CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of economic studies

Reaction of retail investors to financial market movements and sentiment changes

Bachelor's thesis

Author: Jakub Hromčík

Study program: Economics and finance

Supervisor: PhDr. Jiří Schwarz, PhD.

Year of defense: 2024

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, January 3, 2024

Jakub Hromčík

Abstract

This bachelor thesis investigates two areas. First, we study the impact of sociodemographic attributes on retail ivnestors following robo-advice in the choices of ready-made portfolios of passive ETFs with unique risk levels by employing a logistic regression model. Second, we investigate the impact of sociodemographic attributes on retail investors' trading volume adjustments in periods of high expected market volatility as proxied by the VIX index, for which we employ panel data regression methods over 18 consecutive months during a relatively stable period from January 1st 2021 to the end of 2022. We find, in agreement with reasearch on human financial advice, that women are more likely than men to follow risk level recommended by a robo-advisor, while being a man is associated with assuming more risk than recommended. Due to model assumption issues, our results are rather inconclusive in whether men tend to react differently to periods of high expected market volatility.

JEL Classification	D90, D91, G40, G41, J16
Keywords	ETFs, VIX, Robo-advisor, Ready-made portfolio

Title	Reaction of retail investors to financial market movements and sentiment changes

Author's e-mail	`kubahromcik@gmail.com`
Supervisor's e-mail	`jiri.schwarz@fsv.cuni.cz`

Abstrakt

Tato bakalářská práce se zabývá dvěma oblastmi. Nejprve pomocí logistické regrese studujeme vliv sociodemografických veličin na následování rad finančních algoritmů při výběru hotových portfólií pasivních fondů obchodovaných na burzách, takzvaných ETFs, které s sebou nesou různé množství rizika. Dále pomocí panelové regrese zkoumáme vlivy sociodemografických veličin na změny v investičních tocích v obdobích zvýšené očekávané tržní volatility, pro kterou využíváme index VIX. Panelová data zkoumáme v relativně stabilním období od 1. června 2021 do konce roku 2022, tedy celkem 18 měsíců je zahrnuto v analýze. Souhlasně s existující literaturou zjišťujeme, že ženy mají větší pravděpodobnost vybrat si portfolio s menším nebo stejným rizikem, než jim bylo doporučeno algoritmem, jež z dostupných sociodemografických informací vybere portfolio s nejvhodnější mírou rizika. Z důvodů nesplnění některých předpokladů užitých modelů, nejsme schopni vyloučit nulovou hypotézu neexistence rozdílu v investičních reakcích mužů a žen na období očekávané vysoké volatility.

Klasifikace JEL	D90, D91, G40, G41, J16
Klíčová slova	ETFs, VIX, Robo-advisor, Ready-made portfolio
Název práce	Reakce retailových investorů na pohyby finančních trhů a změny sentimentu

E-mail autora	kubahromcik@gmail.com
E-mail vedoucího práce	jiri.schwarz@fsv.cuni.cz

Acknowledgments

The author is grateful to family and to PhDr. Jiří Schwarz, PhD..

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record
Hromcik, Jakub: *Reaction of retail investors to financial market movements and sentiment changes* . Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of economic studies, Prague. 2024, pages 67. Advisor: PhDr. Jiří Schwarz, PhD.

Contents

List of Tables

List of Figures

Acronyms

ETFs Exchange-Traded Funds

VIX Volatility Index

CBOE Chicago Board Options Exchange

ESG Environmental, social and corporate governance

Chapter 1

Introduction

With easier than ever access to technology and the internet, a great rise in retail investment has been taking place in the last decade. More than that, the Covid-19 pandemic has acted as a strong catalyst in the process, catapulting retail investment to an all time high (Deloitte 2021). Becoming competitive to its institutional counterpart in various financial metrics such as trade volume and assets under management (Boston Consulting Group 2021), it is important to study the determinants that drive the behavior of the everyday investor.

Nowadays, retail investors have many options in both the investment products they can choose from and the channels they can invest through. This simple fact leads to an interesting optimization puzzle, which has been troubling researchers for decades ever since Markowitz (1952) fathered Modern Portfolio Theory. Present and future investors, who make investment decisions, face multiple constraints; be it wealth, access to information or time. Research, which investigates decision making under uncertain outcomes, gave rise to the field of behavioral economics, where the idea of heuristics, thought processes that ought to overcome these constraints, takes central place.

Notable examples of specific heuristics employed by individuals include the anchoring heuristic and availability heuristic. Anchoring heuristic essentially means that, when making a guess or an estimate, individuals tend to start their thinking process at a reference point and then adjust towards a final answer or decision. Availability heuristic helps estimate event probabilities based on easiness of imagination or recall of similar events that happened in individual's life recently or in the past (Kahneman *et al.* 1982).

However, heuristics do not always lead to optimal outcomes, several behavioral biases have been identified, that may systematically divert individual investors from reaching their best (not only) financial behavior. Take the two mentioned heuristics. Sales oriented businesses may take advantage of anchoring heuristic by offering multiple size variants of the same product and strategically price them in a way that steers customers towards a specific option. Or, in the case of availability heuristic, one may consider himself more competent in a hazardous endeavour after experiencing positive outcomes despite the underlying odds of success being totally random.

Most interestingly, biases and their extent are not influencing all retail investors equally, demographic attributes of each person may play a role. For example, sex may influence one's (mis)judgment of their trading ability, leading to overconfidence, as demonstrated by Barber & Odean (2001). Another important variable which influences retail investors is their age. People of different ages may have different investment goals, risk preferences, investment skills and susceptibility to various biases. Indeed, Korniotis & Kumar (2011) show that older people tend to prefer more diversified portfolios, trade less and could benefit more from holding passive index funds. More abstract attributes, such as investor sophistication, proxied by (for example) income, are also important in explaining decision making and inclination to biases (Dhar & Zhu 2006).

While the already published studies are extremely relevant to the study of sociodemographic attributes' impact on investment decisions, most papers investigating the issues at hand do so by focusing mainly on the most common financial instruments such as stocks, bonds or retirement plans. Literature investigating retail investor behavior involving historically more recent instruments such as Exchange-Traded Funds (ETFs) or even ready-made portfolios of ETFs, is scarce. A few examples include Bhattacharya *et al.* (2017) who investigate how holding ETFs improves retail investors' portfolio performance and give a general description of ETFs holders as being younger and wealthier compared to non-holders. D'Hondt *et al.* (2023) study the probability and magnitude of trading passive ETFs based on sociodemographic attributes and various bias proxies, they find that investors who trade passive ETFs are generally more sophisticated than those who do not, more likely to have higher education and longer investing horizons and be less prone to overconfidence and home bias (the tendency to hold portfolios of predominantly domestic equities).

Furthermore, with the increasing shift of retail investors towards mobile trading, their behavioral biases may express differently. Studies investigating the two factors - biases and technology, have recently begun to emerge, suggesting that investors using trading apps on their mobile phones are more prone to adverse biases, for example see Kalda *et al.* (2021) or Cen (2021). An attempt to foster a healthier relationship between retail investors and financial markets has been made by implementing so called "robo-advisors", computer algorithms that help retail investors make better informed decisions or go as far as automatically diversify and rebalance their portfolios. Disposition effect; a well documented bias that drives retail investors to 'sell winners early and ride losers too long', is an example of a bias that can be to some extent mitigated by robo-advice (Back *et al.* 2023).

Nevertheless, robo-advisors often work similarly to human financial advisors, offering advice that one may or may not act upon at will. The decision of retail investors to abide by a financial advice is itself influenced by sociodemographic attributes. In following human advice, Reiter-Gavish *et al.* (2021) find that wealthier people are more likely to act upon human financial advice, the same holds for older people and women as opposed to men. To the best of our knowledge, no one has yet documented the effect of sociodemographic variables on retail investors following robo-advice in ready-made portfolio choices.

The objective of this thesis is therefore two-fold. First, we will investigate the impact of market sentiment and sociodemographic atributes on retail investment flows. Specifically, we will study how variables such as age, sex, income and others impact the timing and volume of investment done by retail investors towards ready-made portfolios of ETFs, in connection to market sentiment represented by Volatility index (VIX), and how these may be linked to specific biases. Second, we will investigate whether these and other variables have impact on following robo-advice in portfolio choice, namely the amount of risk assumed when selecting a ready-made portfolio after being (robo)recommended a portfolio of a certain risk level.

The structure of the thesis is following: Chapter 2 conducts behavioral finance literature review, Chapter 3 states hypotheses, Chapter 4 covers the data and methodology, Chapter 5 presents results, Chapter 6 contains discussion and chapter 7 concludes.

Chapter 2

Literature review

This chapter provides an overview of literature that has shaped the underlying area of research. It is split into three sections: Section 2.1 provides a brief historical introduction to the field of behavioral economics, section 2.2 gives an overview of research on aggregate patterns of behavior of retail investors towards financial markets and section 2.3 summarizes research on the relationship between individual investor characteristics and behavior.

2.1 Brief history of behavioral economics

During the 20th century, economics as a science has progressed dramatically. Regarding the focus area of this thesis, the most important evolution came in the 1970s, when Kahneman & Tversky (1979) joined forces to produce Prospect theory, stepping aside from the classical economics theories of *homo economicus* and perfect rationality. They have built upon the previous observations of researchers such as Allais (1953), who has shown inconsistency in choices that goes against the independence axiom of Expected utility theory, generating evidence against perfect rationality. More specifically, Allais has shown that choices are not always consistent in situations involving decision making under uncertainty, counter to the assumptions of expected utility theory.

The work of Kahneman and Tversky is often regarded as the foundation which gave birth to the relatively young field of behavioral economics. Notable outcome of their Prospect theory, relevant for any investment behavior analysis including this thesis, is the finding that in situations involving uncertainty, individuals tend to be risk seeking in avoiding losses and risk averse in acquiring gains.

With the rise of the new millennium, behavioral economics and its subset of behavioral finance have provided a stark contrast to the Efficient market hypothesis, suggesting that cognitive biases, sentiment and irrational behavior can lead to systematic deviations of prices from fundamental values, such as the ones the financial world has experienced during the dot-com bubble and the 2008 financial crisis (Shiller 2000, Rizzi 2008).

These irrational influences are often significant and not fully mitigated by rational investors' counter-movements, as may be the claims of opponents of behavioral economics. Furthermore, rational investors might even time the irrational waves to get in ahead of the market and reap bigger rewards, contributing to further drifts of prices away from fundamental values, as suggested by De Long *et al.* (1990). For the safety of the markets as well as of the individual investors, it is important to continue the study of human behavioral patterns in the domain of finance. Summary of contributions essential to this thesis follows.

2.2 Aggregate patterns of behavior

The literature on retail investment is rich in topics of stocks and mutual funds. The investment product that is sought by investors in our data set (see Chapter 4) is a ready-made portfolio of passive ETFs. However, the author finds it reasonable to assume that literature investigating retail investor behavior towards stocks and mutual funds would be, to some extent, also applicable in passive index funds.

Regarding investor behavior, Sirri & Tufano (1998) found a link between past performance of a mutual fund and subsequent flows (investments and divestments of retail investors towards the specific fund). This link is not symmetrical, though, funds with strong performance enjoy disproportionately more inflows than poor performing funds suffer from outflows. Odean (2000) provides further evidence for this relationship and explains it in terms of two opposing forces rooted in behavioral economics. In the process of selecting a mutual fund to be bought, investors are swayed by the *representativeness heuristic*, meaning they overemphasize the importance of past performance in determining future outcomes. In contrast, when making a decision about selling their mutual funds, the *disposition effect* kicks in and results in holding onto low performing funds for longer than optimal.

The *disposition effect* itself plays a significant role in behavioral finance, linking back to the Prospect theory of Kahneman and Tversky, Barberis & Xiong (2009) demonstrate that it is actually *realized* gains and losses that drive the phenomena rather than *annual* gains and losses. This suggests that investors derive utility at the time a given security is sold at a premium or a loss compared to the original buying price, rather than continuously throughout a year.

