TALLINN UNIVERISTY OF TEHCNOLOGY
School of Economics and Business Administration
Department of Economics and Finance

Jegor Mitrofanski

THE EFFECT OF NEGATIVE EXPERIENCE ON EXPECTED RETURNS AND VOLATILITY: THE CASE OF ESTONIAN INDIVIDUAL INVESTORS

Master's thesis

Supervisor: Pavlo Illiashenko

Tallinn 2018

I declare that I have compiled the paper independently
and all works, important standpoints and data by other authors
have been properly referenced and the same paper
has not been previously been presented for grading.
The document length is 11 057 words from the introduction to the end of summary.

Jegor Mitrofanski……………………………

(signature, date)

Student code: 176695TARM

Student e-mail address: jegormitrofanski@gmail.com

Supervisor: Pavlo Illiashenko:

The paper conforms to requirements in force

……………………………………………

(signature, date)

Chairman of the Defence Committee:

Permitted to the defence

…………………………………

(name, signature, date)

TABLE OF CONTENTS

ABSTRACT

The objective of this master's thesis is to study if past negative experiences of Estonian individual investors have an effect on their expectations about future returns and volatility. Optimism is proxied with expected returns and overconfidence is proxied as expected volatility. As a secondary objective this thesis aims to provide evidence if randomly conditioning the subject to a certain optimistic or pessimistic historical past data will have an effect on their expectations about future returns and volatility. The data for the analysis is collected with a survey among Estonian individual investors and is analysed using ordinary least square method in statistical program Gretl. The results of this thesis can be split in two parts. Firstly, the analysis does not provide strong evidence that Estonian individual investors' future market expectations are affected by their negative past experiences. Neverthless, future return expectations show marginal statistical correlations with investors age and financial knowledge. Future volatility expectations show marginal statistical correlations with past experiences and overconfidence. Secondly, this thesis succeeds at proving evidence that manipulation of provided background infromation influences investor's level of optimism and overconfidence.

Keywords: Optimism, overconfidence, heuristics, experiences

INTRODUCTION

It has been 10 years since the last major financial crisis which resulted in a stock market crash. Prior to that we had experienced the dot-com bubble burst in the start of the century and Russian financial crisis in 1998. There was also European sovereign debt crisis in 2010, which mainly affected the Eurozone and particularly the PIIGS countries. Nevertheless, stock market returns have been historically very positive since the financial crisis of 2008-2009. Annex 1 shows how OMX Tallinn and OMX Baltic Benchmark GI have increased 370% and 382% respectively since market bottom in March 2009.

Since Estonia joined the EU in May 2004, local Gross Domestic Product has more than doubled from 7 124,60 to 17 463, 40 EUR in 2017 (Statistcs Estonia, 2017). Average net salaries have grown nearly 270% from 363 EUR in 2004 to 986 EUR in 2017. Improved economy and higher salary have led the people of Estonia in an even stronger pursuit of better life, higher income and more importantly – financial independence. Facebook group "Financial Independence" serves as great proof of this movement with more than 17 000 member joined as of April 2018. Also there are numerous other Facebook groups with participant numbers from 2000 – 6000 members, where people share their investing ideas. Among these are also riskier assets like cryptocurrencies and crowdfunding. In addition to these groups, approximately two dozen investment blogs are being followed daily (Rahajutud.ee) and Estonian's main financial newspaper has introduced Investor Toomas as a fictional character representing the investment ideas of the newspaper's stock-market journalists.

As approximately 58% of global Facebook users are aged 18-34 (Statisista, 2018), one might also assume that members of above mentioned financial freedom seeking groups fall in the same age category. This in turn means that more than half of financial freedom seekers were aged 8 – 24 years during the financial crisis of 2008-2009 and aged 1-7 years during the dot-com stock bubble burst. It is likely that some of them were not even born during the Russian financial crisis in 1998, which shrank Estonian export to Russia 3 times in half a year (Rei, 2009) and affected local economy in numerous ways.

As Estonian financial freedom seekers were potentially relatively young during the three mentioned economical events with global effects they might be lacking negative financial first or second hand experience. The study of Vissing-Jorgensen (2003) provides evidence that young inexperienced investors expected the biggest returns during the stock market rally of late 1990s.

Several works have investigated if personal experiences of macroeconomic events such as stock market crashes or financial crisis's affect investors risk attitude. For example, it has been found that individuals tend to overweight more recent returns than more distant ones, but distant experiences still have some impact on their risk tolerance (Malmendier & Nagel, 2011).

Likewise, Kaustia and Knüpfer (2008) have shown that there is a positive correlation between past IPO returns and future IPO subscriptions – the more positive have the returns been in previous IPO's, the more subscribers the next IPO's attract.

These studies have mainly focused on previous generations, however Nguyen Vinh Loc (2017) aims to distinguish investment risk profile of the millennial generation. He finds that millennials generally have low risk capacity and low risk need, also that they tend to be risk averse.

In their paper Barone-Adesi *et al*. (2013) find that during periods of stable market growth and low volatility, the representative investor is excessively optimistic and overconfident. During a crisis, he is pessimistic and under confident. During periods in which these biases are strong, the representative investor perceives risk and return to be negatively related, whereas objectively they are positively related.

It has been a sound ride of positive stock market returns since the last economic recession in 2008-2009. Likewise, it has been a decade of extremely low interest rates (Euribor fell sharply after the recession and has continued its decline and ended up in negative territory – Annex 2). Also the evolution of cryptocurrencies is putting investors risk-tolerance to a test.

