacy exceeds their performance on the respective assessments and 0 otherwise. Additionally, the basic and advanced confidence dummies were merged into two categorical confidence variables, basic confidence (basic_confidence) and advanced confidence (advanced_confidence).The former takes 1 if the individual is underconfident in their basic knowledge, 2 if it has an accurate assessment of its basic financial literacy, and 3 if it is overconfident about its basic financial literacy, and the latter takes the same values with respect to the level of confidence on advanced financial literacy.

Finally, some demographic variables were included as main independent variable and additional control variables. Firstly, gender, a categorical variable that takes 0 if the respondent is a male, 1 if the respondent is a female, and 2 if the respondent chose the option "other/prefer not to say". Secondly, a categorical variable for age (age) was included, taking 1 when the respondent is between 18 and 24 years old, 2 when the respondent is between 25 and 29 years old, 3 if he is between 30 and 39 years old, 4 if he is between 40 and 49 years old, and 5 if he is 50+ years old. The selection of these age groups was made based on the assumption that, due to the environment in which I live, a big majority of respondents will be aged between 18 and 30, and very few respondents will be older than 50 years old. Thirdly, a categorical income variable (income) is used, taking 1 if the respondent has an income below Dutch average of 47000€ (CBS, 2023), 2 if the respondent earns approximately the Dutch average of 47000€, 3 if he earns above Dutch average of 47000€, and 4 if the respondent preferred not to disclose that information. Lastly, respondents are asked whether their studies had a Business or Economics major (major), the variable takes 1 if the respondent did study something focused on these subjects, and 0 otherwise.

Some adjustments were made to the dataset to ensure all our observations are valid and usable. Firstly, all observations in which the respondent does not answer all questions in the survey, and hence contain one or multiple missing variables, are removed. Secondly, our gender variable is transformed from a categorical variable that includes the "Other/Prefer not to say" option, to a binary variable (female), taking 1 if the respondent is female and 0 otherwise. This adjustment does not eliminate any usable observations as no

respondents selected the “Other/Prefer not to say” option. After the abovementioned modifications of the dataset, the final sample consists of 165 observations, of which 84 are female, and 81 male.

To provide a clear understanding of the data, we first examine the descriptive statistics. This section summarizes key aspects such as means, medians, standard deviations, and percentiles, offering a foundational overview of the dataset. The descriptive statistics of all relevant variables can be seen in Table 1. The frequency distributions of gender and the overall self-assessed financial literacy levels are visible in Table 2, and the frequency distribution of gender and stock market participation is found in Table3.

With the two most important dummies, stock market participation and gender, an identical distribution is observed, with a minimum of 0 and a maximum of 1, the mean of 0.509 indicates that 50.9% of the observations in our dataset are female, as well as 50.9% of our dataset currently or previously has invested in the stock market. Both variables exhibit a standard deviation of 0.501 and are composed of 165 observations, and coincidentally have a median of 1, the 25th percentile of 0, and the 75th percentile of 1.

Additionally, a difference of almost one whole category between actual basic financial literacy and actual advanced financial literacy is observed, both composed of 165 observations and a range of 1 to 7. The mean for actual basic financial literacy is 5.454, indicating that most respondents have an actual basic financial literacy are between a level 5 “advanced” and 6 “high”. However, the mean for actual advanced financial literacy is 4.509, meaning most respondents has an advanced level of financial literacy that lies between level 4 “average” and 5 “advanced”. The standard deviation for actual basic financial literacy is 1.698 and the standard deviation for actual advanced financial literacy is 1.823, indicating that the difference in overall variation for both variables is not large. The median for basic financial literacy lies at 6, indicating that 60% of observations are at or below high basic financial literacy. The 25th and 75th percentiles for basic financial literacy are 5 and 7, respectively. As for advanced financial literacy, the median lies at 5, indicating that 50% of observations exhibit advanced levels advanced financial literacy or lower, with the 25th and 75th percentiles being 3 and 6, respectively. This indicates that less than 25% of respondents have very high advanced financial literacy.

Furthermore, basic and advanced confidence are analyzed. Both variables are composed of 165 observations and range from 1 to 3. The former has an average of 1.539 and a standard deviation of 0.711, indicating that most respondents lie between the level 1 “underconfident” and 2 “confident” on their level of basic financial literacy. The median for basic financial literacy confidence is 1, indicating that exactly 50% of respondents are underconfident in their basic financial literacy. The 25th and 75th percentiles for this variable lie at 1 and 2, respectively, indicating that less than 25% of respondents are overconfident in their basic financial literacy. Advanced financial literacy has a mean of 1.764, indicating most respondents lie

between the level 1 “underconfident” and level 2 “confident” in their advanced financial literacy. The variable has a standard deviation of 0.855, and a median of 1. This again indicates that 50% of respondents are underconfident in their advanced financial literacy levels. The 25th and 75th percentiles for this variable are 1 and 3, respectively, indicating that at least 25% of respondents are overconfident in their advanced financial literacy levels.

Finally, the control variables, age, income, education, and major, all have 165 observations. The age category has a mean of 1.770, indicating that most respondents are between the age category 1 “18 to 24 years old” and age category 2 “25 to 29 years old”. The variable has a large standard deviation of 1.692, with a range between 1 and 5, and a median of 1, indicating that at least 50% of respondents are part of age group 1. The 25th and 27th percentile for the age category are 1 and 5, respectively, which also indicates that at least 25% of all observations are part of age group 5 “50+ years old”.

The income category has a mean of 1.721, suggesting that most respondents are between income groups 1 “below Dutch national average” and income group 2 “approximately Dutch National average”. The standard deviation for income is 1.016, which is quite large considering a range of 1 to 4. The median for income is 1, indicating that at least 50% of observations are part of income group 1 of income below Dutch national average. By extension, the 25th percentile is also 1, and the 75th percentile is 2, indicating that at least 25% of respondents are of income groups 2 or above.

Table 1 shows a mean for the education category of 3.358, indicating that most respondents lie between education category 3 “Bachelor’s” and education category 4 “Master’s / MBA”. The standard deviation for this variable is relatively low with only .634. The range of observations is 2 to 5, indicating that no respondents are part of category 1 “less than Highschool”. The median for this variable is 3, indicating that at least 50% of observations are part of age category 3 “Bachelor’s” or below. The 25th and 75th percentile of the variable are 3 and 4 respectively, suggesting that at least 25% of observations are part of education category 3 “Bachelor’s”, and that at least another 25% of observations is part of education categories 4 or above, meaning either “Master’s/MBA” or “PhD+”.

Lastly, the major dummy shows a mean of .558, indicating that 55.8% of respondents have a business or economics major. The standard deviation for this variable is .498, and the range is from 0 to 1. The median for this variable is 1, suggesting that the 75th percentile is 1 as well, and indicating that at least 50% of observations have a business or economics major. The 25th percentile is 0, indicating that at least 25% of observations do not have this background.

4 Methodology

The analysis of the data was performed by transferring the responses from Erasmus University Qualtrics to STATA. Once having created the survey "Financial Literacy and Stock Market Participation" and collected all the relevant data, we proceed to testing our 4 sets of hypotheses. All the following regressions were constructed with the relevant dependent and independent variables, using age, income level, education level, and major as control variables.