Grinblatt & Keloharju (2001) show that more sophisticated types of non-retail investors, namely financial institutions, are less moved by past returns of an asset when making a buy-sell decision. One of their findings is also that the less sophisticated retail investors, namely households, are more likely to act as contrarians, a behavior which is to some extent similar to trading under the influence of the *disposition effect*.

On the topic of green funds, Bollen (2006) found that "socially responsible" funds enjoy systematically lower volatility of flows compared to regular funds (inflows that follow a strong performance are greater in SR funds than regular funds).

2.3 Individual characteristics influencing decision making

Barber & Odean (2001) show that sex plays a role in retail investment behavior. Their analysis of stock trading data revealed that men suffer from "overconfidence" more frequently than women do. Overconfidence is a psychological bias that leads one to believe they are more competent in their judgment than may actually be the case. In behavioral finance literature, it is often observed through overtrading; behavior defined by an investor's trade volume being significantly larger than optimal that also goes against profit maximization. Gervais & Odean (2001) analyse overconfidence based on experience and show that retail investors generally tend to be most overconfident at the beginning of their investment journey and, depending on whether they experience success (failure), they become more (less) overconfident.

There is, nevertheless, a much more general pattern at play. A plethora of research has been showing consistently that women are on average more risk averse than men. The best example is a meta-analysis conducted by Byrnes et. al (1998), which also suggested that this difference is not age constant and instead tends to reduce as people get older.

Contributing to the study of the *disposition effect*, Dhar & Zhu (2006) show that there are individual differences in susceptibility to the phenomena among retail investors. They demonstrate that various proxies for investor sophistication such as age, occupation and income, have statistically significant impact on the presence of the *disposition effect* in investment decisions. On average, older and wealthier investors as well as investors from more professional occupations, are less prone to the disposition effect, compared to the members of the opposite groups. Grinblatt & Keloharju (2012) provide more evidence to the influence of investor sophistication by linking IQ to better trading outcomes, less vulnerability to *disposition effect* and more likelihood of engaging in tax-driven selling, i.e. overselling in December.

Agnew *et al.* (2003) analysed the impact of demographic attributes on portfolio allocations and trading activity in 401(k) accounts. Their study shows that men trade more often than women and choose riskier assets. Same tendencies are linked to individuals with higher salary. In their study, age is positively correlated with trading frequency and, perhaps most interestingly for the scope of this thesis, investors are shown to respond to market developments with a one day lag, feedback trading may therefore be present.

Chapter 3

Hypotheses

This chapter states the hypotheses and links them to the underlying area of research.

Hypothesis 1: Women are more likely to follow the recommendations of a robo-advisor when choosing a portfolio.

The literature on whether retail investors actually follow received (human) financial advice is relatively scarce. Nevertheless, Reiter-Gavish *et al.* (2021) have studied the impact of sociodemographic factors on whether one does act on professional financial advice or not. Their results show that women are more likely to comply with financial advice, which is to some extent in line with the observed tendency of women to be more risk averse than men confirmed by Byrnes *et al.* (1998). Studying the drivers of complying with robo-advice is hence an interesting extension to the existing literature by testing whether the observed trends also apply to robo-advice.

Hypothesis 2: There is a significant difference between the sexes in ETFs investment volume adjustments during periods of high market volatility.

VIX index is repeatedly used by researches to proxy for market sentiment (Ding *et al.* 2021). Studies show that market sentiment and volatility as proxied by VIX have significant effect on both following human financial advice and subsequent investment decisions made by retail investors (e.g. Reiter-Gavish *et al.* 2021). Our data set (see Chapter 4) comes from a brokerage which only offers ready-made portfolios of passive ETFs and encourages passive periodic investing. Therefore, we ask whether different sociodemographic groups of investors in our data set differ in their investment responses to high expected market volatility proxied by VIX.

Chapter 4

Data and methodology

In the first part of this chapter, a detailed description of the data will be provided. Section 4.1 gives a general description of the data, section 4.2 describes the variables available, section 4.3 covers summary statistics of the variables and section 4.4 outlines aggregate trading patterns of the investors in the dataset. The second part of the chapter will describe the statistics tools used to analyse the data. Section 4.5 describes the methodology in detail.

4.1 Data set

The data comes from a Czech online based brokerage. The author is bound by a non-disclosure agreement with the company and therefore cannot share the data set. Retail investors, who chose to invest with the company, have filled out a questionnaire regarding socio-demographic information, investment goals, and undergone appropriateness and suitability assessments, as required by MiFID II EU financial regulation. The process of portfolio choice by each investor is the following: after an investor provides information about themselves, they are recommended a specific portfolio (by a robo-advisor) based on their financial knowledge, experience, investment goals and background. They are then able to choose any portfolio they like, but are explicitly warned when choosing a portfolio that involves more risk than recommended. After choosing a portfolio, they are prompted to make their initial investment, by which the set up process is completed and they can continue to invest and divest at will.

The portfolios are on a scale, beginning with very low risk and historically relatively mild performance on one side, and higher associated risk paired up with historically stronger performance of the underlying securities on the other side. We will label these portfolios *I*,*II*,...,*VII*, based on how much risk and potential reward they entail, *I* being the lowest risk and *VII* the highest. In addition, each portfolio has its ESG twin, we will label these *I-ESG*,...,*VII-ESG*. In total, 14 portfolios are offered. The difference in risk is mainly driven by what ratio of the portfolio's value is attributed to bonds ETFs versus stocks based ETFs.

The data itself consists of 3 files, File 1 contains anonymized information about each investor, File 2 consists of dates, volumes and currency codes of deposits of each user and File 3 contains complementary information about their withdrawals. The covered period is between August 2019 and January 2023, therefore it comprises three and a half years, or a total of 42 months worth of investment data.

Two important distinctions separate this data set from most other data sets used by other researchers. The first is that it incorporates information on investors who have only started investing with the brokerage some time during the period the data covers, therefore none of the investors have invested with the company before the start of the period. Consequently, the number of active investors during the period varies, beginning with the very first new investor at the start of the period and concluding with 11,074 investors at the end. Second, based on the product the brokerage offers, investors do not face transaction costs in the form of fees for individual trades (changing portfolios, investing into and divesting from specific portfolios). Instead, they face fixed costs based on the aggregate value of their portfolio(s), irrelevant to the number of trades they make.

The pre-processed version of the data set includes information on 20,767 unique users, 11,074 of whom have made at least one deposit during the mentioned period and the rest have not finished the activation process with the brokerage; they have either filled out some or all of the required information, but have not made an initial or any other subsequent deposits within the time span of the data set.

To clean the data set and for the subsequent analysis, R language has been used together with *tidyverse* packages. Cleaning the data consisted mostly of removing false inputs, for example inputs with '1234567' as monthly income and inputs with completely unrealistic amount of digits. No more than 20 inputs were impacted by removing their values due to mentioned reasons.

4.2 Variables

For each investor, the data set contains the following variables, detailed description of the categorical variables will be provided in the next paragraph. *Postal code* (Czech identification number of a city or a town of their official permanent residence), *Nationality*, *Age*, *Sex*, *Estimated investment length* (the time period the investor expects to invest for at the time of creating their account, in years), *Estimated first investment volume* (the amount they expect to be their first deposit), *Estimated regular investment volume* (the amount they expect to deposit regularly after their first deposit), *Monthly income*, *Monthly expense*, *Investment goal* (the motivation behind the investment), *Source of investment* (the source of money intended for investing with the company), *Main source of income* (the source of their general income), *Occupation*, and finally, the *Recommended portfolio* and *Chosen portfolio* (portfolios recommended by a robo-advisor based on the provided information and the first portfolio the client chose and deposited money into, respectively).

The *Investment goal* variable is a categorical variable of 5 levels: Short term speculation, Medium-term investment, Retirement savings, For-offspring savings and Other long-term goal. The *Source of investment* variable has 5 levels, Employment, Entrepreneurial activity, Pension, Sale of property and Other. On a similar note, *Main source of income* is of 5 levels, Employment, Self-employment, Pension, Real estate rental and Other. Occupation is chosen out of a list of 46 most common areas of occupation such as Real estate, IT, Military, Retail, Healthcare, Agriculture, Financial institutions, Scientific institutions and more.

The data were provided in a form of survey while creating an account with the brokerage. This means that, considering the scale, it will necessarily contain some faulty information entered due to human error or intentional falsehood, this is one of the main limitations of the data set. Moreover, providing some of the information was completely optional, meaning for some investors we may have one or more variables with missing values, apart from *Birth year*, which we have for all 11,074 activated users who have made at least one deposit.

Out of the 11,074 participants, 10,484 have remained active at the end of the time span and 590 have closed their accounts; withdrawing all of their investments. We denominate these groups "Active" and "Churn", respectively. In the analysis, we will study the behavior of all 11,074 investors. When describing summary statistics, we will consider these two groups separately to explicitly show if and how they differ.

4.3 Summary statistics

Table 4.1 displays summary statistics of the numerical variables of Active users. Table 4.2 displays summary statistics of the numerical variables of the Churn group. Table 4.3 shows categorical variables counts for both groups.

Table 4.1: Summary Statistics of numerical variables for the Active group.

ACTIVE	Min	1st Q.	Median	Mean	3rd Q.	Max	NAs
Age	18	27	32	34	39	92	0
Estimated investment length (years)	3	6	10	15.73	20	50	291
Estimated first investment volume	0	1000	10000	45706	50000	4800000	293
Estimated regular investment volume	0	1000	1500	3151	3000	200000	295
Monthly income	0	27500	40000	49991	60000	2000000	303
Monthly expense	0	12000	20000	23494	30000	400000	303

Note: 10,484 observations. All monetary variables are in CZK.

Table 4.2: Summary Statistics of numerical variables for the Churn group.

CHURN	Min	1st Q.	Median	Mean	3rd Q.	Max	NAs
Age	18	26	31	33	38	76	0
Estimated investment length (years)	3	5	10	14.05	20	50	62
Estimated first investment volume	0	1000	5000	44935	20000	1000000	63
Estimated regular investment volume	0	1000	1000	2291	2500	20000	63
Monthly income	0	27000	37000	47150	55000	450000	66
Monthly expense	0	10000	20000	22811	30000	200000	67

Note: 590 observations. All monetary variables are in CZK.

Comparing the Active group with the Churn group, the numerical variables are mostly similar, with the Active group containing more extreme outliers in the right tails of the respective distributions. The ratio of missing values to the total number of members in the Churn group is about three times higher compared to the Active group.

Table 4.3: Summary statistics of categorical variables for Active and Churn groups of retail investors.

	Total	
	Active (10,484)	Churn (590)
Sex		
Female	2,695 (25.7%)	123 (20.9%)
Male	7,753 (74%)	463 (78.4%)
NA	36 (0.3%)	4 (0.7%)
Nationality		
Czech	4,079 (38.9%)	213 (36.1%)
Slovak	20 (0.2%)	1 (0.2%)
Other	10 (0.1%)	0
NA	6,375 (60.8%)	376 (63.7%)
Investment goal		
Short term speculation	0	4 (0.7%)
Medium-term investment	4,381 (41.8%)	223 (37.9%)
Retirement savings	2,667 (25.5%)	126 (21.4%)
For-offspring savings	575 (5.5%)	39 (6.6%)
Other long-term goals	2,560 (24.3%)	131 (22.1%)
NA	301 (2.9%)	67 (11.3%)
Source of investment		
Employment	7,170 (68.4%)	355 (60.2%)
Entrepreneurial activity	2,168 (20.7%)	111 (18.9%)
Pension	90 (0.8%)	3 (0.5%)
Sale of property	133 (1.3%)	9 (1.5%)
Other	610 (5.8%)	39 (6.6%)
NA	313 (3%)	73 (12.3%)
Main source of income		
Employment	6,954 (66.3%)	345 (58.3%)
Self-employment	2,452 (23.4%)	126 (21.5%)
Pension	133 (1.3%)	3 (0.5%)
Real estate rental	73 (0.7%)	3 (0.5%)
Other	561 (5.3%)	40 (6.9%)
NA	311 (3%)	73 (12.3%)
Occupation		
IT	2,008 (19.2%)	79 (13.7%)
Marketing	600 (5.7%)	28 (4.9%)
Healthcare	496 (4.7%)	14 (2.3%)
Retail	375 (3.6%)	25 (4.2%)
Inventory and logistics	291 (2.8%)	35 (5.9%)
Other	6,403 (61%)	339 (56.7%)
NA	311 (3%)	70 (12.3%)

Note: 11,074 observations. Percentage values of the respective groups are given in brackets.