As there might be a strong correlation between past experiences and current investment decisions there might be a broad number of young investors, who are extremely risk averse, potentially insufficiently educated and most importantly, have never been exposed to direct (personal experience) or indirect (friend or family member) effect of a financial crisis and very negative investment returns. Instead they have enjoyed a period of strong growth and very low interest rates (which has for example fuelled their real estate investments). These would most likely be investors who were at the minimum age of 14 and maximum age of 23 during the financial crisis (2008) and who currently are in the age of 24 – 33.

As Estonian financial freedom seekers were potentially relatively young during the three mentioned economical events with global effects they might be lacking negative financial first (or second) hand experience and have unrealistic stock market expectations.

This thesis aims to study how exposure to negative financial experience correlates with investors optimism and overconfidence. Optimism is proxied with expected returns and overconfidence is proxied as expected volatility. This approach follows a distinction between optimism and overconfidence suggested by Barone-Adesi *et al.* (2013).

Additionally, this thesis searches for evidence if manipulation of provided background information will influence investors levels of optimism and overconfidence. Based on previous findings in behavioral economics and psychology (framing effect, anchoring bias, and priming) it is expected that investors in optimistic condition will report higher level of optimism (higher expected returns) and higher level of overconfidence (lower level of expected volatility).

Therefore, two hypotheses are constructed:

Hypothesis H1:
H0: Past negative financial experience does not have an effect on expectations about future market returns (optimism) and volatility (overconfidence).
H1: Past negative financial experience does have an effect on expectations about future market returns (optimism) and volatility (overconfidence).

Hypothesis H2:
H0: Conditioning investors to an optimistic or a pessimistic past does not have an effect on their expectations about future market returns and volatility.
H1: Conditioning investors to an optimistic or a pessimistic past does not have an effect on their expectations about future market returns and volatility.

First part of this work focuses on literature overview of behavioral economics. The second part conducts an empirical analysis via survey among Estonian investors. Finally, the conclusion is made and additional suggestions for research are proposed.

The author would like to thank Pavlo Illiashenko and Cliona Georgia Dalberg for their contribution to this thesis.

1. LITERATURE OVERVIEW

1.1. Optimism

Optimism can be defined as the difference between expectations and actual reality. According to Sharot (2000), if expectations are better than reality, one might be considered optimistic and pessimistic if the reality expectations are inferior.

According to the findings of Puri and Robinson (2007), there is a relationship between an individual's level of optimism and their choice of portfolio. Also optimists tend to have higher savings than less optimistic people. Both Barber and Odean (2000) and Puri and Robinson (2007) findings confirm that individual stocks play a higher role in an optimist's investment portfolio. Puri and Robinson (2006) suggest that a moderate level of optimism can relate to better decision-making but also find that excessive optimism can be closely linked to overconfidence.

Carver *et al*. (2010) divide optimist and pessimists into two concrete types – people that expect good things and people that expect bad things. Scheier and Carver (1992) developed the LOT (Life Orientation Test) as a mean for measuring optimism by asking the respondents directly what are they expectancies about the future. While the initial LOT had one dimention, Carver *et al*. (1994) revised the test two years later by adding a second dimention – the respondent had to agree with optimistic statements and disagree with pessimistic ones in order to show his or her optimism.

Studies have shown that optimism can also be directly linked to two socio-economic variables – income and education. Moreover, the relationship between them is a two way street – Solberg *et al.* (2009) provide evidence that optimism may lead to better educational results and higher future income Segestrom (2007), however Heinonen *et al*. (2006) show that better socioeconomic conditions in a person's childhood lead to more optimism when they become adults (or vice-versa, bad coditions during childhood are more likely to raise pessimistic adults).

This thesis uses the LOT-R (Carver *et al*., 1994) test to measure a respondents level of optimism and gives a brief overview of this Likert scale type test in the second chapter.

1.2. Overconfidence

Fabre and Heude (2009) define overconfidence as an individual belief that he or she are able to achieve the desired result merely by believing in his or hers abilities. Although some researchers argue that optimism and overconfidence are closely related and appear frequently together (Taylor and Brown, 1988), others tend to distinguish them stating that confidence relates to a desirable outcome based on skills, while optimism is simply a positive outcome of external factors Malmendier and Tate (2005).

According to the findings of Larwood and Whittaker (1977), Svenson (1981), Alicke *et al.* (1995), when one must assess its skills, then usually he or she overestimates them compared to the mean of the results. Miller and Ross (1975) find that overconfident people tend to think that a positive outcome was achieved due to their actions and a negative outcome was a result of bad luck.

In their paper "The Trouble with Overconfidence" Moore and Healy (2008) distinguish three types of overconfidence: (1) overestimation of an individual's actual abilities, (2) overestimation of an individual's abilities compared to others, (3) excessive belief in his or hers abilities. They show that on difficult tasks people tend to think they do better than they actually do, but at the same time think they are worse than average. When it comes to easier tasks, people tend to be underconfident and think they are better in comparison with others.

As Moore and Healy (2008) provide evidence that all three types of overconfidence differ from one to other, then this work focuses primarily on the second type of overconfidence, which states that people think of their skills as superior compared to the median. There have been several major empricial studies about "better-than-average" effect. Zenger (1992) empirical study showed how 37% of managers placed themselves as top 5% performers in the firm. A typical "better-than-average" question is related to assesment of one's driving skills. This originates from Svenson (1981) work which showed that 93% of American drivers and 69% of Swedish drivers considered themselves to be better-than-average driver in their country.