To test the first null hypothesis of no gender effect on stock market participation, a Probit regression is run, with the dependent binary variable stock market participation and the independent binary variable female, while controlling for age, income, major, and education. The final Probit regression used to test our first null hypothesis is:

$$Probit\big(P(stockmarket_i = 1)\big) = \phi^{-1}\big(P(stockmarket_i = 1)\big)$$
$$= \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 income_i + \beta_4 education_i + \beta_5 major_i + \varepsilon_i$$

Subsequently, a test for normality of residuals was run by manually generating the residuals as the difference between the prediction and the actual observation, and consequently plotted into a histogram, visible in Figure 1. In this case, robust standard errors were added to the regression. The results from this regression can be seen in Table 4a.

After estimating the first Probit regression model, we calculated the average marginal effects of the female coefficient. The marginal effects provide the change in the probability of past or current stock market participation associated with an individual being female rather than male, holding other variables constant. This approach allows for a more intuitive interpretation of the results, as it translates the coefficients from the Probit model into changes in probability. These results can be seen in Table 4b. Then, a Wald test was run on the significance of the female coefficient of this regression, the results of which are reported in Table 4c.

To test the second set of null hypotheses of no gender effect on basic financial literacy and no gender effect on advanced financial literacy two Ordered Probit regressions were constructed. Firstly, the gender effect on basic financial literacy is tested, controlling for age, income, education, and major, leading us to the following regression:

$$actual\_basic_i^* = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 income_i + \beta_4 education_i + \beta_5 major_i + \varepsilon_i$$

Next, the normality of the residuals was tested by manually creating the residuals on STATA by subtracting the actual observation from the prediction and plotting a histogram of the residuals, found on Figure 2. Again, robust standard errors were added to the regression.

The results from this Ordered Probit regression can be found in Table 5a. Consequently, the average marginal effects of the female variable were calculated to ease the interpretation by transforming the z scores found in Table 5a into probabilities. These results are found in table 5b. Additionally a Wald test on the significance of the female variable was performed, the results of which are shown in Table 5c.

Secondly, null hypothesis 2b is tested, namely the gender effect on advanced financial literacy, through an Ordered Probit model with advanced financial literacy as dependent and gender, age, income, education, and major as independent variables. The Ordered Probit regression to test hypothesis 2b is:

$$actual\_advanced_i^* = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 income_i + \beta_4 ieducation_i + \beta_5 major_i + \varepsilon_i$$

Thereafter, the normality of residuals was tested by subtracting the actual observation from the predicted one, hence manually creating the residuals on STATA and then plotting a histogram, found in Figure 3. Once more, robust standard errors were added to the regression. The results of the regression can be found in Table 6a. To conclude the testing of our second set of hypotheses, the marginal effects of the female variable for the second Ordered Probit regression were calculated and depicted in Table 6b. A Wald test was conducted to test the significance of the female coefficient and its results were depicted in Table 6c.

Finally, the third set of null hypotheses of no gender effect on investor confidence is tested. First, hypothesis 3.1 is tested with an Ordered Probit model, using age, income, education, and major as control variables. The regression to test hypothesis 4.1 is:

$$basic\_confidence_i^* = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 income_i + \beta_4 education_i + \beta_4 major_i + \varepsilon_i$$

After testing for the normality of residuals as previously described, robust standard errors were added to the regression. See Figure 4 for the distribution of the residuals. The results of this ordered Probit regression can be found in Table 7a. Additionally, the average marginal effects for the female variable were calculated to ease the interpretation of results, said marginal effects can be seen in Table 7b. Again, a Wald test for significance of the female variable was run and its results are shown in Table 7c.

Subsequently, null hypothesis 3.2 of no gender effect on investor advanced financial literacy confidence is tested. The chosen regression is an Ordered Probit due to the categorical quality of the dependent variable, controlling for actual age, income, education, and major. The final regression to test hypothesis 4.2 is:

$$advanced\_confidence_i^* = \beta_0 + \beta_1 female_i + \beta_2 age_i + \beta_3 income_i + \beta_4 education_i + \beta_4 major_i + \varepsilon_i$$

The normality of residuals was tested by manually creating the residuals on STATA through subtracting the actual observation from the prediction, and said residuals were then plotted into a histogram, as seen in Figure 6. The results of the regression with the robust standard errors are seen in Table 8a. To ease interpretation, the average marginal effects of the independent variable of interest were calculated and shown in Table 8b. To test the significance of the female variable a Wald test was conducted and reported in Table 8c.

5 Results

The results of this study are discussed below, aiming to provide interpretations of the statistical analysis to reject or accept our three sets of hypotheses and aid the answering of our research question.

Hypothesis 1

Null hypothesis 1 of no gender effect on stock market participation was tested with Probit regression 1, depicted in Table 4a. It is found that the female coefficient has a z value of -3.68, and hence is significant at the 1% significance level. The coefficient of -.277 in the marginal effects table, as seen in Table 4b, indicates that the probability of participating in the stock market decreases by 27.7% when the individual is a female compared to when it is a male, holding all other variables equal. The Wald Test shown in Table 4c has a test statistic of 13.57, making it significant at the 1% significance level with a p value of 0.0002. The therefore reject the null hypothesis 1 of no gender effect on stock market participation at the 1% significance level.

Table 4a additionally shows that education level 5 is significant at the 5% significance level, with a z score of 2.23. The positive coefficient suggests that having a PhD or even more education has a positive effect on the likelihood of investing in the stock market. The dummy variable major also is significant at the 5% significance level, with a positive coefficient suggesting that having a major in economics or business has a positive effect on the likelihood of stock market participation, as the z score is 2.21.

Finally, Figure 1 shows the distribution of the stock market participation variable after running the Probit regression. The distribution if residuals is plotted against a normal distribution curve to aid interpretation. As the residuals of the variable have a distribution completely unrelated to that of the normal distribution, the null hypothesis of equal residuals is rejected and robust standard errors are added to the Probit regression.

Hypothesis 2.1

The second set of null hypotheses of no gender effect on financial literacy is divided into two sub hypotheses, each tested with an Ordered Probit regression.

The first Ordered Probit regression results are depicted in Table 5a, and its marginal effects in Table 5b. It is found that gender has a significant negative effect on basic financial literacy, with a z value of -1.68. Table 5b shows the marginal effects of gender on basic financial literacy, indicating that gender has a significant effect on basic financial literacy level 2 “low” with a coefficient of .037 and 7 “very high” with a coefficient of -.091. Specifically, being female increases the probability of having low basic financial literacy by 3.7% compared to being male, all else equal, and it decreases the probability of having very high

basic financial literacy by 9.1%, all else equal. The respective z scores for these marginal effects are 1.70 and -1.68, making both significant at the 10% significance level. Additionally, Table 5c shows the Chi2 test statistic of the Wald test on the significance of female to be 2.84, with 1 degree of freedom, making the variable significant at the 10% significance level with a p-value of 0.0922. Therefore, we reject the null hypothesis of no gender effect on basic financial literacy at the 10% significance level.

Furthermore, Table 5a shows that age categories 2 and 5 have significant effects at the 5% significance levels. The respective z scores are -1.68 and 2.08, suggesting that being 25 to 29 years old, compared to being 18 to 24 years old, on average decreases one's level of basic financial literacy, while being 50 years old or older in comparison to being 18 to 24 years old on average increases one's basic financial literacy, all else equal.

Moreover, it is found that education levels 4 "Master's/MBA" and 5 "PhD+" are significant at the 10% and 5% levels, respectively, with z scores of 1.45 and 1.98. The positive coefficients suggest that, in comparison to having only a Highschool degree, having a Master's or MBA on average increases one's basic financial literacy, all else equal. Similarly, having a PhD or more, compared to having only a Highschool degree, on average increases one's level of basic financial literacy, all else equal.