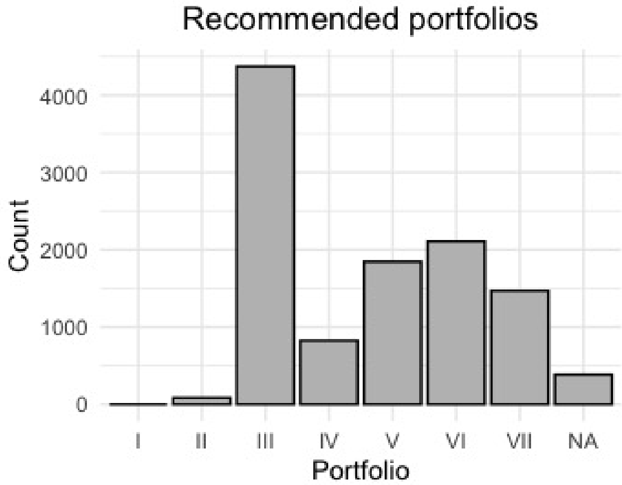
To inspect whether and how individuals in our data set differ from the average population, the following paragraphs will contrast demographic statistics of investors from the aggregated group (Active and Churn) with demographic averages of the Czech republic, which are gathered from official report of the Český statistický úřad (2021). The average age among the investors equaled 34 years, while the median age was 32. This is significantly lower than the national average of 42.7 years. This could be due to the company's platform being solely online, attracting younger generations more. Only about 26% of all investors were female, while 74% were male, this is similar to several data sets of the studies mentioned in Chapter 2, for example, Dhar & Zhu (2006) documented a ratio of 21% to 79% in their data set.

The median monthly income was 40,000 CZK, while the mean was 50,626 CZK per month. These are greater than the average (41,265 CZK) and median (34,741 CZK) gross monthly income of the general Czech population. It should be noted here that the main source of income for the investors in our data set is not strictly employment, which is the baseline of the provided statistic of the general population. If we consider only those employed, the median does not change, while the mean decreases to 46,409 CZK, a number closer to that of the general population, but still quite larger.

66% (7299) of investors have claimed to be employees, 23.3% (2578) self-employed and 1.2% (136) pensioners. Therefore, the ratio of self-employed individuals is higher than in the (working) Czech population (13.6%, against 76.1% employed).

The distributions of portfolio recommendations and selections are influenced by the following technicalities. Each investor is recommended at most one standard portfolio (ESG portfolios were not included in the recommendation process) and they can choose several portfolios at will. Therefore, the sum of all selected portfolios will necessarily be higher than or equal to the sum of recommended portfolios. Out of the 11,074 users, we lack selected portfolio information for 590 (100%) from the Churn group and 94 (1%) from the Active group. The remaining 10,390 users have selected in total 12,375 portfolios. The exact distributions are in Figure 4.1 (Recommended portfolios), Figure 4.2 (Selected standard portfolios) and Figure 4.3 (Selected ESG portfolios).

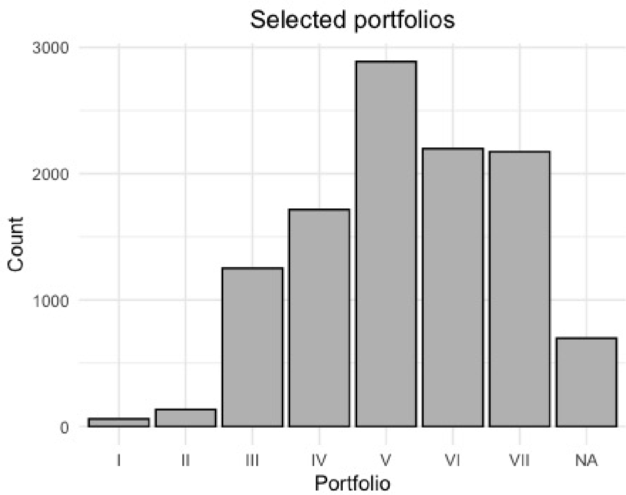
Figure 4.1: Regular portfolios recommended to all investors by a robo-advisor.



Note: 11,074 observations. Portfolios range from *I* (lowest risk and associated possibility of reward), to *VII* (highest implicit risk and historically stronger performance).

In the recommended portfolios distribution, we can immediately see that the robo-advisor has produced two groupings. One is a straight peak at portfolio *III*, and the other is a more normally looking distribution around portfolios *IV-VII*.

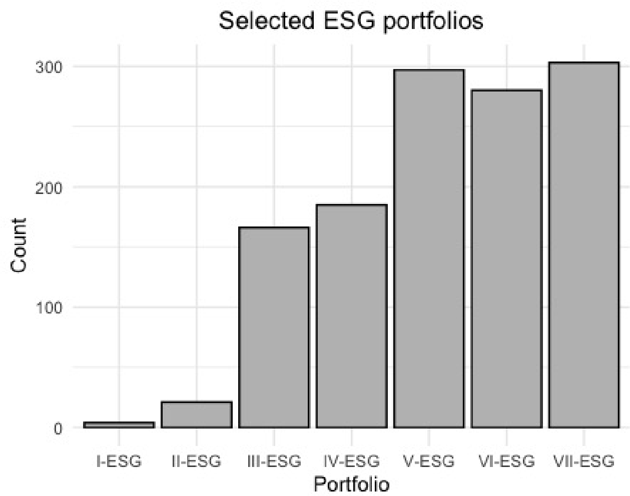
Figure 4.2: Regular portfolios selected by Active group investors.



Note: 10,484 observations, all Churn group investors have missing values. Portfolios range from *I* (lowest risk and associated possibility of reward), to *VII* (highest implicit risk and historically stronger performance).

The selected portfolios reveal that investors, on average, decided for a portfolio with more risk than recommended. Portfolio *III*, which was recommended to 4373 investors, has been picked by only 1253, less than a third. Most users selected portfolios *V*, *VI* and *VII*.

Figure 4.3: ESG portfolios selected by Active group investors.



Note: Portfolios range from *I-ESG* (lowest risk and associated possibility of reward), to *VII-ESG* (highest implicit risk and historically stronger performance).

In the ESG realm, the preference of a more risky portfolio with higher potential rewards was even more pronounced than among the standard portfolios, with the most popular portfolio being *VII-ESG*.

4.4 Trading patterns

Table 4.4 shows the summary statistics for the trading records of all 11,074 investors. In total, the investors made 100,909 deposits and 5,696 withdrawals. Interestingly, the ratio of deposits made by men and women is very similar to the ratio of men to women in the dataset. Men made altogether 78,980 deposits and 4,296 withdrawals, while women made 21,688 deposits against 1,069 withdrawals. This goes against the observed tendency for men to have disproportionately more trades than women and likely stems from the investment product being intended for passive regular investing over time rather than active trading.

Table 4.4: Summary statistics of deposits and withdrawals for all investors.

	Min	1st Q.	Median	Mean	3rd Q.	Max
Women						
Amount of deposits	1	2	4	7.7	11	145
Amount of withdrawals	0	0	0	0.4	0	25
Trades per 30 days	0.02	0.17	0.46	0.62	0.96	11.32
Men						
Amount of deposits	1	2	5	9.6	13	242
Amount of withdrawals	0	0	0	0.52	1	28
Trades per 30 days	0.02	0.19	0.52	0.70	1.03	19.42

Note: 11,074 observations, both Active and Churn group are included.

4.5 Methodology

This section will describe the methodology used to test the hypotheses stated in Chapter 3. It is split into two subsections, Subsection 4.5.1 will cover methodology for testing the first hypothesis and subsection 4.5.2 will go over the methodology for the second hypothesis.

4.5.1 Hypothesis 1

To test Hypothesis 1, the impact of sex on the likelihood of following robo-advice, we transformed the chosen portfolios and recommended portfolios to numerical values from 1 to 7 based on implicit risk level, following the logic of the roman numerals denoting the underlying portfolios (see section 4.2). If a person chose more portfolios, we calculated the average portfolio by taking the mean of the levels of all chosen portfolios. We did not make any distinctions between the same levels of ESG portfolios and their standard portfolio equivalents, i.e. portfolio *VI* has the same risk level as portfolio *VI-ESG*, that is 6. We then compared the risk levels recommended and chosen to create a binary variable *Risky*, taking value 1 if the risk level chosen is greater than recommended and 0 if less than or equal to recommended.

To test the hypothesis, we have opted for a logistic regression. Logit and probit models are frequently used in retail investment research to examine the effects of sociodemographic attributes on binary dependent variables (e.g. Bhattacharya *et al.* 2017, Bailey *et al.* 2011, Korniotis & Kumar 2011, Reiter-Gavish *et al.* 2021). Our logistic regression equation is following:

$$\begin{aligned}\log\left(\frac{P(\text{Risky}_i = 1)}{1 - P(\text{Risky}_i = 1)}\right) = {} & \beta_0 + \beta_1 \times \text{Age}_i + \beta_2 \times \text{InvLen}_i \\ & + \beta_3 \times \text{Sex}_i + \beta_4 \times \log(\text{Income}_i) \\ & + \beta_5 \times \log(\text{Expense}_i)\end{aligned}$$

where *Risky* is a binary variable taking values 1 if a person's selected portfolio (or the mean of chosen porftolios) risk level is greater than that recommended to them by the robo-advisor. *Sex* is a binary variable, 1 for men, 0 for women. Variables *Income* and *Expense* are the individual's monthly income and expense, respectively, log transformed to account for right skews in their distributions. $Income$ and $Expense$ assumed null values, hence before going forward with the log transformation, we have added 1 to all observations, e.g. $log(Income_i + 1)$ for all i. Variable *InvLen* is the estimated (by the investor) investment length in years that the individual expects to invest for at the time of creating their account, we have capped this variable so that in total with the investor's age, they may only expect to invest until they are 85 years old, about 5 years above the current life expectancy at birth in Czechia. The reason to introduce the cap was the existence of unrealistic combinations of *Age* and $InvLen$, where some (older) individuals entered up to 50 years as their estimated investment length.

Expense, although theoretically highly correlated with *Income*, is an infrequent regressor in the literature and may hold valuable information about a person's risk preference. Similarly, a person may adjust his risk preferences based on the horizon they plan to invest for, hence the inclusion of their estimated investment length *InvLen*.

After removing observations with missing values for any of the variables in the regression, the remaining number of observations equaled 10,040, which is well satisfactory for our method of choice. We could not calculate the risk level for the churn group due to the missingness of the chosen portfolio data for all of its members, the implications and possible bias will be discussed in Chapter 6.