1.3. Risk preference

According to Horcher (2005), risk is something necessary to take in order to follow an opportunity or seek a potential profit or gain. However one must differentiate risk and exposure, as the first one is related to loss probability and the latter to possibility.

Different economical researchers have studied what exactly affects an individual's tolerance and approach to risk. Modern portfolio theory pioneer Markowitz (1952) suggests that investors are risk averse by nature and if presented with two portfolios with similar expected returns, are more likely to choose the one with lower risk. Empirical study of of Weber *et al.* (2012) provide strong evidence that changes in one's risk attitude have an impact of his or her's expected return and volatility.

Past experiences is also one possible way of explaining risk attitude and is one of the key variables in the context of this study. Malmendier and Nagel (2011) suggest that higher past returns attract higher risk tolerance. Moreover, they specify that near-term experiences have a greater effect than those of further past and it is even more evident with younger investors.

The new study of Andersen (fourthcoming) analyses the impact of bank defaults in the financial crisis of 2008-2009 to risk preferences of Danish investors. The study focuses on personal experiences and divides the experiences into first-hand experiences, which arise from personal gains or losses, second-hand experiences which arise from gains or losses of close relatives or friends and third-hand experiences which reflects the locations of the defaulted banks. Andersen (fourthcoming) succeeds at proving that personally gained experiences make investors less risk tolerant. This thesis shares similar approach to differentiation of experiences and uses first-hand and second-hand experiences as main independant variables of interest.

This thesis follows the risk measurement steps of Malmendier and Nagel (2011) which are described in more detail in chapter two.

1.3. Heuristics - availability and representativeness

Pioneers in the research of heuristics were Tversky and Kahneman, who found that indivduals tend to adhere to certain mental shortcuts when taking decisions. They identified these shortcuts as heuristics and defined reprsentativness and availability.

In some situations these mental shortcuts might seem as an easy way out or and easy solution, but they can actually lead to a biased decisions. For example if one has to take a decision about either one will travel today from point A to B by airplane or by car, it can be affected by availability heuristic if there has been an airplane accidnet in the recent history. Althought statistically an airplane accident is far less likely to occur than a car accident, the person might consider taking the car, merely because they recall a recent event.

Another example from Tversky & Kahneman (1973) argue that an individual may asses a rate of divorce in his community based on recalling the divorces among his social network.

Khan *et al*. (2017) divide the availability heuristic in four distinct categories: retrievability, categorization, narrow range of experience, and resonance.

When it comes to investing, Barber and Odean (2002) show that investors tend to pick stocks that are more in the spotlight and frequently on the news. This may be caused by insufficent amount of time, resources, or knowledge. However, they prove that these stocks, affected by availability bias do not bring superior returns. In their follow up work, Barber and Odean (2008) find additional evidence to support this assumption.

An important finding from Lee *et al*. (2008) shows that there is a positive correlation between analysts' long term growth predictions and the economical cycle.

Kliger and Kudryavtsev (2010) test the availability heuristic on investor reactions to analyst recommendation revisions and find that there is a strong relation between the two.

Kahneman and Tversky (1972) show that in a process of decision making, an individual might overweight more recent information. In their study, a person named Linda was discribed as 31 years old, single, outspoken and very bright, having a major in philosophy, having concerns about discrimination and social justice, attending anti-nuclear demonstrations during her studies in university. After this brief introduction, respondents were asked if Linda is more likely to be (A+B) a bank teller and active in a feminist movement or (B) a bank teller. The so called "Linda problem" is a good example of conjuction fallacy, where option (A+B) has actually a conjuction of two events, which can not be more probable than just one event: $P(A+B) \leq P(B)$ (Tversky and Kahneman, 1983).

According to Kahneman and Tversky (1972), representativeness means that in situations of uncertainty, people *"evaluate the probability of an uncertain event, or sample, by the degree to which it is: (i) similar in essential properties to its parent population; and (ii) reflects the salient features of the process by which it is generated"*.

In investing, representativeness heuristic can be considered as a psychological bias that causes the investors to extrapolate excellent performance from the past to the future. Barberis *et al*. (1998) model shows that investor with a representativeness herustic might overreact to a string of similar information.

2. DATA AND METHODOLOGY

2.1. Data collection

This work evaluates how different variables are related to an investors expectation for future returns and volatility based on the example of Estonian investors. To obtain the data a survey has been conducted over the internet among Estonian investors. The survey was distributed in different Estonian investing related groups on Facebook, LHV bank forum and via other channels:

- "Financial freedom group" on Facebook (17 000+ members)
- "Female investors club" on Facebook (10 000+ members)
- Rahajutud.ee Facebook group (2000+ members) and website (1500+ unique visitors daily)
- Estonian cryptocurrency group (5000+ members)
- LHV bank forum
- Äripäev and Investor Toomas editorial board
- Other channels (personal contacts and social network of the author)

For this study, an assumption is made, that a typical investor of interest is aged between 18 to 63 years. Several authors like Heaton and Lucas (2000) and Yoo (1994) have shown that the share of riskier assets like stocks increases until the age of retirement and starts to decrease or maintain its share afterwards (Poterba and Samwick, 2001; Heaton and Lukas 2000). As the age of retirement in Estonia is currently from 63 to 65 depending on the year of birth, then there is a maximum target population size of 780 000 people according to Estonian Statistics.