Additionally, the major coefficient is significant at the 1% significance level with a z score of 5.91. The positive z-score and coefficient suggest that individuals with a business or economics major have significantly higher basic financial literacy levels, on average, compared to those who do not have that major, all else equal.

Figure 2 shows the distribution of the residuals of the basic literacy coefficients based on the Ordered Probit regression plotted against a normal distribution curve for ease of interpretation. Since the distribution of the residuals does not follow the curve, robust standard errors are added to the regression.

Hypothesis 2.2

Null hypothesis 2.2 of no gender effect on advanced financial literacy is tested in a second Ordered Probit regression, the results of which are depicted in Table 6a, and the marginal effects of the female variable are depicted in Table 6b. Again, a negative gender effect on advanced financial literacy is detected, with the variable female having a z score of -2.47. Table 6b shows that the independent variable of interest, female, has a significant effect on all advanced financial literacy levels except level 5 "advanced". All significant coefficients are significant at the 5% significance level, with their z scores being 2.32, 1.97, 2.25, 2.19, -2.37 and -2.33 going from lowest to highest level. Specifically, this means that being female, compared to being male, increases the probability of having very low, low, moderate, and average advanced financial

literacy by 6.2%, 2.1%, 3.4%, and 2.1%, respectively, all other factors equal. Additionally, it reduces the probability of having high or very high advanced financial literacy by 6.7% and 7.2%, respectively, all else equal.

Additionally, the Wald test statistic of 6.12 with 1 degree of freedom suggests that the female coefficient is significant at the 5% significance level, with a p value of 0.0134. We therefore reject our null hypothesis of no gender effect on advanced financial literacy at the 5% significance level.

Table 6a shows that the coefficient of age is significant at the level 5, with a z-score of 2.08, making it significant at the 5% level. The coefficient indicates that being 50 years old or above on average increases the advanced financial literacy level compared to those who are between ages 18 to 24. Additionally, the education variable is significant at levels 3, 4, and 5, with the first and the last significant at the 5% level, and level 4 being significant at the 1% level. The respective z-scores are 2.03, 2.58, and 2.24 and they indicate that having a Bachelor's degree, a Master's degree, or a PhD significantly increase, on average, the advanced financial literacy levels, compared to having only a Highschool degree, all else equal. Finally, major has a significant positive effect at the 1% significance level, with a z-score of 6.68, indicating that having a major in economics and business on average increases advanced financial literacy compared to not having such a major, leaving all else equal.

Figure 3 shows the distribution of residuals of the dependent variable advanced financial literacy in this regression plotted against a normal distribution curve. Seeing as the residuals do not follow the same distribution as the plotted curve, robust standard errors were added to the equation.

Hypothesis 3.1

The first null hypothesis of our last set tests the effect of gender on basic investor confidence levels and is tested via an Ordered Probit model. The results of the model are depicted in table 7a and the marginal effects of the variable of interest in Table 7b. It is found that being female has a negative effect on basic investor confidence, with a z score of -3.29 it is significant at the 1% significance level.

Table 7b shows that the female coefficient has marginal effect coefficients on under confidence, accurate confidence, and advanced confidence of .321, -.109, and -.122, respectively. This indicates that being female increases the probability of being under confident in one's basic financial literacy by 23.1%, compared to being male, as well as decrease the probability of being confident and overconfident in one's basic financial literacy levels by 10.9% and 12.2% respectively, all else equal. These results are all significant at the 1% significance level, with z scores of 3.65, -3.32, and -3.22, respectively.

Additionally, Table 7c shows that the Wald test conducted on the significance of the female coefficient has a test statistic of 10.82, with 1 degree of freedom, making it significant at the 1 % significance level with a p value of 0.0010. Therefore, we can reject the null hypothesis of no gender effect on basic investor confidence at the 1% level.

Furthermore, the age coefficients for levels 2 and for are significant at the 1% significance levels. The respective z-values are 2.85 and -12.30, indicating that being in age group 25 to 29 on average increases investor confidence levels compared to age group 18 to 24, and being in age group 40 to 49 on average decreases confidence compared to age group 18 to 24, all else equal.

Figure 4 shows the distribution of residuals of the dependent variable basic confidence based on this model plotted against a normal distribution curve. Since the distribution of the residuals does not match the normal distribution curve, the assumption of normality of residuals is rejected and robust standard errors are added to the regression.

Hypothesis 3.2

Null hypothesis 3.2 of no gender effect on advanced financial literacy confidence is tested in an Ordered Probit regression, with its results depicted in Table 8a and the marginal effects of the female variables in Table 8b. It is found that the variable female has a significant negative effect at the 1% level, with a z score of -3.57. Table 8b shows that females, compared to males, have an increased probability by 24.8% of being underconfident in their advanced financial literacy, all else equal and significant at the 1% significance level. In contrast, it is shown that being female reduced the probability of having accurate or excessive confidence in one's advanced financial literacy by 3.9% and 20.8%, respectively, all else equal and significant at the 1% significance level. Table 8c shows the results of the Wald test for the significance of the female coefficient. The test statistic of 12.75 is significant at the 1% significance level, with a corresponding p-value of 0.0002. Therefore, the null hypothesis of no gender effect on advanced financial literacy confidence is rejected at the 1% significance level.

Additionally, the age coefficient is significant at level 5, with a z-score of -2.05 it is significant at the 5% significance level. The coefficient and z-score suggest that being in age group 50 years old and above on average decreases advanced financial literacy confidence compared to age group 18 to 24 years old.

Figure 5 shows the distribution of the dependent variable advanced confidence plotted against a normal distribution curve. Since the distribution of residuals does not follow that of the normal distribution, the assumption of normally distributed residuals is rejected, and robust standard errors are added to the regression.

6 Discussion

The abovementioned results from the statistical analysis align with our expectations stated in section 2. Each hypothesis has relevant ties to the existing research described in this paper, and based on the drawn conclusions the leading research question can be answered.

The first hypothesis in analysis directly tackles the leading research question, testing the effect of gender immediately on stock market participation. The rejection of this hypothesis takes a big step towards answering the research question, although it does not examine the underlying mechanisms. The finding that gender has a negative effect on stock market participation is in line with our expectation based on the previous research on gender and financial behavior (Bannier et al., 2019; Hasler & Lusardi, 2017; Potrich et al., 2018). As such, no research had been conducted directly on the gender effect on stock market participation, although van Rooij et al. (2011), as well as Potrich et al. (2018) briefly touched on the potential effect of gender on multiple financial metrics such as stock market participation, risk perception, and financial literacy.

The aim of the remaining 4 hypotheses was to analyze and understand the effect found in hypothesis 1. Firstly, it was analyzed whether gender has a significant effect on financial literacy, which would explain the significant negative relationship between females and stock market participation, in line with van Rooij et al.'s (2011) finding that financial literacy has a significant effect on investment activity. Since the finding also stated that the significant effect on said activity is particularly strong in the advanced financial literacy levels, the analysis was split to see whether gender influences both basic and advanced financial literacy. Rejecting hypothesis 2.1 of no gender effect on basic financial literacy is in line with our expectation of women having lower basic financial literacy. Coincidentally, rejecting hypothesis 2.2 of no gender effect on advanced financial literacy corresponds with our expectation of a negative relationship between females and advanced financial literacy, which was also driven by our observation in Table 2 of no women stating they have very high financial literacy. This is in line with the findings of Hasler and Lusardi (2017) and Potrich et al. (2018) of women having significantly lower financial literacy levels. It is important to note that, though both gender effects on basic and advanced financial literacy are significant, the latter is more significant at the 5% level compared to the former, which is only significant at the 10% level. This suggests that the negative effect of gender on financial literacy is more pronounced in advanced financial literacy questions, which is in line with the findings of van Rooij et al. (2011) that the most determining factor in determining stock market participation was the lack of advanced financial literacy, rather than that of basic financial literacy. Seeing that females have a more negative relationship with advanced financial literacy therefore explains the highly significant negative relationship between gender and stock market participation.