Logistic regression

Logistic regression is a regression method that transforms the regression equation such that the output is a probability between 0 and 1. It makes use of the logit link function $ln\left(\frac{P}{1-P}\right)$, where

$$P(Y=1)=\frac{1}{1+e^{-(\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3)}},$$

in which $P(Y=1)$ is the probability that the dependent variable takes value 1 (in our model the probability of a person choosing riskier than recommended portfolio), β_0 is the intercept and β_i and X_i, $i>0$ are the coefficient and variable i. The logit function is hence the natural logarithm of the odds of the dependent variable, set equal to the regular expression of the linear regression equation.

$$ln\left(\frac{P}{1-P}\right)=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3.$$

Therefore, the interpretation of the coefficients is not as straightforward as in the case of regular linear regression. Each single unit increase in the explanatory variable i now changes the log odds of the outcome by its respective coefficient β_i. Further, log odds can be transformed to odds ratios by simply exponentiating the coefficient β_i, e^{β_i}.

Naturally, if the odds ratio equals 1, the variable i has no impact on the odds of the outcome. Odds ratio greater than one implies positive effect of the variable on the odds of the outcome and vice-versa. Specifically, a one-unit change in X_i, say one-unit increase, leads to the increase in odds of the outcome occuring by e^{β_i} times.

Model assumptions

In this subsection, we will conduct tests and discuss whether the assumptions required for logistic regression are satisfied. All tests have been conducted using R language.

Logistic regression has 6 essential assumptions that need to be met. The first one is that the dependent variable is binary, which is satisfied in our case given the specification of *Risky*. The second assumption is independence of observations, this is also satisfied as we are observing unique individuals and we have no base to expect that one has any influence on the decisions and characteristics of the others. The third assumption and the first one we should validate by a test is the assumption of no perfect multicollinearity between the explanatory variables. A way to detect multicollinearity is to calculate the variance inflation factor (VIF). VIF is essentially a calculation involving the R_i^2 of a regression of an explanatory variable i on all other explanatory variables,

$$VIF_i = \frac{1}{1 - R_i^2.}$$

In general, VIF values above 5 are considered indicators of strong presence of multicollinearity (Daoud 2017). The VIF test was conducted using vif function from R's *car* package, table 4.5 shows the results of the test.

Table 4.5: Variance Inflation Factor (VIF) for each explanatory variable in the logistic regression model.

Explanatory variable	VIF
Age	1.06186
Sex	1.018326
InvLen	1.036143
log(Income)	2.700516
log(Expense)	2.706253

Note: Generally, VIF values above 5 are considered concerning levels, suggesting multicollinearity.

As expected, a cause of concern comes from the logs of the variables $Income$ and *Expense*. Despite, their VIF values are not too large and significantly concerning. As a result, this test suggests that the assumption of no perfect multicollinearity is satisfied.

The fourth assumption is the assumption of linearity between the independent variable and the log odds of the dependent variable. To test this assumption, we have decided for a formal test using the Box-Tidwell method and an informal test by plotting each continuous independent variable (that is all variables except the binary *Sex* variable) against the predicted log odds and adding a locally weighted scatterplot smoothing (Lowess) line. The Box-Tidwell test requires all variables to be positive, therefore we added one unit (+1) to the log transformed variables *Income* and *Expense*, only to perform the test. Table 4.6 and Figure 4.4 show the outputs of these tests.

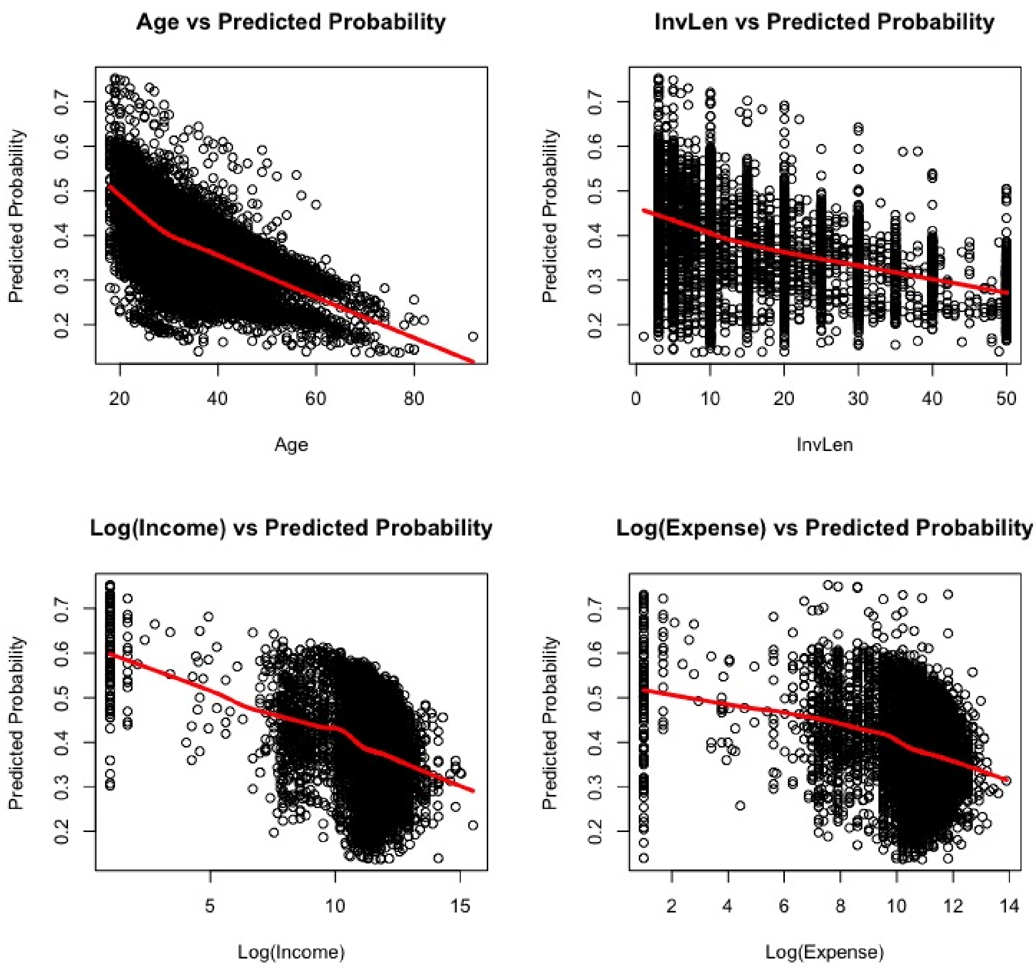
Table 4.6: Results of Box-Tidwell test for non-linearity between independent variables and log odds

Explanatory variable	MLE of lambda	P-value
Age	0.74388	0.2762256
InvLen	-2.01831	0.0006377 ***
log(Income)	6.15377	0.2326151
log(Expense)	-0.92598	0.0017521 ***

Significance codes: * for 10%, ** for 5%, *** for 1%.

Two of the variables tested show significant P-values, indicating non-linearity and consequently suggest a different transformation may be more optimal. The MLE of lambda statistic quantifies what (power) transformations of the associated variable may be more suitable to achieve linearity, e.g. for a better fit, we could consider transforming $log(Expense)$ to the power of -1. However, implementing the transformations suggested by the Box-Tidwell test may not always be plausible due to changes in interpretation. Following is Figure 4.4 showing why a transformation may not be necessary.

Figure 4.4: Scatterplots of independent variables against predicted log odds with Lowess lines.

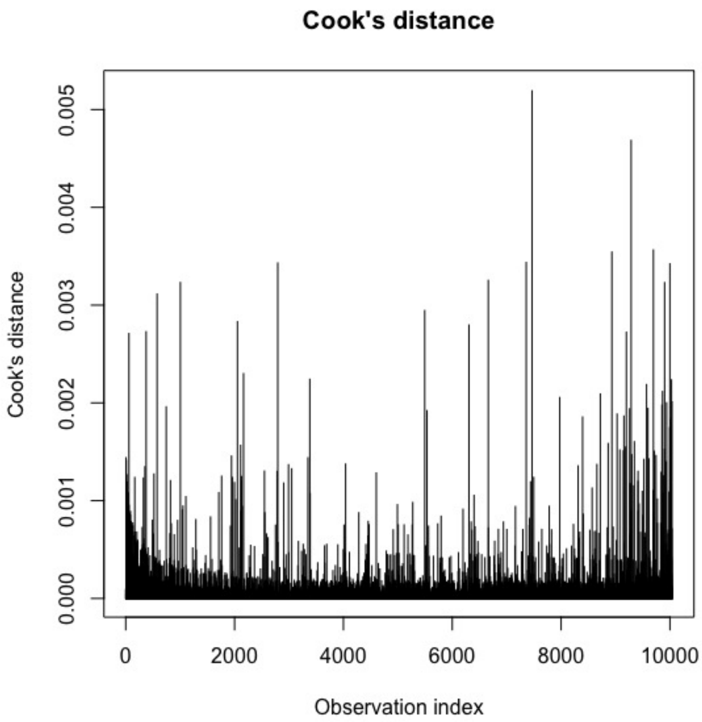


Note: All included variables have retained the adjustments done before perfoming the Box-Tidwell test.

Despite the results of the Box-Tidwell test, the scatterplots and the fitted Lowess lines resemble a relatively linear relationship, with minor deviations at the beginning and in the middle of the respective distributions. Therefore, despite the statistically significant presence of non-linearity, the amount in the whole data may not be strong enough for us to conclude that the fourth assumption is fully invalidated. Based on the results of the visual inspection, we decided to consider the fourth assumption to be satisfied, as the Lowess lines are rather linear in nature. Possible implications of the contra-indicative results of the Box-Tidwell test will be discussed in Chapter 6.

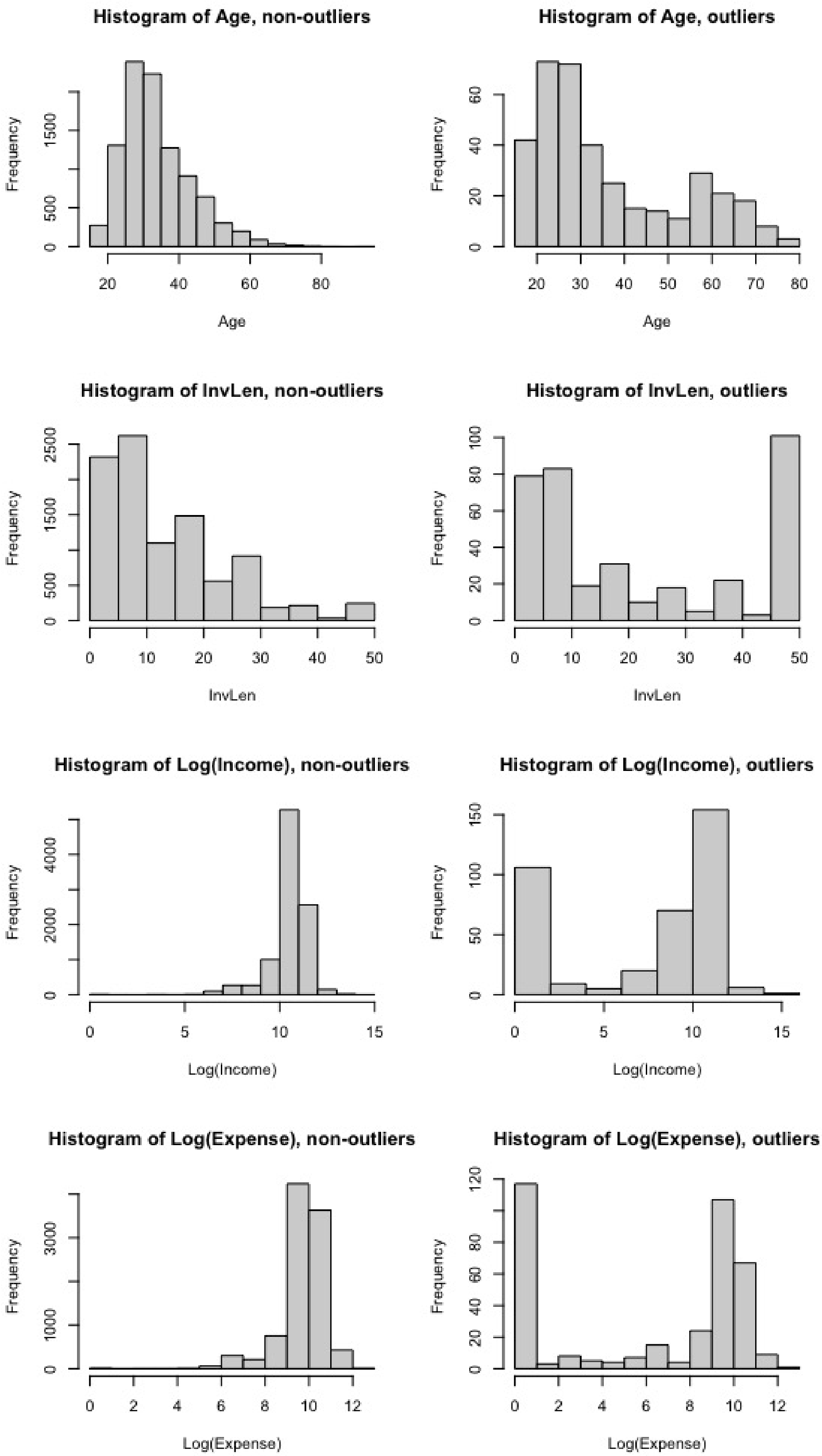
Fifth assumption is that there are no extreme outliers in the data. This assumption can be checked multiple ways, we have opted for the most common one - calculating Cook's distance for each observation in the data set.