Furthermore, according to Estonian Statistics, about 38% of people aged 18 – 74 in Estonia save money. This reduces the maximum target population size to approximately 300 000. However, according to a survey conducted by Swedbank and Estonian Institute of Economic Research in March 2017, only up to 20% of people who make savings, do that for retirement or investing. This brings the maximum number of potential investors to about 60 000 people. As there is no information about how many of these people who save money are actually investing in order to grow their savings we assume that there are approximately 30 000 – 60 000 active investors in Estonia. According to the data from Nasdaq Baltic stock exchange there are currently over 23 700 investment accounts registered in the stock exchange, however as some individuals may have several business accounts, the actual number of individuals might be less. Although it is impossible to know how many individuals have investments with foreign brokers or how many crypto investors there are in Estonia, we can assume, that our estimated target population size is fair.

In order to draw conclusions on a 95% confidence level with 6% interval the survey required a minimum of 264 answers. The survey was conducted during three days in April 2018 and reached 953 people, from which 565 people started replying to the survey. There was 199 drop-outs and thus with a completion rate of 64,78%, the survey collected 366 completed sets of answers. To guarantee that the sample would comprise of 100% Estonian investors, the survey was conducted in Estonian language. Following the data collection, a preliminary analysis of the collected data was made and additional 30 answers were dropped which in the last two questions of the survey had not respected the numerical answer requirement and/or had extreme values such as 100% volatility.

The survey was divided into five main parts as presented in Figure 1:

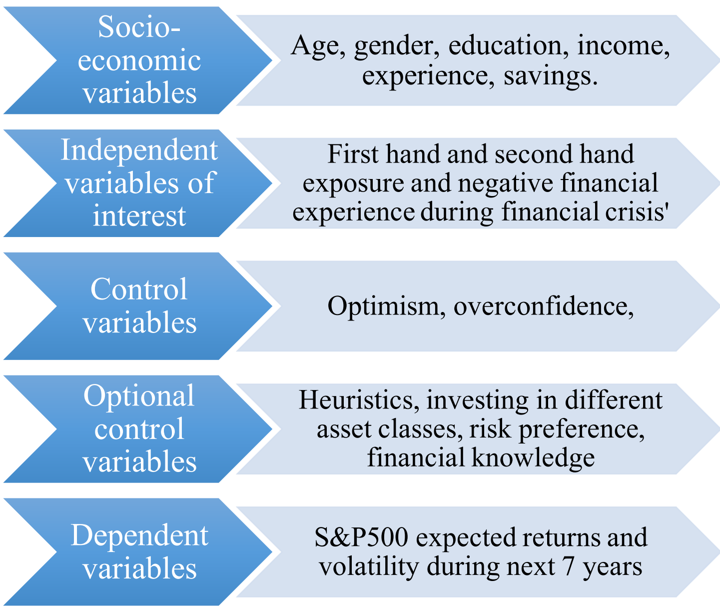


Figure 1. Five main parts of the survey conducted in April 2018
Source: Author

2.2. Dependent variables: optimism and overconfidence

Dependent variables in this thesis are expected market return and expected market volatility. Easton and Sommers (2007) provide evidence that there exists a correlation between level of optimism and implied expected returns. Likewise, eempirical findings of Jlassi *et al.* (2014) suggest that excessive volatility can be explained by overconfidence.

In the survey conducted for this master's thesis the respondents were randomly presented with two different scenarios of S&P500 historical returns. The pessimistic condition told investors: "The average annual rate of return of the S&P 500 for the period of 2000-2017 was 4,2%.", while the optimistic condition stated: "The average annual rate of return of the S&P 500 for the period of 1977-2017 was 8,7%.". After being exposed to one or the other condition, respondents were then asked to state what would they expect the S&P500 returns to be in the following 7 years. The mean result for both conditions was 5,23%, however the respondents with pessimistic condition had a mean expected return of 4,25% while the same figure with optimistic condition was 6,21%.

After being conditioned to an optimistic or pessimistic historical past returns, the respondent was presented with an equivalent condition about S&P500 volatility. The pessimistic condition stated "The Standard deviation of S&P500 for the period of 2000-2017 was 17,5%" and the optimistic "The Standard deviation of S&P500 for the period of 1977-2017 was 15,7%". The mean expected return volatility for the next 7 years was 18,26% and 16,56% respectively.

2.3. Negative financial experience and a set of controls

This sub-chapter gives an overview about main independent variables of interest, followed by a set of control variables used in this thesis. There are two main independent variables of interest which aim to provide reasoning for the two hypothesis presented.

First hypothesis (H1) states that lack of negative financial experience leads to more optimistic expectations about future returns and higher overconfidence which is proxied by lower volatility expectations. The survey tested for direct (first hand) and indirect (second hand) financial exposure and negative financial experiences. Regarding exposure, the respondents were asked if they were holding an investment portfolio during the Russian financial crisis of 1998, stock market crash of 2000/2001 and/or financial crisis of 2008/2009. Their exposure was given a binary coding of yes

(1) and no (0). Same question was also asked about their close friends and relatives to measure for indirect exposure. Regarding negative financial experiences, the respondents were asked if they had suffered any negative financial experience such as stock loss or job loss during or immediately after the Russian financial crisis of 1998, stock market crash of 2000/2001 and/or financial crisis of 2008/2009. Their financial losses experience was given a binary coding of yes (1) and no (0). Same question was also asked in regards to their close friends and relatives to measure for indirect negative financial losses.

It is eminent from the collected data 68% of all validated respondents had experienced a direct negative financial loss at the time of the above mentioned crises. However only 28% of the respondents owned a portfolio during any of these crises and 49% of respondents were exposed via their personal or close relative portfolio. These numbers seem realistic as the average age of the respondent was 34,5 years.