Secondly, the effect of gender on investor confidence was analyzed, following the findings by Zhu et al. (2021) and Yang and Li (2016) that females exhibit significantly higher levels of financial under confidence and risk aversity, compared to males. The construction of these two hypotheses assumes that lower confidence is directly related to lower investment activity (Zhu et al, 2021). Again, due to the finding that basic and advanced financial literacy should be differentiated, two analyses were conducted, on gender effect on basic financial confidence and gender effect on advanced financial confidence. Rejecting hypothesis 3.1 of no gender effect on basic investor confidence at the 1% significance level is in line with our expectation of a negative gender effect on basic investor confidence (Yang and Li, 2016; Zhu et al., xx) and adds a layer to understanding why females are underrepresented in the stock market (Muir, 2024). Similarly, rejecting hypothesis 3.2 of no gender effect on advanced financial confidence at the 1% significance level is in line with our expectation of a negative such effect (Yang and Li, 2016; Zhu et al., 2021). Finding such a strongly negative relationship between gender and investor confidence suggests that the gender investment gap is not only rooted in lacking financial literacy amongst women, but also in a significantly lower level of confidence amongst those who do have sufficient financial literacy.

Therefore, the answer to the leading research question is yes, gender does have a significant negative effect on stock market participation, and this phenomenon is driven by women having lower financial literacy levels (Bannier et al, 2019; Hasler and Lusardi, 2017; Potrich et al., 2018; van Rooij et al, 2011), and by their reduced levels of confidence in comparison to their male counterparts (Yang and Li, 2016; Zhu et al., 2021).

It is important to note that the results found in this paper might suffer from certain biases and limitations and could therefore not be representative for the population. Firstly, due to the use of a self-run survey, the sample size of 196 observation suffers from multiple biases, seeing as it was shared mainly to friends and family, which suggests most respondents have approximately the same background as me. For example, more than 50% of the population have a Bachelor's degree, multiple age and education groups have no observations, and most respondents has received a sort of economics of business education. Particularly this last factor might bias the financial literacy regression results. Additionally, due to privacy reasons multiple respondents selected the "prefer not to say" option for the income category, potentially skewing the distribution of income in this sample and therefore invalidating the effect of the control variable. Regarding the data collection, the variable for actual basic financial literacy does not have a level 4 "average", which resulted in the data analysis lacking in some areas and the distribution being irregular. This problem arose due to the van Rooij et al. (2011) financial literacy scales having 7 layers, but there only being 5 questions in the basic financial literacy assessment. Since it was chosen to follow the van Rooij et

al. (2011) survey design to avoid different results due to measurement differences, this problem was not solved.

Nevertheless, the conclusions drawn above are of help to economists and policy makers for two reasons. Firstly, seeing a significant relationship both between gender and financial literacy and gender and investor confidence can be used by policymakers to tackle the problem at the root and invest more in making financial education widespread across all social classes and gender, perhaps through mandatory classes at the high school level. Secondly, having identified a significant effect of gender on stock market participation adds a layer to the existing research on financial gender studies, and provides economists with an insight into the mechanisms of gender and financial behavior. Identifying the different effect of gender on basic versus advanced financial literacy, and the strong effect of gender on investor confidence adds an insight to economists and policymakers which will in the future help them to understand and reduce the phenomenon that is the gender investment.

7 Conclusion

In this thesis I have looked at the gender effect on investment activity, specifically, on the negative effect of gender on stock market participation. Previous research had found that financial literacy has a significant effect on stock market participation, that females have on average lower levels of basic and advanced financial literacy compared to men, as well as reduced confidence levels in their financial literacy levels. However, no ties had been studied between gender and stock market participation. Therefore, this thesis aims to connect the findings of financial literacy effect on investment activities and the findings on the gap in stock market participation. The leading research question in this paper was "Gender Investment Gap: does gender have a negative effect on stock market participation?".

To answer this research question a survey was constructed, consisting of a basic and advanced financial literacy assessment inspired by van Rooij et al.'s (2011) survey. The survey had 196 responses and the participants were asked to state their self-perceived level of basic and advanced financial literacy to assess their self-confidence. Five Probit and Ordered Probit regressions were run with the aim of testing the gender effect on stock market participation, financial literacy, and investor confidence. All regressions showed a significant negative gender effect on the respective dependent variables.

This study therefore concludes that females do exhibits lower levels of stock market participation compared to males, which is driven by their lower levels of basic and advanced financial literacy, and particularly by their highly significant lower confidence. Future research should test the models in this paper on a bigger, more representative sample for the population. Additionally, a more extensive assessment of basic and advanced financial literacy might give more precise results and counteract the issue encountered due to the lacking amount of basic financial literacy questions. Furthermore, it would be interesting to test the causality pattern of the three observed relationships: females and stock market participation, females and financial literacy, and females and confidence. It is not clear in this thesis whether females are less confident because of their lower levels of financial literacy, or whether their literacy is lower because they do not even try due to lower confidence.

References:

Bannier, C. E., Meyll, T., Röder, F., & Walter, A. (2019). The gender gap in 'Bitcoin literacy'. *Journal Of Behavioural And Experimental Finance (Print)*, *22*, 129–134. https://doi.org/10.1016/j.jbef.2019.02.008

CBS. (2023, May 3). *Inkomen - De arbeidsmarkt in cijfers 2022*. Inkomen - De Arbeidsmarkt in Cijfers 2022 | CBS. https://longreads.cbs.nl/dearbeidsmarktincijfers-2022/inkomen/#:~:text=Terwijl%20het%20gemiddeld%20persoonlijk%20inkomen,1%20duizend%20euro%20beduidend%20lager.

Hasler, A., & Lusardi, A. (2017). The gender Gap in Financial Literacy: A Global perspective. In *Global Financial Literacy Excellence Center*. The George Washington University School of Business. https://gflec.org/wp-content/uploads/2017/07/The-Gender-Gap-in-Financial-Literacy-A-Global-Perspective-Report.pdf

Lanciano, E., Previati, D. A., Ricci, O., & Lopes, F. S. S. (2024). Gender differences and measurement error in financial literacy. *Journal Of Behavioral And Experimental Finance*, *41*, 100896. https://doi.org/10.1016/j.jbef.2024.100896

Liang, P., & Guo, S. (2015). Social interaction, Internet access and stock market participation—An empirical study in China. *Journal Of Comparative Economics (Print)*, *43*(4), 883–901. https://doi.org/10.1016/j.jce.2015.02.003

Muir, M. (2024). Women's lack of confidence widens gender investment gap, study finds. *Financial Times*. https://www.ft.com/content/120ba6b9-b8f5-4000-a618-07c6f8f7d845

O'Keefe, T. (2013). Feminist Identity Development and Activism in Revolutionary Movements. In *Palgrave Macmillan UK eBooks*. https://doi.org/10.1057/9781137314741