Figure 4.5: Cook's distance values for each observation in the logistic regression model



The interpretation of Cook's distances varies. Some suggest a general threshold of 1, where any observation crossing this level is considered an outlier. This is the case with Weisberg (2005), who also recommends to remove influential observations and run the analysis again to examine how obtained coefficients have changed. Others recommend a threshold weighted by the total number of observations n, for example 4/n-k-1, where k is the number of independent variables (Fox 2002). For our purposes, due to the scale of the data set, since each observation may have relatively less impact on the final outcome, it is more appropriate to use the latter and set a threshold based on the number of observations. For simplicity, we will set the threshold to $4/n$. Since we have 10,040 observations, the threshold equals $4/10040 = 0.00039841$. Despite the highest Cook's distance in the dataset being only 0.0052 (rounded to the fourth decimal place), in total, 371 observations have surpassed the calculated threshold. Figure 4.6 displays differences in distributions of the explanatory variables between the group of outliers versus non-outliers (as delimited by the threshold).

Figure 4.6: Explanatory variables distributions, outliers versus non-outliers



Not surprisingly, the outliers are mostly observations from the tails of the aggregate distributions. For *Age*, it is often people younger than 20 or older than 60 that fall into the outlier group. The $InvLen$ outliers are driven by people who expect to invest for 50 years. For $Log(Income)$, it is investors who have reported no monthly income and similarly for $Log(Expense)$.

Outliers can be dealt with either by adjusting the data (removing the observations or replacing them with mean or median) or adjusting models, however, we believe most of the outliers in our data set carry important information about the real distributions. There is no reason to believe that *Age* or $InvLen$ are wrong data entries and discard all older people or younger people who expect to invest for a long time. On the same note, we cannot discern whether a person really does have no monthly income, although improbable, or is living off saved funds.

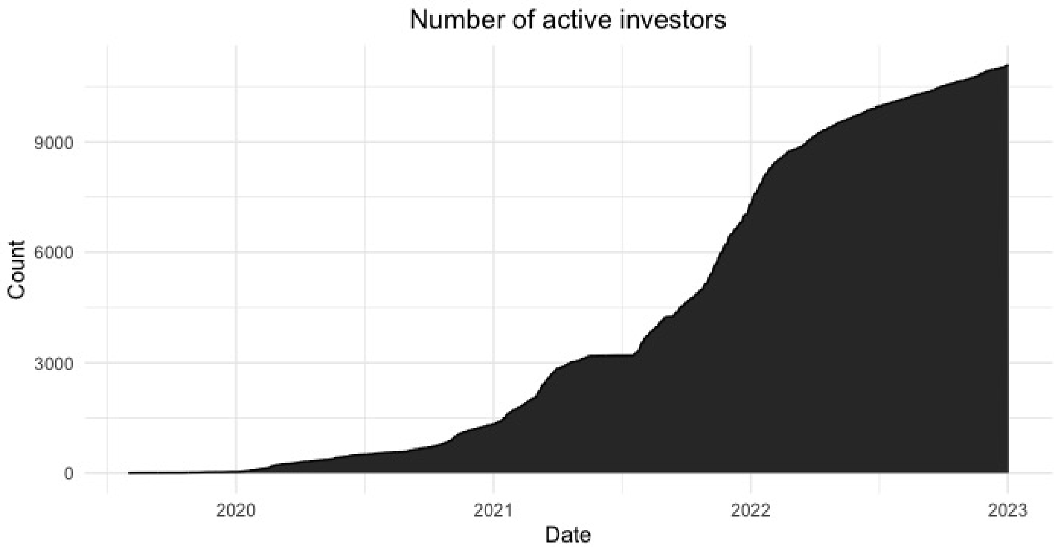
To address the impact of outliers, we will run the logistic regression twice, once with all observations and second time without the outliers, to compare how the obtained coeffiecients and other statistics change.

4.5.2 Hypothesis 2

To test Hypothesis 2, specifically whether there is a difference between the sexes in investment volume adjustments during periods of high market volatility, we have decided for panel data regression. Due to the relatively low frequency of trades per month (less than 1 on average, see Table 4.4) for both men and women, we have aggregated deposit and withdrawal volumes data into monthly $Total\_volume_{i,t}$, measuring the total sum of all flows investor i has conducted in a given month t.

One crucical obstacle in running panel data regression on our data set is that the number of investors is not constant throughout the 42 months. Figure 4.7 displays the evolution in number of active investors. The challenge is therefore in the choice between the number of periods t and investors i we run the regression on.

Figure 4.7: Number of active investors in the data set

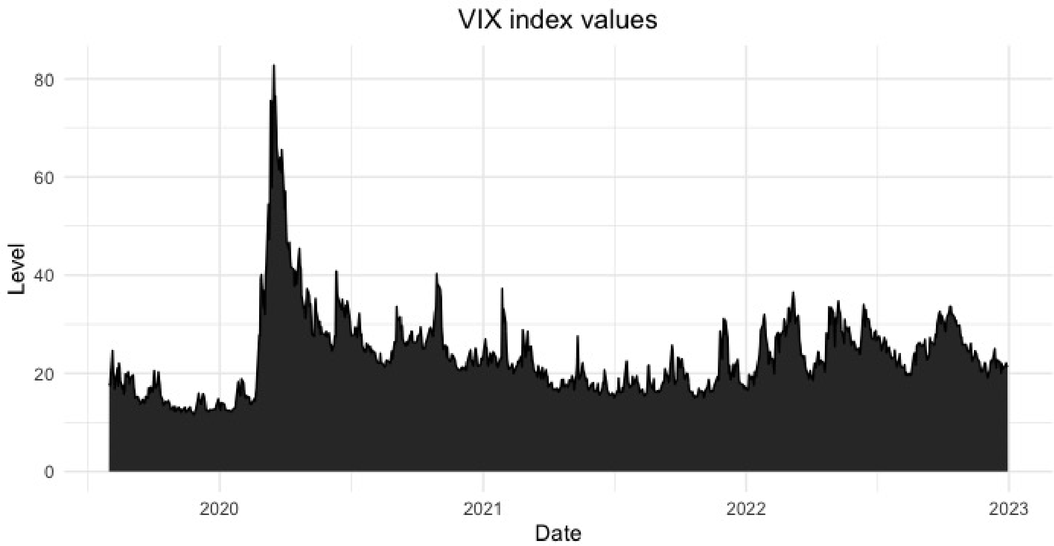


Note: The sample period is from August 2019 to January 2023 through which the total number of users reached 11,074, 590 of whom have closed their accounts some time during the period.

Notably, the most appealing subset of the sample period is right near the end of Q2 2021, when the number of investors crosses 3000, this subset would allow us to retain up to 22 periods, depending on the exact cutoff time, while allowing for a decent sample size.

Nevertheless, the decision should be contrasted with the levels of the VIX index, which is the essential variable to the analysis. VIX is a market index made available by the Chicago Board of Options Exchange (CBOE), it is calculated continually from S&P 500 options and aims to predict the expected volatility of the S&P 500 index in the next 30 days (Chicago Board Options Exchange 2019). It is considered a good representative of the market sentiment in regards to expected volatility (Ding *et al.* 2021). Figure 4.8 highlights VIX values throughout the full sample period.

Figure 4.8: VIX levels in the full sample period



Note: The figure displays VIX close values for specific days of the sample period. Data was collected from the Chicago Board Options Exchange who maintain the index.

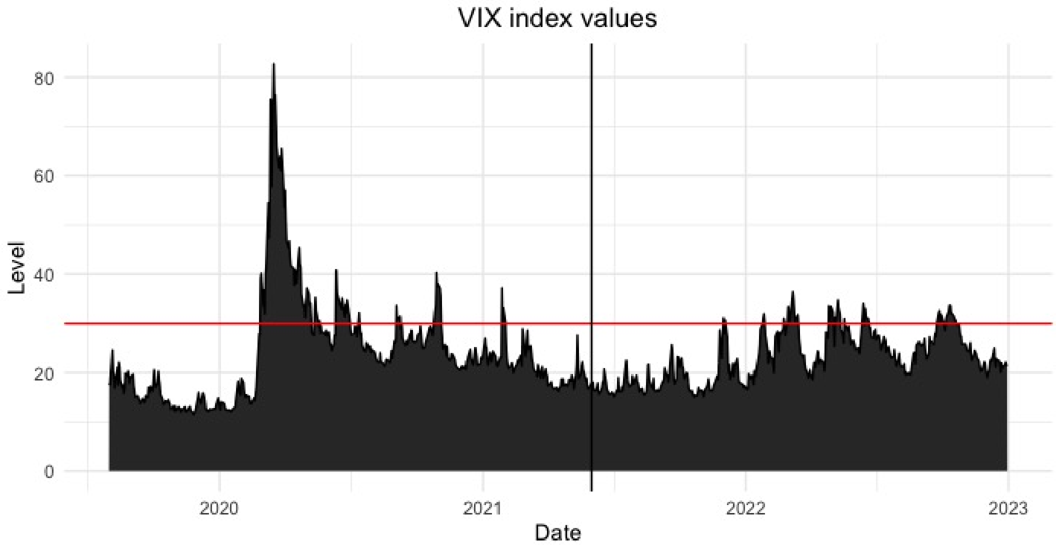
Unfortunately, during the height of the Covid-19 pandemic, where VIX peaked (March 2020), only about 200 investors were active, a sample not large enough for the type of analysis we want to conduct. At the suggested cutoff point in Q2 2021, the VIX levels were already returning to those before the pandemic, a manifestation of its tendency for mean-reversion (Nielsen & Posselt 2021).

From the figure, especially after Q1 2021, one can notice a long-term, rather stable evolution disrupted by peaks relatively high in magnitude, but short in duration. Since we have aggregated our flows data into total monthly volumes, we needed to transform the information in daily VIX data to fit our monthly panel structure. Due to the short durations of the peaks, the information they carry may be lost if we simply average the days to get monthly data. Instead, we wanted to preserve the significance of the peaks and opted for a dummy variable $Volatile_t$ that equals one if VIX crosses a certain threshold in the given month and 0 otherwise.

The threshold value itself is worthy of discussion, the aim is to set it so that only peak levels of VIX cross it, while the general trend remains below. After careful data examination, we chose 30 to be the delimiting value. Figure 4.9

explicitly shows the chosen threshold over the data from Figure 4.8

Figure 4.9: VIX threshold



Note: The figure displays VIX close values for specific days of the sample period. Data was collected from the Chicago Board Options Exchange who maintain the index. The horizontal red line refers to the chosen threshold of 30. The vertical black line refers to the chosen cut-off point (June 1st, 2021) for subsetting active investors who have set up accounts before then.