Although the survey measured in detail the levels of exposure and financial loss during the crises, the detailed variables did not prove to be significantly important during first regression analysis and two first level general variables created: first one included both direct and indirect exposure (variable name: all_exposure) and second one both direct and indirect losses (variable name: all_losses). Once again, neither of the two first level general variables proved to be significant, but as their p-value decreased, the author decided to pool both first level general variables together into one second level general variables (variable name: all_experience). As we can see from the third chapter, this proved to be efficient as this variable became significant in one of the suggested model specifications.

Second hypothesis (H2) states that conditioning a respondent to a certain condition will have an immediate effect on his or her response. The respondents were presented with two different conditions using a randomization technique (variable: condition). One part was presented with a condition that was optimistic and long term – respondents were told that the S&P 500 annual return over a period of 40 years (1977 – 2017) was 8,7%. The other part was pessimistic and short-term – respondents were given annual returns over 17 years (2000 – 2017) which was 4,2%. With either condition, the respondents were asked what would they expect the annual return of the S&P500 to be in the following 7 years. The mean expected return was 5.3% with a standard deviation of 2,7%. However, it is remarkable how conditioning the respondents affected their response. The mean expected return for respondents who were presented with an optimistic and longer term past was

6,2% and 4,3% for those who were presented with a pessimistic and shorter term past performance data.

Based on the randomly selected past scenario of the question about returns, the respondents were automatically taken to a question about volatility with the same scenario. The optimistic condition in this case was the volatility of S&P500 during the period of 40 years (1977 – 2017) which was 15,7% and the pessimistic condition where the S&P500 volatility was 17,5% during the period of 17 years (2000 – 2017). The mean expected volatility was 17,2 %, with a standard deviation of 5,6%. Although the differences here are not as big as in the case of returns, here once again conditioning seems to have an immediate effect, as the optimistically conditioned respondents had a mean excepted volatility of 16,6% and the pessimistically conditioned had a mean of 17,8%.

In order to test the relationship between above mentioned independent variables and the dependent variables mentioned in previous subchapter, a set of control variables was created and gathered from the respondents. The set of control variables comprised of variables for optimism and overconfidence, socio-economic factors, availability and representativeness heuristics, risk preference, financial knowledge and portfolio composition.

General level of optimism as a control variable was measured via a revised LOT-R test (Carver *et al.*, 1994). Respondents were asked to evaluate their own feelings on a scale from A (I agree a lot) to E (I disagree a lot), answering 10 questions (Appendix 3), from which 4 questions were filler questions, 3 questions were optimistic and 3 questions were pessimistic or reversed optimistic. The answers were coded on a scale from 0 (E) to 4 (A) and an overall score was calculated. LOT-R test does not have a benchmark, but the higher the total value, the more optimistic is the respondent. According to Carver *et al.* (1994) an average person's score is between 14 and 15 points (Table 3). The mean of the general level of optimism of the respondents of the survey for this thesis was 15.58 and standard deviation 3.46, which indicates that the average respondent was generally slightly more optimistic than an average person. It is interesting to note, that in the case of Carver *et al.* (1994), the female respondents of current survey were also more optimistic averaging 16.65 points compared to 15.32 points of men.

Group	Average on the LOT-R
Students	14.33
Female students	14.42
Male students	14.28
Bypass patients	15.16
Female bypass patients	14.92
Male bypass patients	15.24

Table 1. Average LOT-R score in Carver *et al*. (1994) research
Source: Carver *et al*. (1994)

General level of overconfidence (variable gen_overconf) was measured via self-assessment better-than-average questions about one's abilities. The respondents were asked to assess on a 1-5 point Likert scale if they disagree (1) or agree (5) that they are better than average driver, more informed, get along better with others and have better leadership skills than an average person. The question about driving was taken from Svesson (1981), while the other questions were created by the author. The mean value for all the respondents is 2,9 with a standard deviation of 0,7, which indicates that an average respondent considers himself as slightly better than average. Here also the difference in gender is present – while female respondents considered themselves with a score of 2,7 more towards average, then male respondents had an average score of 3,0. The survey distinguished general and financial overconfidence (variable fin_oveconf) and tested for the latter with a similar self-assessment approach asking the respondents if they consider themselves having better than average returns, set of investment skills, investment experience and if they consider themselves to be better informed than an average investor. The mean value for all the respondents is 2,6 with a standard deviation of 0,7, which indicates that an average respondent considers his or her investment abilities as more or less average. Likewise, the difference in gender is also present here with female investors considering themselves modestly being below average with a mean of 2,4.

The survey also determined if several socio-economic variables might have an effect on an investor's return and volatility expectations. The respondents were asked to indicate their age, gender, level of education, net monthly income, savings which are not part of the portfolio and investment experience in years. The average age (variable: age) of the respondents was 34,5 years, with the youngest being 17 years (one respondent) and the oldest 63 years of age (two respondents). This relates well with an assumption made in subchapter 2.1. where the author assumes that the range of investors age is from 18 to 63 years. It is eminent from the survey data that investor's with above average age expected an average S&P500 return of 4,9% while the younger investors were more optimistic and expected 5,5%. Furthermore, conditioning to optimistic or pessimistic past among young investors had a lower mean difference (1,7%) than among the older than mean respondents (2,0%).

Although communities like "Female Investors Club" attract females to start investing their representation in this study remains faint as only 68 respondents were female representing a 20% of the total sample.