Potrich, A. C. G., Vieira, K. M., & Kirch, G. (2018). How well do women do when it comes to financial literacy? Proposition of an indicator and analysis of gender differences. *Journal*

Of Behavioural And Experimental Finance (Print), *17*, 28–41. https://doi.org/10.1016/j.jbef.2017.12.005

Ubide, A. (2024). Risks grow for investors as a new economic framework emerges. *Financial Times*. https://www.ft.com/content/30bfc73b-55aa-4a60-9440-11b22428e927

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

Yang, X., & Li, Z. (2016). Ambiguity vs risk: An experimental study of overconfidence, gender and trading activity. *Journal Of Behavioural And Experimental Finance (Print)*, *9*, 125–131. https://doi.org/10.1016/j.jbef.2016.01.003

Zhu, D., Hodgkinson, L., & Wang, Q. (2021). Interaction and decomposition of gender difference in financial risk perception. *Journal Of Behavioural And Experimental Finance (Print)*, *30*, 100464. https://doi.org/10.1016/j.jbef.2021.100464

Appendix A: Tables and Figures

Table 1: Descriptive Statistics Stockmarket, female, basic and advanced financial literacy, basic and advanced confidence, age, education, income, and major.

Name	Obs.	Mean	Std. Dev.	Min.	Max.	25%	50%	75%
stockmarket	165	.509	.501	0	1	0	1	1
Basic FL	165	5.455	1.698	1	7	5	6	7
Advanced FL	165	4.509	1.823	1	7	3	5	6
Basic confidence	165	1.539	.711	1	3	1	1	2
Advanced	165	1.764	.855	1	3	1	1	3
female	165	.509	.501	0	1	0	1	1
age	165	1.770	1.492	1	5	1	1	5
Income	165	1.721	1.016	1	4	1	1	2
education	165	3.358	.634	2	5	3	3	4
Major	165	.558	.498	0	1	0	1	1

Table 2: pre-test self-assess financial literacy and gender.

Financial literacy	male	female	Total
Very low	2	7	9
low	13	21	34
moderate	1	17	18
average	10	17	27
advanced	20	9	29
high	20	13	33
Very high	15	0	15
Total	81	84	165

Table 3: Frequency distribution of gender and stock market participation.

Stock market	male	female	total
0	26	55	81
1	55	29	84
total	81	84	165

Table 4a: Probit regression Gender effect on stock market participation.

Number of obs = 165

Wald chi2(11) = 33.33

Prob > chi2 = 0.0005

Pseudo R2 = 0.1455

Log pseudolikelihood = -97.707

Stock market	coefficient	Robust SE	z	P>|z|	95% confidence interval	
Female	-.822**	.223	-3.68	0.000	-1.259	-.384
Age Group						
25 - 29	.279	.404	0.69	0.490	-.513	1.072
40 - 49	-.586	.962	-0.61	0.542	-2.472	1.299
50+	.051	.526	0.12	0.904	-.783	.886
Education						
Bachelor's	.339	.598	0.57	0.571	-.833	1.512
Master's /	.510	.627	0.81	0.416	-.720	1.739
PhD+	1.813**	.815	2.23	0.026	.216	3.410
Major	.501**	.227	2.21	.027	.056	.946
Income						
~ Dutch	.209	.317	0.66	0.510	-.412	.829
> Dutch	.041	.407	0.10	0.919	-.756	.838
Undisclosed	-.473	.372	-1.27	0.204	-1.203	.257
_cons	-.271	.561	-0.48	0.629	-1.369	.828

Note: The asterisks indicate the significance of the coefficient, with * p<0.10, ** p<0.05, *** p<0.01. Gender effect depiction with age, education level, income, and major as control variables

Table 4b: Marginal Effect of gender on Stock Market Participation.

	dy/dx	Delta SE	z	P>z	95% conf. interval	
female	-.277	.066	-4.22	0.000	-.405	-.148

Note: dy/dx for factor levels is the discrete change from the base level.

Table 4c: Walt test gender effect on stock market participation

[stockmarket] female = 0	
Chi2(1)	13.57
Prob > Chi2	0.0002

Table 5a: Ordered Probit Gender effect on basic financial literacy

Number of obs = 165

Wald chi2(6) = 81.76

Prob > chi2 = 0.0000

Pseudo R2 = 0.2209

Log pseudolikelihood = -19620074

Basic FL	coefficient	Robust SE	z	P>|z|	95% confidence interval	
Female	-.309*	.184	-1.68	0.093	-.670	.052
Age Group						
25 - 29	-.595***	.354	-1.68	0.092	-1.288	.098
40 - 49	-.141	.590	-0.24	0.812	-1.300	1.015
50+	.694**	.334	2.08	0.038	.040	1.349
Education						
Bachelor's	.274	.428	0.64	0.523	-.566	1.113
Master's/MBA	.671*	.453	1.48	0.138	-.216	1.558
PhD+	1.124**	.679	1.95	0.051	-.006	2.654
Major	1.178***	.199	5.91	0.000	.787	1.569
Income						
~ Dutch	-.232	.271	-0.86	0.391	-.764	.299
> Dutch	.187	.305	0.61	0.539	-.410	.785
Undisclosed	-.261	.319	-0.82	0.414	-.887	.365
/cut1	-1.682***	.474			-2.612	-.752
/cut2	-.556	.410			-1.360	.248
/cut3	-.169	.407			-.966	.627
/cut4	.665*	.410			-.137	1.468
/cut5	1.493***	.419			.671	2.314

Note: The asterisks indicate the significance of the coefficient, with * p<0.10, ** p<0.05, *** p<0.01. Gender effect depiction with age, education level, income, and major as control variables

Table 5b: marginal effect Gender actual basic financial literacy

		dy/dx	Delta SE	z	P>z	95% conf. interval	
female							
	1	.011	.009	1.21	0.228	-.007	.028
	2	.037	.022	1.70	0.090	-.006	.079
	3	.017	.012	1.48	0.139	-.006	.040
	5	.030	.019	1.59	0.112	-.007	.066
	6	-.003	.005	-0.53	0.595	-.013	.008
	7	-.091	.054	-1.68	0.093	-.198	.015

Note: dy/dx for factor levels is the discrete change from the base level.

Table 5c: Wald Test gender effect on basic financial literacy

[actual_basic]female = 0	
Chi2(1)	2.84
Prob > Chi2	0.0922

Table 6a: Ordered Probit regression of gender effect on advanced financial literacy

Number of obs = 165

Wald chi2(11) = 84.71

Prob > chi2 = 0.0000

Log pseudolikelihood = -265.578

Pseudo R2 = 0.1240

Advanced FL	coefficient	Robust SE	z	P>|z|	95% confidence interval	
Female	-.461**	.186	-2.47	0.013	-.826	-.096
Age Group						
25 - 29	-.444	.351	-1.27	0.206	-1.131	.244
40 - 49	-.113	.573	-0.20	0.844	-1.235	1.010
50+	.553**	.266	2.08	0.037	.032	1.073
Education						
Bachelor's	.893**	.439	2.03	0.042	.031	1.754
Master's/MB	1.190***	.461	2.58	0.010	.287	2.092
PhD+	1.170**	.521	2.24	0.025	.148	2.191