We have chosen the first of June 2021 as the point at which we start analysing all active investors and their investment decision over the rest of the period, resulting in total of 18 time periods (January 2023 was discarded due to lack of data for most of the month). The final data set is a balanced panel of 2828 investors who have made their initial deposits at latest in May (to omit the impact of the first initial deposit) and for whom we have all the data to specify the following preliminary model:

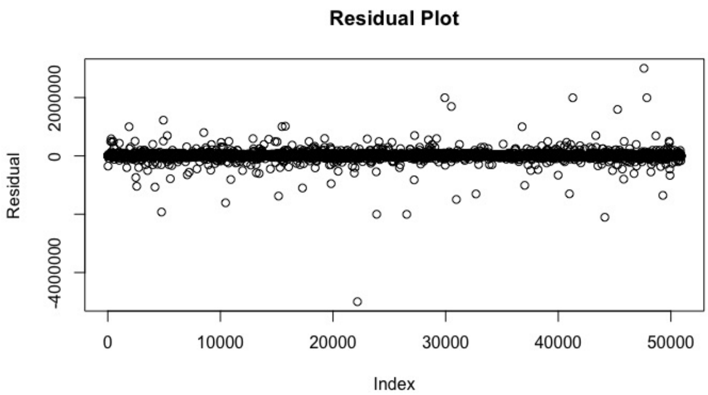
$$
\begin{aligned}
Total\_volume_{i,t} = &\beta_0 + \beta_1 \times Sex_i + \beta_2 \times Age_i + \beta_3 \times InvLen_i \\
&\beta_4 \times \log(Income_i) + \beta_5 \times Volatile_t + \beta_6 \times Volatile_{t-1} \\
&+ \beta_7 \times Sex_i \cdot Volatile_t + \beta_8 \times Sex_i \cdot Volatile_{t-1} + u_{i,t},
\end{aligned}
$$

where $Total\_volume_{i,t}$ is the total flows (positive or negative) a given investor i makes in month t, Sex_i is the sex of investor i, 1 for males and 0 for females and $Volatile_t$ specifies, whether VIX has breached the threshold level of 30 in the month t. We have decided to include the first lag of the dependent variable as well as the first lag of $Volatile$ to account for a possible delay in investor reaction to the heightened volatility. Finally, the interaction term between Sex and $Volatile$ (and its first lag) are the variables of interest, we want to see whether their coefficient is statistically significant.

Model assumptions

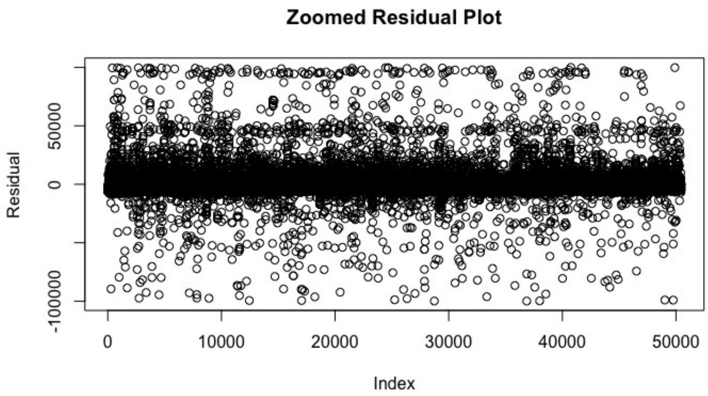
Just as for the first hypothesis, this subsection will cover the assumptions required for panel data analysis. We have a balanced panel data set with $N = 2828$ and $T = 18$. To achieve the same amount of periods for all investors, we had to drop all those from the churn group, in total 125. The first assumption is the assumption of linearity between the dependent variable and the independent variables, this can be observed visually by inspecting the residuals. Figures 4.10 displays the residuals, while Figure 4.11 is a "zoom-in" to provide more detail.

Figure 4.10: Residual plot of random effects model



Note: Residual plot of random effects model for visual inspection of linearity.

Figure 4.11: Residual plot of random effects model, zoomed



Note: Zoomed residual plot of random effects model for visual inspection of linearity.

Given the inspection, there is no base to suggest non-linearity, however, the first plot sheds light on the presence of outliers which should be addressed by further examination.

The second assumption is the assumption of zero correlation between the independent variables and individual-specific effects. This assumption is usually investigated using Hausman test to decide between random effects and fixed effects model (Hausman 1978). We have conducted Hausman test using *R* function *phtest* from *plm* package. The null hypothesis of the Hausman test is that there is no correlation between the individual-specific effects in the errors and explanatory variables in the model. The result of the Hausman test we conducted bore a P-value of almost 1, meaning we fail to reject the null hypothesis and cannot therefore say that the fixed effects model is a better fit.

The third assumption is the assumption of no perfect multicollinearity (similarly to the logistic model regression). We have again used Variance Inflation Factor to assess the collinearity between the independent variables. Table 4.7 displays their VIF values.

Table 4.7: Variance Inflation Factor (VIF) for each explanatory variable in the random effects model.

Explanatory variable	VIF
Sex	2.316229
Age	1.072968
InvLen	1.049024
Income	1.047758
Volatile	5.860315
lag(Volatile)	5.860315
Sex*Volatile	7.306744
Sex*lag(Volatile)	7.306744

Note: Generally, VIF values above 5 are considered concerning levels, suggesting multicollinearity.

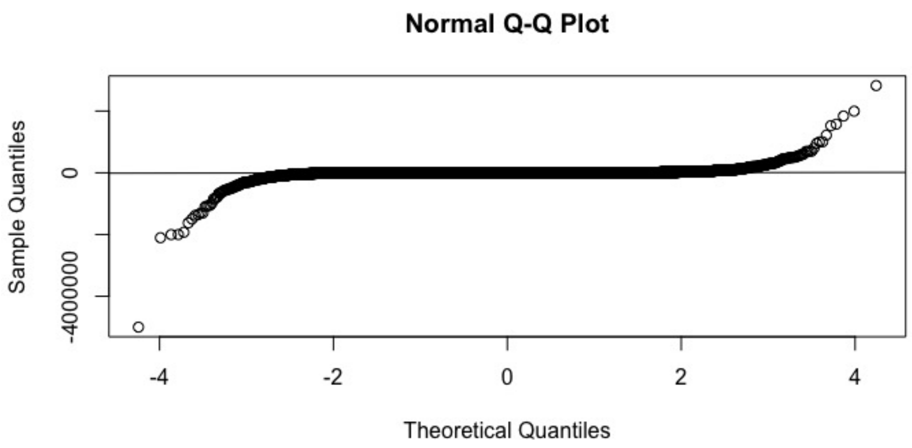
In our random effects model, the VIF values breach the threshold of 5. However, our primary variables (except $Volatile$) remain well-behaved, the only case for collinearity comes from the lags and interaction terms, but this is natural given the fact that they are generated from the primary variables. Calculating the VIF values without the presence of lags and interaction terms results in near perfect non-collinearity levels (close to 1 in VIF value) in the primary variables. Therefore, we consider the third assumption met.

The fourth assumption is the assumption of homoskedasticity. This can be tested using Wald test for groupwise heteroskedasticity (Greene 2008). From running the test, we have obtained a P-value <0.0001, suggesting presence of groupwise heteroskedasticity. We therefore consider this assumption violated and will continue by using robust standard errors to account for the presence.

Assumption number 5 is no autocorrelation in the residuals. This assumption can be tested using the Woolridge test, its null hypothesis being no first-order serial autocorrelation in idiosyncratic errors (Wooldridge 2010). We conducted the test and received a P-value <0.0001, strongly suggesting the presence of first-order autocorrelation. To provide at least some remedy, we will include lags of the dependent variable in the model, however this results in the Hausman test result becoming statistically significant and suggests we should switch to fixed effects model, which we will do and in addition also incorporate some lags of the dependent variable.

The sixth assumption is normality of errors. We can test this assumption both by visual inspection and the formal Shapiro-Wilk test (Shapiro & Wilk 1965). Figure 4.12 displays the Q-Q plot. For normality to hold, the points should follow the line well along the quantiles, which they do not, suggesting this assumption is violated. This is confirmed by the Shapiro-Wilk's test P-value of <0.0001.

Figure 4.12: Q-Q plot for normality of errors in random effects model



Before testing Hypothesis 1, we have met some assumptions of logistic regression fully and some remain open to interpretation and therefore hold implications to our model specification. Specifically, the assumption of linearity between the variables $InvLen$ & $Log(Expense)$ and the log odds has been shown to not hold using the formal Box-Tidwell test, but visual inspection suggested the issue does not necessarily prevent us from conducting the analysis. The data contains outliers, detected using Cook's distance. The possible impact of any single one of them is not large, though, and most of them seem to hold valuable information and therefore should not be removed. For comparison, we will conduct the test twice, once with the full data set and second time without outliers.

Regarding Hypothesis 2, some of the assumptions were met, unfortunately others, namely those concerning residuals, were not. In detail, the assumption of homoskedasticity was ruled out by Wald test, the assumption of no autocorrelation in the residuals was rejected by Woolridge test and the assumption of normality of errors was shown not to hold by using both visual inspection and Shapiro-Wilk test. We will calculate robust standard errors using the *vcovHC* function in R's *plm* package, which should correct for heteroskedasticity and we will conduct both random effects and fixed effects estimation.

Chapter 5

Results

In this chapter results of the tests described in Chapter 4 will be presented. Section 5.1 covers the results of Hypothesis 1 testing while Section 5.2 reports on findings regarding the test of Hypothesis 2.

5.1 Results of logistic regression

Table 5.1 presents classification metrics measuring model fit.

Table 5.1: Classification metrics, All Observations model and Without Outliers model

Metric	All Observations	Without Outliers
Accuracy	0.6154 (0.6058, 0.6250)	0.6278 (0.6181, 0.6374)
Kappa	0.0588	0.0853
McNemar's Test P-Value	<0.0001	<0.0001
Sensitivity (Recall for '0')	0.1378	0.1976
Specificity (Recall for '1')	0.9130	0.8771
Positive Predictive Value	0.4967	0.4825
Negative Predictive Value	0.6296	0.6535
Prevalence	0.3836	0.3669
Detection Rate	0.0529	0.0725
Detection Prevalence	0.10647	0.1503
Balanced Accuracy	0.5254	0.5374
N	10,040	9,669

Note: The cut-off line for classifying 'Risky' as 1 (the positive class) based on the predicted probabilities was set to 50%. N refers to the number of observations used.

Visibly, both models have low sensitivity with the chosen classification threshold of 0.5. The balanced accuracy is only marginally better than a fair coin flip, with the model Without Outliers having a slight edge at identifying risky investors. Table 5.2 shows the odds ratios (calculated from the coefficient estimates) with their respective 95% confidence intervals. The McFadden's and Nagelkerke's pseudo R^2 are small, but in similar range for logit (or probit) models utilized in existing research (see for example Bhattacharya *et al.* 2017, Bailey *et al.* 2011).