Regarding their level of education (variable: education), the respondents had a multiple choice question with eight different values to indicate their education. As this wide distribution provided too much noise in the empirical part of this work, the levels were consolidated into three major levels. There were 47 respondents who did not have a university degree, 165 respondents who had a bachelor's degree or an unfinished master's degree and 125 respondents with a master's degree or higher. The mean value of education was 2,3 which also indicates that majority of respondents were highly educated people. When asked about their level of net monthly active and passive income (variable income), the survey gave the respondents 6 different levels which proved inefficient for the empirical part of this work. Thus, the levels were consolidated into three levels with 132 respondents belonging to level 1 (monthly net income until 1500 EUR), 152 respondents belonging to level 2 (monthly net income 1501 – 3000 EUR) and 52 respondents in level 3 (monthly net income over 3000 EUR). The mean category was 1,76 with a standard deviation of 0,70.

While people might be hesitant to provide accurate data about their income an additional variable measuring savings (variable: savings) was introduced. The respondents were asked to indicate how many months' worth of savings they have in their emergency fund apart from their investment portfolio. The survey offered initially 6 levels: no emergency fund, 1-2 months, 3-4 months, 5-6 months and more than 6 months, which were later consolidated by the author to 3 levels. A total of 59 respondents did not have any emergency savings apart from their investment portfolio, 168 respondents had savings that would last the 1-4 months and 109 respondents would have emergency fund savings that would last longer than 5 months.

Finally, the respondents were asked also to indicate their investing experience (variable: experience) in years which had a mean of 7,7 years with a standard deviation of 6,3 years. This might indicate well why only 28% of the respondents had direct exposure during the crisis – the average respondent only started investing at the age of 26,8 years and did that in 2011 – well after the last major financial downturn. Moreover, this average 34,5-year-old investor would have been around 17-years old during the dot-com bubble and 15-year-old during the Russian financial crisis, which might have given him or her a second-hand exposure or losses, which they would still remember a while later.

Risk preference was measured via standard self-reported question about risk preferences. According to Dohmen, *et al.*, (2011) and Treibich (2015), the self-reported questions provide reliable information about an individual's risk aversion. Results of the survey are coded from 1 (not willing to take any financial risk) to 4 (willing to take significant financial risk, in order to receive substantial returns).

Financial knowledge is also tested with self-reported questions about knowledge of financial mathematics, stock markets, portfolio theory, peer to peer lending, block chain technology. The answers were presented on a traditional five level Likert scale ranging from strongly disagree to strongly agree. The mean result was 2.63 with a standard deviation of 0.70.

Representative and availability heuristic were also measured during the survey. Representative heuristic was measured via two questions, one of them was so-called "Linda problem" from Tversky and Kahneman (1983). Respondents were presented with background information about Linda's youth and asked if it was more likely that Linda is a bank teller or a bank teller who is also a feminist. About 46% of the respondents were affected by the representative heuristic and chose the second, less probable option. Second question about representativness was taken from Khan *et al.* (2017) where the respondents was asked to state on a Likert scale if they agree with a statement that taken into account that IPO stocks are either (A) succesful long-term investments or (B) stocks that fail as succsful long-term investments, the IPO of stock ABC is most likely (A) because IPO-s are a good long term investment. Availability heuristic was tested with a similar question on a

Likert scale where respondents were asked if they agree that they are likely to purchase a stock, if their friend who is very good at stock picking would suggest to do that. The result was 2.7 on a scale from 1 to 5.

Finally respondents were also asked to state what type of assets and in which approximate proportion are held in their investment portfolios. Proposed assets for assesment were stocks, cryptocurrencies, real estate, peer-to-peer lending, start-up equity and fixed income securities. Althought the respondents were asked to indicate the specific share of the specific asset, this study gives the data a binary coding of (not in portfolio(0) and present in portfolio(1).

2.3. Model specification

The data in the previous subchapter does not give a definitive answer if first or second hand exposure to financial crisis's and negative financial experiences affect investor's expectations about returns and volatility.

In order to conduct a regression analysis, an econometrical model needs to be constructed. As there are several variables of interest, the model constructed needs to be a multiple linear regression model.

The following general econometric model has been constructed:

$$Y_i = f\left(Controls_i; DV_{Controls_i}; opt_{Controls_i}; Exposure_i; \right),$$

where i subscript designate a respondent. Y_i is a measure of either optimism or overconfidence proxied by expected returns and expected volatility correspondingly. $Controls_i$ refers to a set of standard socio-economic controls such as age, gender, education and savings. $DV\_Controls_i$ include general optimism and general overconfidence. $opt\_Controls_i$ refers to a set of optional controls such as availability and representativeness heuristics, financial knowledge, risk preferences, investments in different asset classes. Whereas, $Exposure_i$ include measures of direct and indirect exposure to negative financial experience and condition.