Major	1.280***	.192	6.68	0.000	.904	1.655
Income						
~ Dutch	-.220	.251	-0.88	0.381	-.711	.272
> Dutch	.257	.277	0.93	0.354	-.286	.800
Undisclosed	-.093	.261	-0.36	0.722	-.603	.418
/cut1	-1.162	.416			-.977	.653
/cut2	.191	.416			-.624	1.006
/cut3	.781*	.418			-.039	1.601
/cut4	1.374***	.420			.551	2.198
/cut5	1.845***	.424			1.014	2.677
/cut6	3.054***	.445			2.183	3.925

Note: The asterisks indicate the significance of the coefficient, with * p<0.10, ** p<0.05, *** p<0.01. Gender effect depiction with age, education level, income, and major as control variables

Table 6b: Average marginal effects of gender on advanced financial literacy

		dy/dx	Delta SE	z	P>z	95% conf. interval	
female							
	1	.062	.027	2.32	0.020	.010	.115
	2	.021	.011	1.97	0.049	.000	.043
	3	.034	.015	2.25	0.024	.004	.064
	4	.021	.010	2.19	0.028	.002	.041
	5	-.000	.004	-0.02	0.981	-.008	.008
	6	-.067	.028	-2.37	0.018	-.123	-.012
	7	-.072	.031	-2.33	0.020	-.133	-.012

Note: dy/dx for factor levels is the discrete change from the base level.

Table 6c: Wald Test on Gender effect on advanced financial literacy

[actual_advanced]female = 0	
Chi2(1)	6.12
Prob > Chi2	0.0134

Table 7a: Ordered Probit regression of gender effect on basic investor confidence

Number of obs = 165

Wald chi2(11) = 829.03

Prob > chi2 = 0.0000

Log pseudolikelihood = -135.910

Pseudo R2 = 0.1166

basic confidence	coefficient	Robust SE	z	P>|z|	95% confidence interval	
Female	-.692**	.211	-3.29	0.001	-1.105	-.280
Age Group						
25 - 29	1.059***	.371	2.85	0.004	.332	1.786
40 - 49	-4.324***	.352	-12.30	0.000	-5.013	-3.635
50+	-.468	.372	-1.26	0.209	-1.197	.261
Education						
Bachelor's	.117	.461	0.25	0.800	-.787	1.021
Master's/MB	-.182	.502	-0.36	0.718	-1.166	.803
PhD+	-.080	.711	-0.11	0.910	-1.473	1.314
Major	.063	.202	0.31	0.755	-.332	.458
Income						
~ Dutch	.385	.296	1.30	0.194	-.196	.966
> Dutch	.021	.328	0.06	0.949	-.621	.663
Undisclosed	-.055	.357	-0.15	0.877	-.755	.645
/cut1	-.030	.433			-.878	.818
/cut2	1.040**	.440			.179	1.902

Note: The asterisks indicate the significance of the coefficient, with * p<0.10, ** p<0.05, *** p<0.01. Gender effect depiction with age, education level, income, and major as control variables

Table 7b: Marginal effects of gender of basic investor confidence.

		dy/dx	Delta SE	z	P>z	[95% conf. interval]	
female							
	1	.231	063	3.65	0.000	.107	.355
	2	-.109	.033	-3.32	0.001	-.173	-.045
	3	-.122	.038	-3.22	0.001	-.197	-.048

Note: dy/dx for factor levels is the discrete change from the base level.

Table 7c: Wald Test for significance coefficient

[basic_confidence] female = 0	
Chi2(1)	10.82
Prob > Chi2	0.001

Table 8a: Ordered Probit regression for the gender effect on advanced investor confidence

Number of obs = 165

Wald chi2(11) = 27.17

Prob > chi2 = 0.0043

Log pseudolikelihood = -159.322

Pseudo R2 = 0.0804

Advanced confidence	coefficient	Robust SE	z	P>|z|	95% confidence interval	
Female	-.702***	.196	-3.57	0.000	-1.087	-.317
Age Group						
25 - 29	.558	.468	1.19	0.233	-.359	1.475
40 - 49	-.314	.842	-0.37	0.709	-1.964	1.336
50+	-.806**	.394	-2.05	0.041	-1.578	-.035
Education						
Bachelor's	.349	.510	0.69	0.493	-.650	1.348
Master's/MB	.192	.536	0.36	0.719	-.857	1.242
PhD+	.686	.702	0.98	0.329	-.691	2.063
Major	-.036	.199	-0.18	0.856	-.427	.354

Income						
~ Dutch	.303	.292	1.04	0.298	-.268	.875
> Dutch	.114	.324	0.35	0.725	-.521	.749
Undisclosed	-.185	.282	-0.66	0.512	-.738	.368
/cut1	-.108	.493			-1.075	.859
/cut2	.544	.496			-.429	1.517

Note: The asterisks indicate the significance of the coefficient, with * p<0.10, ** p<0.05, *** p<0.01. Gender effect depiction with age, education level, income, and major as control variables

Table 8b: Marginal effects of gender on advanced investor confidence

	dy/dx	Delta SE	z	P>z	95% conf. interval	
Female						
1	.248	.063	3.95	0.000	.125	.371
2	-.039	.012	-3.18	0.001	-.063	.015
3	-.208	.058	-3.60	0.000	-.322	-.095

Note: dy/dx for factor levels is the discrete change from the base level.

Table 9c: Wald Test for significance of female coefficient of gender effect

[advanced_confidence] female = 0	
Chi2(1)	12.75
Prob > Chi2	0.0004

Appendix B: Normality of Residuals graphs

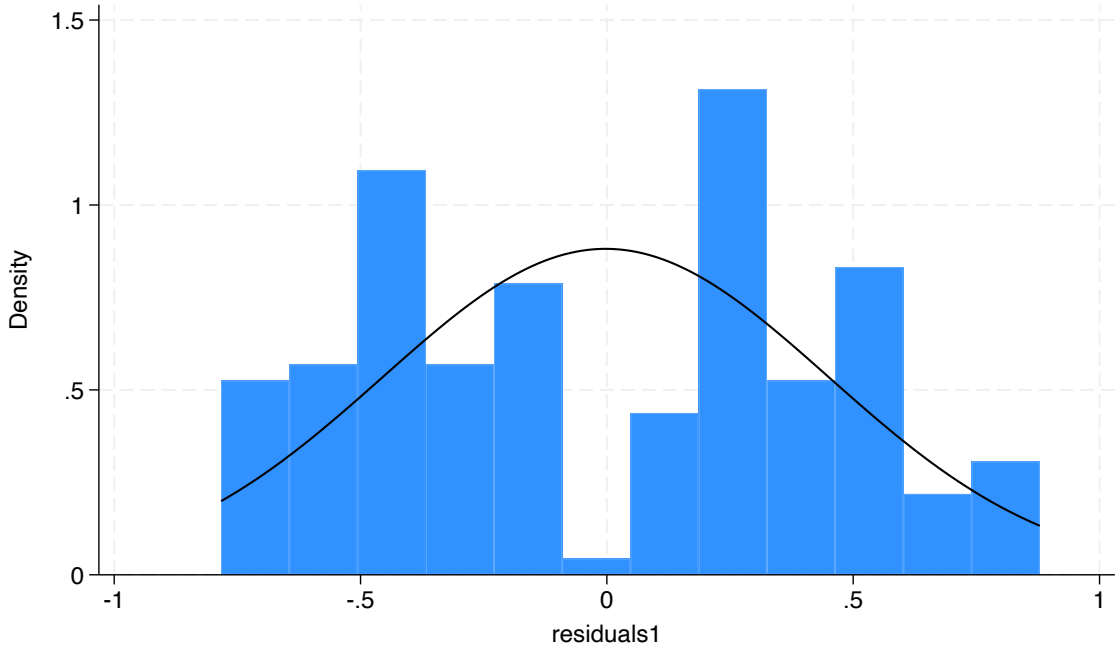

Figure 1: normality of residuals Probit Regression 1, gender effect on stock market participation, against a normal distribution curve.