Table 5.2: Logistic regression results

	Dependent variable: Risky	
	All Observations	Without Outliers
Age	0.974*** (0.970, 0.978)	0.961*** (0.956, 0.966)
Sex	1.658*** (1.561, 1.754)	1.874*** (1.771, 1.977)
Log(Income)	0.921*** (0.877, 0.965)	0.959 (0.864, 1.053)
Log(Expense)	1.017 (0.974, 1.061)	1.070 (0.978, 1.162)
InvLen	0.979*** (0.975, 0.983)	0.967*** (0.963, 0.971)
Constant	2.881*** (2.568, 3.194)	1.824*** (1.391, 2.257)
N	10,040	9,669
Log Likelihood	-6,507.49	-6,069.76
Akaike Inf. Crit.	13,026.98	12,151.52
McFadden's R^2	0.027	0.045
Nagelkerke's R^2	0.047	0.078

Note: Results are reported as odds ratios, 95% confidence intervals are in brackets. Sex is a binary variable equal to '0' for women and '1' for men. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Age, *Sex*, *Log*(*Income*) and *InvLen* have statistically significant coefficient estimates at the 1% level, while *Log*(*Expense*) is insignificant. The odds ratios show how the probability of *Risky* taking value 1 changes given one unit increase in the dependent variable. Our results suggest that as people get older, the likelihood that they will choose a riskier than recommended portfolio decreases, the same goes for having longer investment horizon. While the *InvLen* result may seem contra-intuitive, given the odds ratios of *Age* and their possible relationship, it may simply be the case that people who are 'in the game' for a shorter horizon may have a larger propensity towards risk, this result goes against the findings of Veld-Merkoulova (2011), who finds the opposite (and more intuitive) relationship between self-reported expected investment horizon and assumed portfolio risk. This finding requires more investigation, however, it is beyond the scope of this thesis. Being a man is associated with a 65.8% larger probability of picking a riskier than recommended portfolio. $Log(Income)$loses its statistical significance with the removal of outliers. However, as noted in the previous chapter, each individual outlier has almost no effect on the result of the regression, only their summed influence impacts the model significantly. The threshold for calculating Cook's distance is sensitive to the amount of observations, the outliers identified seem as plausible representations of the population, however the P-value of the $Log(Income)$ coefficient equals 0.38 in the Without Outliers model, implying a a strong shift towards small significance.

Our findings about following robo-advice are therefore in line with the research on following human financial advice, e.g. Reiter-Gavish *et al.* (2021) show positive effect of age, sex (female) and the log of wealth on adhering to received advice.

5.2 Results of panel data regression

In chapter 4, we tested the most important assumptions of the estimations and stated that assumptions of homoskedasticity, no serial autocorrelation in errors and normality of errors are all violated. While these violations do not bias the coefficients, they have negative influence on our ability to correctly calculate standard errors. To redeem, tables 5.3 and 5.4 show results of random effects and fixed effects estimations, respectively, with robust standard errors. To remind, Sex is a binary variable, 1 for males and 0 for females. Volatile is a binary variable equatl to 1 if the given callendar month experienced VIX levels greater than 30 and 0 otherwise.

In the random effects model, Sex, $InvLen$, $Income$, $lag(Volatile)$ and $Sex * Volatile$ are all statistically significant. Sex, $InvLen$ and $Income$ are positive, indicating that men, people with longer investment horizons and individuals with greater income are all expected to deposit more each month than members of the opposite groups. On the other hand, in a month subsequent to that which experienced high volatility ($VIX > 30$), individuals are expected to decrease their deposits by 2,400 CZK. Inside the same month where VIX crosses the level of 30, men are expected to decrease their volume by 2,600 CZK more than women. Meanwhile, in the fixed effects model, which differences the individual specifics away, the third lag of $Total\_volume$, the first lag of $Volatile$ and the interaction term $Sex * Volatile$ are all statistically significant. Similarly to the random effects model, in a month subsequent to the one experiencing a VIX peak above 30, investors are expected to decrease their volume by 1,991 CZK, while men decrease their volume by 2,738 CZK more than women inside the same month where the peak occurs.

Both models achieve a very low R^2, but as is the case with the logistic regression model, this is quite usual for behavioral finance panel data analysis.

Table 5.3: Random effects estimation

	Dependent variable:
	Total_volume
Sex	1,758.869**
	(792.761)
Age	29.163
	(24.195)
InvLen	43.087**
	(18.361)
Income	0.018***
	(0.004)
Volatile	466.784
	(1,087.377)
lag(Volatile)	−2,404.531**
	(1,087.377)
Sex*Volatile	−2,595.508**
	(1,252.939)
Sex*lag(Volatile)	−485.472
	(1,252.939)
Constant	454.197
	(1,131.493)
Observations	50,904
R^2	0.002
Adjusted R^2	0.002
F Statistic	113.459*** (df=8)

Note: Monetary variables are in CZK. *p<0.1; **p<0.05; ***p<0.01

Table 5.4: Fixed effects estimation

	Dependent variable:
	Total_volume
lag(Total_volume, 1)	0.025
	(0.054)
lag(Total_volume, 2)	0.040
	(0.034)
lag(Total_volume, 3)	-0.030^{***}
	(0.011)
Volatile	863.887
	(786.157)
lag(Volatile)	$-1{,}991.408^{**}$
	(809.754)
Sex*Volatile	$-2{,}738.604^{***}$
	(1,034.029)
Sex*lag(Volatile)	-465.000
	(963.481)
Observations	45,248
R^2	0.004
Adjusted R^2	-0.062
F Statistic	26.731^{***} (df = 7; 42413)
Note: Monetary variables are in CZK.	*p<0.1; **p<0.05; ***p<0.01

The interaction term $Sex*Volatile$ is significant at 5% level in both models, suggesting that men in our data set do react in volatile months more severely than women do. Based on these results, we would be able to reject the null hypothesis of no difference in adjustments between the sexes in callendar months experiencing high expected volatility (VIX>30).

While these results have significant coefficients in the variables of interest for our hypotheses, they retain problematic attributes that distort our ability to draw conclusions. These attributes will be summarized in the following chapter.

Chapter 6

Discussion

This chapter will include robustness checks of our results, it will follow the same structure as previously, starting with the logistic regression robustness check and following with panel data regressions robustness checks. The end of the chapter contains section about limitations.

To validate the calculated coefficients of the logistic regression (previously reported as odds ratios in the results chapter), we will perform bootstrapping. Bootstrapping is essentially resampling the original sample with replacement and running the regression with the original model specifications, repeated many times over. We have opted for 1000 iterations of the process and Table 6.1 reports the obtained 95% confidence intervals of the coefficients.

Table 6.1: Bootstrapping results for logistic regression

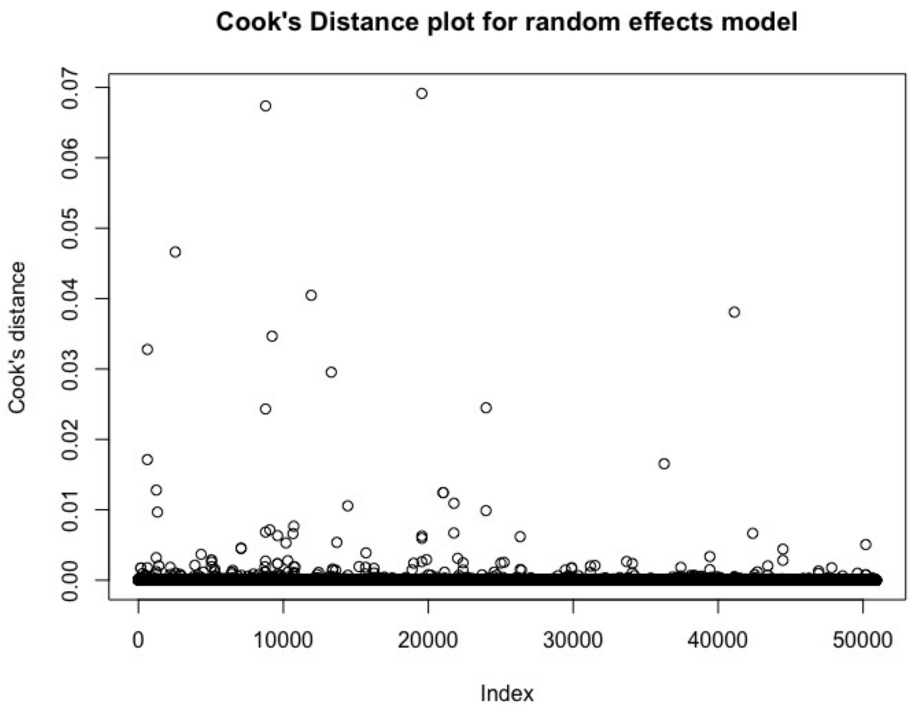
	95% Confidence intervals	
	All Observations	Without Outliers
Age	(-0.0311, -0.0221)	(-0.0443, -0.0349)
Sex	(0.4091, 0.6047)	(0.5340, 0.7271)
Log(Income)	(-0.1264, -0.0371)	(-0.1305, 0.0529)
Log(Expense)	(-0.0267, 0.0638)	(-0.0202, 0.1513)
InvLen	(-0.0248, -0.0173)	(-0.0372, -0.0293)
Constant	(0.744, 1.366)	(0.1592, 1.0131)

Note: Bootstrapping-derived 95% CIs based on 1000 iterations of resampling and running the logistic regression specified in Chapter 4. As opposed to results in Chapter 5, we decided to report these results as regular coefficients instead of odds ratios to better capture which intervals contain 0. Sex is a binary variable equal to '0' for women and '1' for men.

Just as in the original sample, the only coefficients which include 0 in their 95% confidence intervals are $Log(Expense)$for the All Observations model and $Log(Expense)$ with $Log(Income)$, suggesting some robustness to the coefficient results reported in the previous chapter.

Regarding our panel data regressions, we have explicitly stated in Chapter 4 the presence of outliers based on visual inspection of residual plots. We should therefore investigate their impact on the results presented in Chapter 5. To do so, we calculated the Cook's distance using the random effects model formula on the panel, but using standard linear regression function, disregarding the panel structure. This has resulted in identifying 470 (<1%) observations being flagged by using the same method as in the case of the logistic regression, e.g. all observations with Cook's value greater than $4/N$, where N is the number of observations. Figure 6.1 displays the plot of calculated Cook's distances.

Figure 6.1: Cook's distance values for each observation in the panel data set



We will now execute the two regressions again with robust standard errors and without the outliers to see the impact on the resulting statistics, mind that removing problematic rows will lead us to an unbalanced panel. Tables

6.2 and 6.3 report random and fixed model estimations respectively, without the outliers flagged in the previous step.

Table 6.2: Random effects estimation, removed outliers

	Dependent variable:
	Total_volume
Sex	1,628.727**
	(676.569)
Age	30.005
	(33.810)
InvLen	43.473***
	(15.878)
Income	0.018**
	(0.007)
Volatile	487.556
	(792.192)
lag(Volatile)	−2,446.695***
	(785.245)
Sex*Volatile	−2,506.296***
	(942.134)
Sex*lag(Volatile)	−128.442
	(942.912)
Constant	422.689
	(1,156.685)
Observations	50,173
R^2	0.002
Adjusted R^2	0.002
F Statistic	116.978*** (df = 8; 50164)
Note: Monetary variables are in CZK.	*p<0.1; **p<0.05; ***p<0.01

Table 6.3: Fixed effects estimation, removed outliers

	Dependent variable:
	Total_volume
lag(Total_volume, 1)	-0.050^{***}
	(0.019)
lag(Total_volume, 2)	0.017
	(0.041)
lag(Total_volume, 3)	-0.026^{**}
	(0.012)
Volatile	987.590
	(813.754)
lag(Volatile, 1)	$-2{,}110.349^{***}$
	(801.417)
Sex*Volatile	$-2{,}416.375^{**}$
	(958.040)
Sex*lag(Volatile, 1)	−210.454
	(951.049)
Observations	44,076
R^2	0.006
Adjusted R^2	−0.062
F Statistic	35.880^{***} (df = 7; 41241)
Note: Monetary variables are in CZK.	*p<0.1; **p<0.05; ***p<0.01

The practical difference in the random effects model without outliers is almost non-existent, $lag(Volatile)$, $Sex * Volatile$ and $InvLen$ are now statistically significant at the 1% level, but without great impact to their coefficients magnitude. For the fixed effects estimation, the first lag of $Total_volume$ becomes greatly significant (at 1% level) from no significance at all, suggesting the outliers had great impact. $Sex * Volatile$ increases in magnitude, while $lag(Volatile)$ decreases. Therefore, the outliers had zero to no impact on the random effects specification and significant impact on some of the variables in the fixed effects model.