The following variables were not included in the regression:

Variable name	Definition	Regression choice
DEPENDENT VARIABLES		
returns	Numerical value of expected returns	Used in regression as main dependent variable
volatility	Numerical value of expected returns	Used in regression as main dependent variable
NEGATIVE FINANCIAL EXPERIENCE AND SET OF CONTROLS		
age	Numerical value of respondents age	Used in regression, as age is significant
gender	Binary coding: 0 – male 1 – female	Used in regression as a socio-economic variable
education	Coding: 1 – Up to unfinished BSc degree 2 – Finished BSc until unfinished MSc degree 3 – Finished MSc and higher	Used in regression as a measure of respondent's general level of education.
savings	Coding: 1 – no savings 2 – One to four months' savings 3 – More than five months' savings	Used in regression as a measure of respondent's financial strength.
optimism	LOT-R test numerical value: Highest possible: 24 Lowest possible: 6	Used in regression, as optimism is proxied via dependent variable returns.
gen_overconf	Numerical value on 1-5 Likert scale	Used in regression, as overconfidence is proxied via dependent variable returns.
condition	Randomly selected by survey program. Binary coding: 0 – optimistic 1 – pessimistic	Used in regression to test H2
all_experience	All negative financial experience and exposure grouped together. Binary coding: 0 – no experience 1 – negative experience	Used in regression to test H1
finlit	Numerical value on 1-5 Likert scale of self-assessed financial knowledge	Used in regression as important control variable

repr_heur	Numerical value on 1-5 Likert scale of representativeness heuristic	Used in regression as important control variable
availab_heur	Numerical value on 1-5 Likert scale of availability heuristic	Used in regression as important control variable
stocks	Binary coding: 0 – portfolio does not contain stocks 1 – portfolio contains stocks	As stocks are considered riskier assets, then stocks are used as a measure of risk in the regression.
crypto	Binary coding: 0 – portfolio does not contain crypto currencies 1 – portfolio contains crypto currencies	As cryptocurrencies are considered riskier assets, then crypto currencies are used as a measure of extreme risk in the regression.

Table 2. Dependent, independent and control variables included in regression

Source: author

The following variables were not included in the regression:

Variable name	Definition	Regression choice
DEPENDENT VARIABLES		
returns2	Log-transformed dependant variable returns	Left out from regression, as log-transformed variable did not improve the regression, but excluded some significant variables.
volatility2	Log-transformed dependant variable volatility	Left out from regressioon, as log-transformed variable did not improve the regression, but excluded some significant variables.
NEGATIVE FINANCIAL EXPERIENCE AND SET OF CONTROLS		
all_exposure	Both indirect and direct exposure grouped together. Binary coding: 0 – no exposure 1 – exposure	Left out from regressioon, as regression analysis did not find any statistical significance with specified variable. Instead, a general variable all_experience was used, which includes both negatiive financial experience and losses during crises.
all_losses	Both indirect and direct losses grouped together. Binary coding: 0 – no direct losses 1 – direct losses	Left out from regressioon, as regression analysis did not find any statistical significance with specified variable. Instead, a general variable all_experience was used, which includes both negatiive financial experience and losses during crises.

direct_exposure	Binary coding: 0 – no direct exposure 1 – direct exposure	Left out from regressioon, as regression analysis did not find any statistical significance with specified variable. Instead, a general variable all_experience was used, which includes both negatiive financial experience and losses during crises.
Indirect_exposure	Binary coding: 0 – no indirect exposure 1 – indirect exposure	Left out from regressioon, as regression analysis did not find any statistical significance with specified variable. Instead, a general variable all_experience was used, which includes both negatiive financial experience and losses during crises.
direct_losses	Binary coding: 0 – no direct losses 1 – direct losses	Left out from regressioon, as regression analysis did not find any statistical significance with specified variable. Instead, a general variable all_experience was used, which includes both negatiive financial experience and losses during crises.
Indirect_losses	Binary coding: 0 – no indirect losses 1 – indirect losses	Left out from regressioon, as regression analysis did not find any statistical significance with specified variable. Instead, a general variable all_experience was used, which includes both negatiive financial experience and losses during crises.
income	Coding: 1 – Up to 1 500 EUR net monthly income 2 – 1501 – 3000 EUR net monthly income 3 – Over 3000 EUR net monthly income	Left out from regression, as it is highly correlated with savings and people might be more honest about savings than income.
fin_overconf	Numerical value on 1-5 Likert scale	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.
risk	Numerical value on 1-5 Likert scale	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.
finlit2	Binary value: 0 – incorrect answer 1 – correct answer	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.
linda	Binary value: 0 – incorrect answer 1 – correct answer	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.

r_estate	Binary coding: 0 – portfolio does not contain real estate 1 – portfolio contains real estate	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.
p2p	Binary coding: 0 – portfolio does not contain p2p investments 1 – portfolio contains p2p investments	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.
start_ups	Binary coding: 0 – portfolio does not contain start-up equity investments 1 – portfolio contains start-up equity investments	Left out from regression, as it did not prove significance in any regression. Very high p-values indicated no correct fitting in the regression.

Table 3. Dependent, independent and control variables not included in regression

Source: author

3. DATA AND METHODOLOGY

These variables were then divided into variables that were important for the construction of the basic model specification (discussed in detail in Chapter 3.1.) and optional control variables which were added to the specification.

In order to test the two hypotheses six different econometric specifications were constructed – three for dependent variable returns and three for dependent variable volatility.

For both dependent variables the author started with constructing a basic model, which would include a constant, selected socio-economic variables and a control variable for the dependent variable.

As such, the following basic specification (Model R_Basic) was constructed for variable returns:

$$Y_i(returns) = f\big(age, gender, education, savings)_i; optimism_i;\ opt_{Controls_i};\big)$$

where optimism acts as control variable for dependent returns and socio-economic variables act as control variables (Annex 5).

For dependent variable volatility the following basic specification (Model V_Basic) was constructed:

$$Y_i(volatility) = f(age, gender, education, savings)_i; overconfidence_i;)$$

where general overconfidence acts as control variable for dependent returns, socio-economic variables act as control variables (Annex 8).