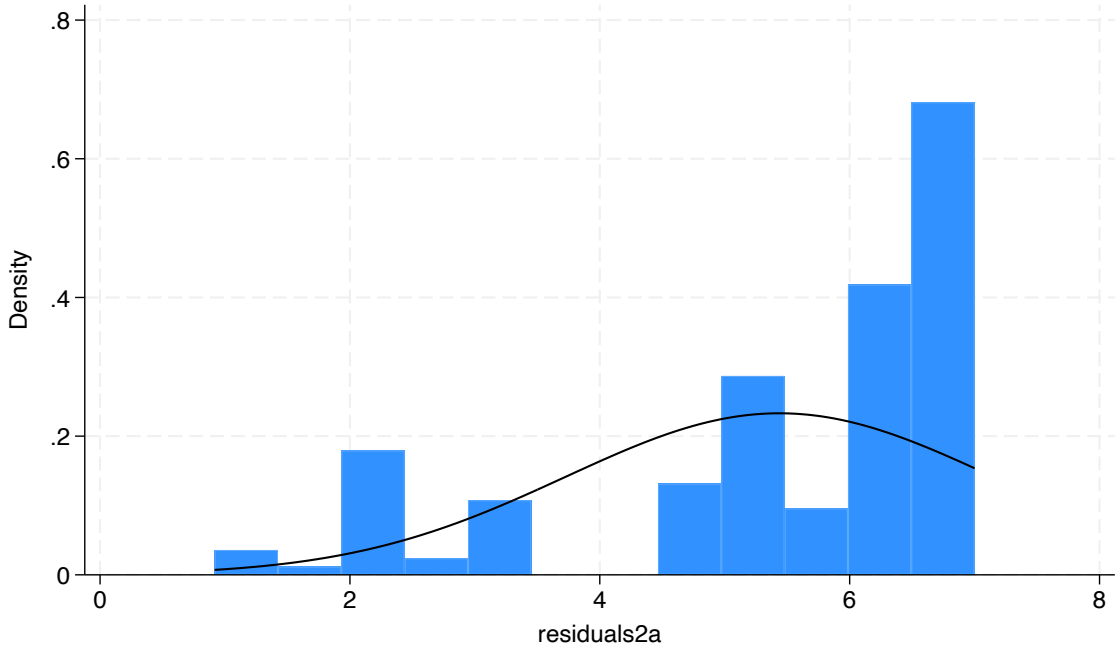

Figure 2: residuals distribution from Ordered Probit regression 2a, gender on basic financial literacy, against a normal distribution curve.

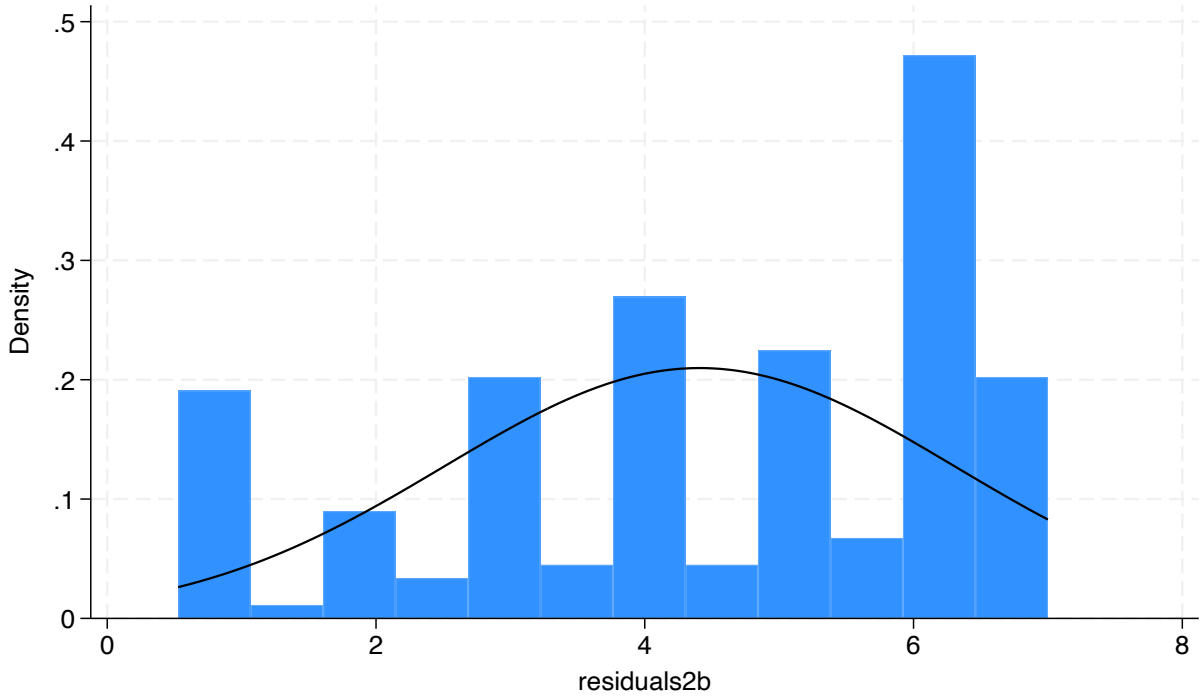

Figure 3: distribution of residuals of Ordered Probit regression 2b, gender on advanced financial literacy, against a normal distribution curve.

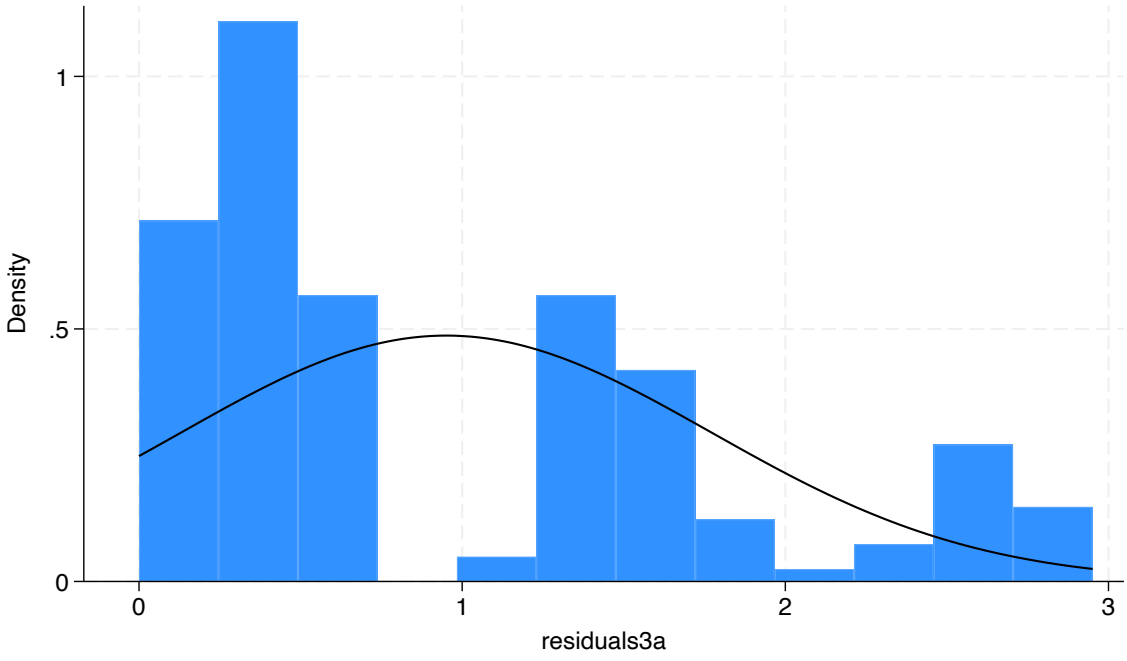

Figure 4: distribution of residuals of Ordered Probit Regression 3a, gender on basic investor confidence, against a normal distribution curve