6.1 Limitations

In Chapter 4, we have formally shown that the assumption of linearity between the independent variables and log odds is violated using the Box-Tidwell test. After visual inspection, we have decided to continue with the analysis. However, the assumption is crucial to the interpretation of the results. If the violation is indeed significant, it may lead to incorrectly reported relationships between the dependent variables and the log odds as well as incorrect P-values. Also our bootstrapping results would be invalidated as they use the same specification of the regression equation.

Furthermore, we have realized our binary dependent variable to not be well defined, the issue being the inability of investors with the highest-risk portfolio being recommended to select even riskier portfolio. Table 6.4 displays results of running the logistic regression without the mentioned problematic class of investors. In total, 1441 observations were removed from the full data set reported in Chapter 5.

Table 6.4: Logistic regression results without mis-specified class

	Dependent variable: Risky
Age	0.979*** (0.974, 0.983)
Sex	1.906*** (1.807, 2.005)
Log(Income)	0.968 (0.923, 1.012)
Log(Expense)	0.991 (0.946, 1.035)
InvLen	0.992*** (0.988, 0.995)
Constant	1.809*** (1.485, 2.133)
N	8,639
Log Likelihood	-5,805.480
Akaike Inf. Crit.	11,622.960

Note: Results are reported as odds ratios; 95% confidence intervals are in brackets. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

From the table, one can observe that the relationships in the significant variables have equal directions as in the All Observations and Without Outliers models. The $Log(Income)$ becomes insignificant, similarly as in the case of removing outliers. $InvLen$, despite being significant, now has very little practical impact (compared to other variables) on the likelihood of choosing a riskier-than-recommended portfolio.

For the panel data regressions, we did not satisfy homoskedasticity assumption, no autocorrelation in the residuals assumption and the normality of errors assumption. We have run the regression using robust standard errors to account for heteroskedasticity. However, the presence of autocorrelation may lead the coefficients to be biased and the standard errors to be unreliable together with the P-values. This suggests we may have derived some information from the analysis, but the results of hypothesis testing are inconclusive.

Chapter 7

Conclusion

We have investigated the impacts of sociodemographic attributes on compliance with robo-advice in choosing ready-made portfolios of ETFs that vary in risk. Each investor was recommended a specific portfolio based on suggested risk level determined by an algorithm, we then studied whether specific variables such as sex, age, estimated investment horizon and the natural logarithms of income and expense play a role in the decision of an individual to select a riskier-than-recommended portfolio.

In following financial advice, the literature is scarce. A recent study conducted by Reiter-Gavish *et al.* (2021) found that, on average, older, wealthier people and women are more likely to comply with financial advice received from a human financial advisor. To our best knowledge, the relationship between robo-advice and retail investor advice compliance is not well-established, particularly in the realm of modern investment products such as ETFs and much less in ready-made portfolios of ETFs. We use logistic regression model to estimate coefficients of the mentioned variables and find that, in compliance with Reiter-Gavish *et al.* (2021), older, wealthier (as proxied by income) people and women in particular are more likely to comply with robo-advice in selecting a ready-made portfolio of ETFs that carries less or equal risk to that recommended by the algorithm.

However, we show that linearity between the independent variable and the log odds does not hold using Box-Tidwell test, which may bias and invalidate our results. Based on visual inspection of the regressors against predicted log odds, resembling a linear realtionship, we have decided to proceed with the analysis. Furthermore, using Cook's distance, about 3% of the data set was determined to be outliers, we have conducted the test again without the outliers and this resulted in the coefficient of $Log(Income)$ to become statistically insignificant. Other coefficients remain significant with minor differences in their magnitude, except for the sex variable. Including outliers in the data set lead to men having a 65.8% greater odds of selecting a riskier-than-recommended portfolio compared to women and excluding the outliers increased the statistic to 87.4%.

Contrary to relationship between investment horizon and risk preference documented by Veld-Merkoulova (2011), we report a slight (but statistically significant) positive relationship between the estimated investment horizon (self-reported by the investors) and the probability of selecting a portfolio of at most the recommended risk level. This is not very intuitive, as one would expect investors who plan to invest for (possibly) decades to be less sensitive to short term risks. One possible explanation would be a specific risk profile affecting most of the investors in our data set. Majority of investors in our data set live in the Czech republic, a country with a very recent introduction (thanks to the socialist regime ending in 1993) to financial markets compared to other more established western states. Veld-Merkoulova (2011) documents the relationship in Dutch households.

To quantify the impact of market sentiment on retail investors' flows adjustments in connection to sociodemographic variables, we have estimated fixed effects and random effects models. Our results of random effects show negative relationship between total volume of flows in a given month and a dummy variable equal to one if the preceding month experience VIX levels greater than 30. Importantly, being male was negatively associated with flows inside the same month where VIX crossed the threshold of 30, suggesting men react more to expected volatility than women do. The fixed effects estimation supported the stated relationships.

These results are somewhat in line with the observed tendency of men to perform more transactions than women in general and being prone to overconfidence in their trading ability Barber & Odean (2001). In our random effects model, being a man is also associated with greater monthly volume (deposits into investment account) of 1,759 Czech Koruna.

Nevertheless, we have shown the assumptions of homoskedasticity, no autocorrelation in the residuals and normality of residuals to not hold. We have opted for robust standard errors, but we have not fully remedied autocorrelation and this would mean our coefficients might be biased. Our results therefore remain inconclusive as to whether men do adjust their trading patterns differently than women in times of high expected market volatility.

Bibliography

Agnew, J. R., P. Balduzzi, & A. E. Sunden (2003): "Portfolio choice and trading in a large 401(k) plan." *American Economic Review* **93(1)**: pp. 193–215.

Allais, M. (1953): "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole Americaine." *Econometrica* **21(4)**: pp. 503–546.

Back, C., S. Morana, & M. Spann (2023): "When do robo-advisors make us better investors? the impact of social design elements on investor behavior." *Journal of Behavioral and Experimental Economics* **103**: p. 101984.

Bailey, W., A. Kumar, & D. Ng (2011): "Behavioral biases of mutual fund investors." *Journal of Financial Economics* **102(1)**: pp. 1–27.

Barber, B. M. & T. Odean (2001): "Boys will be boys: Gender, overconfidence, and common stock investment." *The Quarterly Journal of Economics* **116(1)**: pp. 261–292.

Barberis, N. & W. Xiong (2009): "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation." *The Journal of Finance* **64(2)**: pp. 751–784.

Bhattacharya, U., B. Loos, S. Meyer, & A. Hackethal (2017): "Abusing etfs." *Review of Finance* **21(3)**: pp. 1217–1250.

Bollen, N. P. B. (2006): "Mutual fund attributes and investor behavior." *Journal of Financial and Quantitative Analysis* **41(4)**: pp. 683–708.

Boston Consulting Group (2021): "Global asset management 2021: The $100 trillion machine." Accessed: June 09 2023.

Byrnes, J. P., D. C. Miller, & W. D. Schafer (1998): "Gender differences in risk taking: A meta-analysis." *Psychological Bulletin* **125(3)**: pp. 367–383.

Cen, X. (2021): "Smartphone trading technology, investor behavior, and financial fragility." Available at SSRN. `https://ssrn.com/abstract=3424405` or `http://dx.doi.org/10.2139/ssrn.3424405`.

Český statistický úřad (2021): *Základní výsledky Sčítání lidu, domů a bytů, Česká republika*. Praha: Český statistický úřad, kód publikace: 170389-23, Č.j.: csu-008523/2023-61 edition. Sčítání lidu, domů a bytů.

Chicago Board Options Exchange (2019): "Vix index." `https://www.cboe.com/tradable_products/vix/`. Accessed: 2024-01-03.

Daoud, J. I. (2017): "Multicollinearity and regression analysis." In "Journal of Physics: Conference Series," volume 949, p. 012009.

De Long, J. B., A. Shleifer, L. H. Summers, & R. J. Waldmann (1990): "Noise Trader Risk in Financial Markets." *Journal of Political Economy* **98(4)**: pp. 703–738.

Deloitte (2021): "The rise of newly empowered retail investors." https://www2.deloitte.com/us/en/pages/financial-services/articles/the-future-of-retail-brokerage.html. Accessed: June 08 2023.

Dhar, R. & N. Zhu (2006): "Up Close and Personal: Investor Sophistication and the Disposition Effect." *Management Science* **52(5)**: pp. 726–740.

D'Hondt, C., M. Petitjean, & Y. Elhichou (2023): "Uncovering the profile of passive exchange-traded fund retail investors." Available at SSRN: https://ssrn.com/abstract=3522963.

Ding, W., K. Mazouz, & Q. Wang (2021): "Volatility timing, sentiment, and the short-term profitability of vix-based cross-sectional trading strategies." *Journal of Empirical Finance* **63**: pp. 42–56.

Fox, J. (2002): *An R and S-Plus Companion to Applied Regression*. Sage.

Gervais, S. & T. Odean (2001): "Learning to be overconfident." *Review of Financial Studies* **14**: pp. 1–27.

Greene, W. H. (2008): *Econometric Analysis*. Prentice Hall.

Grinblatt, M. & M. Keloharju (2001): "What makes investors trade?" *The Journal of Finance* **56(2)**: pp. 589–616.

Grinblatt, M. & M. Keloharju (2012): "IQ and stock market participation." *The Journal of Finance* **67(6)**: pp. 2121–2164.

Hausman, J. A. (1978): "Specification tests in econometrics." *Econometrica: Journal of the Econometric Society* pp. 1251–1271.

Kahneman, D., P. Slovic, & A. Tversky (editors) (1982): *Judgment Under Uncertainty: Heuristics and Biases.* Cambridge University Press.

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

Kalda, A., B. Loos, A. Previtero, & A. Hackethal (2021): "Smart(phone) investing? a within investor-time analysis of new technologies and trading behavior." SAFE Working Paper No. 303. https://ssrn.com/abstract=3765652 or http://dx.doi.org/10.2139/ssrn.3765652.

Korniotis, G. M. & A. Kumar (2011): "Do older investors make better investment decisions?" *Review of Economics and Statistics* **93(1)**: pp. 244–265.

Markowitz, H. (1952): "Portfolio selection." *The Journal of Finance* **7(1)**: pp. 77–91.

Nielsen, O. L. & A. M. Posselt (2021): "Betting on mean reversion in the vix? evidence from etp flows." SSRN Electronic Journal. https://ssrn.com/abstract=3529285 or http://dx.doi.org/10.2139/ssrn.3529285.

Odean, T. (2000): "Do Investors Trade Too Much?" *American Economic Review* **89(5)**: pp. 1279–1298.

Reiter-Gavish, H. *et al.* (2021): "Financial advice: Who Exactly Follows It?" *Research in Economics* **75(3)**: pp. 244–258.

Rizzi, J. V. (2008): "Behavioral bias of the financial crisis." *Journal of Applied Finance (Formerly Financial Practice and Education)* **18(2)**. Available at SSRN: https://ssrn.com/abstract=2698652.

Shapiro, S. S. & M. B. Wilk (1965): "An analysis of variance test for normality (complete samples)." *Biometrika* **52(3/4)**: pp. 591–611.

Shiller, R. J. (2000): *Irrational Exuberance.* Princeton, NJ: Princeton University Press.

Sirri, E. R. & P. Tufano (1998): "Costly search and mutual fund flows." *The Journal of Finance* **53(5)**: pp. 1589–1622.

Veld-Merkoulova, Y. V. (2011): "Investment horizon and portfolio choice of private investors." *International Review of Financial Analysis* **20(2)**: pp. 68–75.

Weisberg, S. (2005): *Applied Linear Regression.* New York: Wiley, 3 edition.

Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data.* MIT Press.