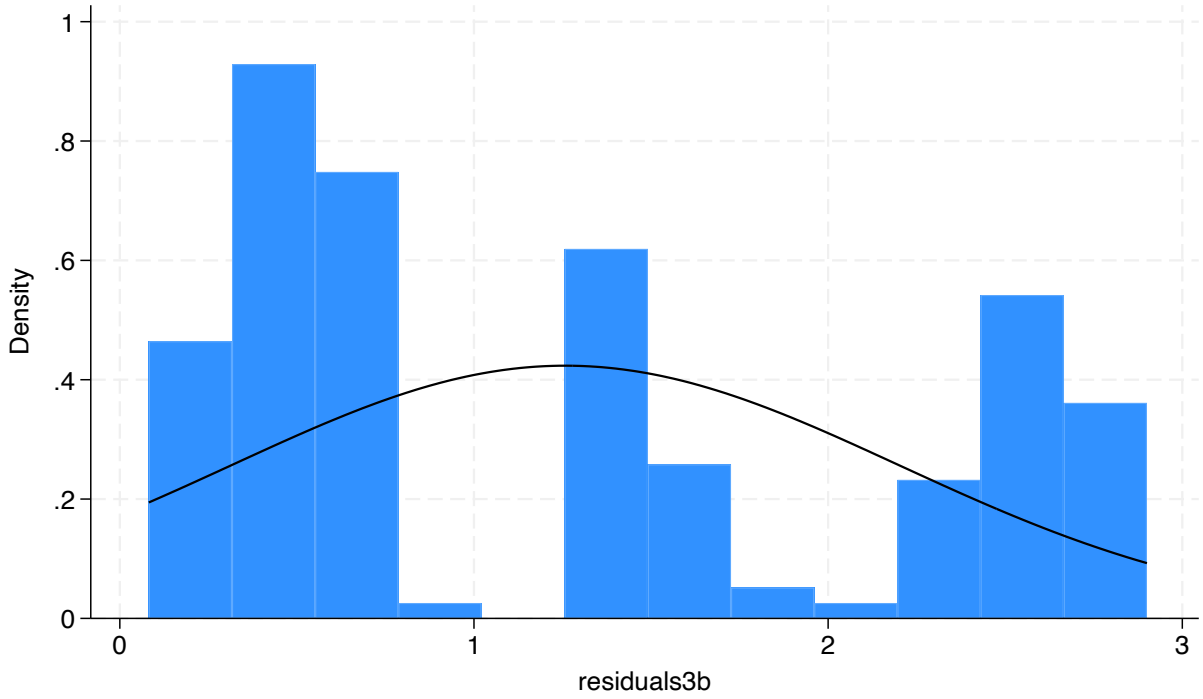

Figure 5: distribution of residuals of Ordered Probit regression 3b, gender on advanced investor confidence levels, against a normal distribution curve.

Appendix C: Financial Literacy Assessment Questions

Basic Literacy Questions.

- *Numeracy:* Suppose you had €100 in a savings account with a 2% yearly interest rate. How much would you have in the account after 5 years if you left the money there to grow?
    i. More than €102 (correct)
    ii. Exactly €102
    iii. Less than €102
    iv. Do not know
- *Interest compounding:* Suppose you had €100 in a savings account with a 20% yearly interest rate. How much would you have in this account after 5 years, assuming you never make any withdrawals or interest payments?
    i. More than €200 (correct)
    ii. Exactly €200
    iii. Less than €200
    iv. Do not know
- *Inflation:* Suppose the yearly interest rate on your savings account is 1% and the yearly inflation rate is 2%. Compared to today, how much would you be able to buy with this money after 1 year?
    i. More than today
    ii. Exactly the same
    iii. Less than today (correct)
    iv. Do not know
- *Time value of money:* Suppose your sibling inherits €10,000 today and your cousin inherits €10,000 3 years from now. Who is richer because of the inheritance?
    i. Your sibling (correct)
    ii. Your cousin
    iii. They are equally rich
    iv. Do not know
- *Money illusion:* Suppose that in the year 2030 both your income and prices of all goods have doubled. Compared to today, how much are you able to buy with your income in 2030?
    i. More than today
    ii. The same (correct)
    iii. Less than today
    iv. Do not know

Advanced Literacy Questions

- Which answer most accurately describes the main function of the stock market?
    i. Predicting stock earnings
    ii. The stock market is the result of an increase in stock prices
    iii. Bringing together people who want to buy and sell stocks (correct)
    iv. None of the above
    v. Do not know

- Which of the following statements about stocks is correct? Buying the stock of firm B in the stock market means:
    i. I now own a part of firm B (correct)
    ii. I have lent money to firm B
    iii. I am liable for firm B's debt

iv. None of the above
v. Do not know

- Which of the following statements about mutual funds is correct?
  i. Once I invest in a mutual fund, I cannot withdraw money in the first year
  ii. Mutual funds can invest in several assets (e.g. both stocks and bonds) (correct)
  iii. Mutual funds pay a guaranteed rate of return, which depends on their historical performance
  iv. None of the above
  v. Do not know

- Which of the following statements about bonds is correct? Buying a bond of firm B:
  i. I now own a part of firm B
  ii. I have lent money to firm B (correct)
  iii. I am liable for firm B's debt
  iv. None of the above
  v. Do not know

- In the long run (imagine 10 or 20 years), which asset yields the highest return?
  i. Savings accounts
  ii. Bonds
  iii. Stocks (correct)
  iv. Do not know

- Which asset typically displays the highest fluctuations over time?
  i. Savings accounts
  ii. Bonds
  iii. Stocks (correct)
  iv. Do not know
- By spreading money among multiple different assets, the risk of losing money:
  i. Increases
  ii. Decreases (correct)
  iii. Stays the same
  iv. Do not know
- (True or False?) When you buy a 10-year bond, it means you cannot sell it after 5 years without incurring a penalty.
  i. True
  ii. False (correct)
  iii. Do not know
- (True or False?) Stocks are typically riskier than bonds.
  i. True (correct)
  ii. False
  iii. Do not know
- (True or False?) Buying a company stock typically yields a safer return than a stock mutual fund.
  i. True
  ii. False (correct)
  iii. Do not know
- When the interest rate falls, what should happen to bond prices?
  i. Rise (correct)
  ii. Fall
  iii. Stay the same

iv. None of the above
v. Do not know

Self-reported fin. Literacy questions:

- Self assessed basic financial literacy 1 (very low), 2(low), 3 (moderate), 4 (average), 5 (advanced), 6 (high), 7(very high)
- Self assessed advanced financial literacy 1 (very low), 2(low), 3 (moderate), 4 (average), 5 (advanced), 6 (high), 7(very high)