Although both basic specifications had statistically significant variables, their descriptive abilities were poor and their p-values high (Chapter 3.1.). In the following step, main independent variables of interest were introduced in order to evaluate if null hypothesis of H1 and H2 can be rejected and if specifications would become significant.

For this variables condition and all_experience were introduced in the equation for both dependent variables and specifications Model R_H1+H2 and Model V_H1+H2 were constructed:

$$Y_i(returns) = f\begin{pmatrix} age, gender, education, savings)_i; optimism_i; \\ opt_{Controls_i}; condition_i; experience_i \end{pmatrix}$$

$$Y_i(volatility) = f\begin{pmatrix} age, gender, education, savings)_i; overconfidence_i; \\ opt_{Controls_i}; condition_i; experience_i \end{pmatrix}$$

As a final step five optional control variables were added to the specification: availability and representativeness heuristics, crypto currency and stock investments and financial knowledge.

Model R_H1+H2+OC and model V_H1+H2+OC were constructed as follows:

$$Y_i(returns) = f\begin{pmatrix} age, gender, education, savings)_i; optimism_i; \\ opt_{Controls_i}; condition_i; experience_i \end{pmatrix}$$

$$Y_i(volatility) = f\begin{pmatrix} age, gender, education, savings)_i; overconfidence_i; \\ opt_{Controls_i}; condition_i; experience_i \end{pmatrix}$$

As the third model for volatility which included the optional controls became statistically insignificant, a fourth model V_H1+H2+OC_2.0 was constructed without heuristics.

The seven regressions are constructed as shown in the Table 4:

Model name	R_BASIC	R_H1+H2	R_H1+H2+OC	V_BASIC	V_H1+H2	V_H1+H2+OC	V_H1+H2+OC_2.0
Dependent	Returns	Returns	Returns	Volatility	Volatility	Volatility	Volatility
const	6.245*** (1.071)	6.965*** (1.029)	8.147*** (1.292)	13.072*** (2.140)	12.113*** (2.166)	13.331*** (2.577)	13.367*** (2.379)
age	−0.034* (0.018)	−0.030* (0.017)	−0.032* (0.018)	−0.003 (0.037)	−0.014 (0.037)	−0.025 (0.039)	-0.026 (0.038)
gender	0.208 (0.374)	0.151 (0.353)	0.0480 (0.357)	0.272 (0.775)	0.317 (0.770)	0.145 (0.781)	0.160 (0.776)
education	−0.064 (0.222)	0.042 (0.210)	0.053 (0.213)	0.780* (0.460)	0.735 (0.460)	0.713 (0.466)	0.704 (0.464)
savings	−0.313 (0.219)	−0.221 (0.207)	−0.115 (0.212)	−0.125 (0.455)	−0.209 (0.453)	−0.139 (0.465)	-0.121 (0.461)
optimism	0.063 (0.043)	0.067 (0.041)	0.063 (0.041)				
gen_overconf				0.922** (0.442)	0.887** (0.441)	1.022** (0.455)	1.014** (0.447)
condition		−1.787*** (0.281)	−1.876*** (0.284)		1.168* (0.615)	1.059* (0.623)	1.062* (0.621)
all_experience		−0.549 (0.400)	−0.452 (0.403)		1.35331 (0.876)	1.495* (0.886)	1.515* (0.881)
repr_heur			−0.058 (0.158)			0.094 (0.345)	
availab_heur			0.085 (0.136)			-0.084 (0.301)	
crypto			−0.372 (0.213)			-0.622 (0.820)	-0.639 (0.811)
stocks			−0.378 (0.374)			0.674 (1.003)	0.690 (0.461)
finlit			−0.372* (0.213)			-0.718 (0.475)	-0.705 (0.468)
Observations	336	336	336	336	336	336	336
Adjusted R^2	0.015	0.124	0.129	0.006	0.019	0.015	0.021
F Statistic	0.072*	1.01e-08***	8.52e-08***	0.219	0.067*	0.147	0.075*
Note: *p<0.1; **p<0.05; ***p<0.01							

Table 4. Model specifications in regression

Source: author’s calculations

3.1. Empirical results

After conducting an OLS (Ordinary Least Squares) regression analysis to a sample of 336 surveys, it is evident that empirical model's adjusted R2 was approximately around 0,12 – 0.13 or 12% – 13% for models with dependent variable return and 0.01 – 0.02 or 1% – 3% for dependent variable volatility, which shows how much of returns variability is explained with the current model. If usually adjusted R2 is considered to be important and implies a value from 70%, then in some occasions adjusted R square can be misleading. For example, in the fields of social sciences and arts a low R2 of 10% or lower can generally be accepted, as human behaviour is difficult to predict.

All models with expected return as dependent variable were significant at a level of alpha at least 0.1. However, the model with volatility as dependent became only significant after independent variables of interest were added and kept its statistical significance only with a selection of optional controls.

Model R_Basic (Annex 5) is comprised of a constant, four socio-economic variables (age, gender, education and savings) and optimism which serves as control variable for dependent variable returns. The models adjusted R^2 is at a low level of 0.015 which indicates that the basic specification for dependent variable returns is quite weak to explain returns variability. Nevertheless, the specification itself proves to be statistically significant on a level of alpha of 0.1. Variable age has a p-value of 0.052 and is significant on a level of 0.1 and the constant has a p-value of <0.0001 and is significant on the level of 0.01. Although the other variables are statistically not significant, variables savings and optimism have p-values of 0.155 and 0.149 respectively. This specification allows us to conclude that older people tend to have lower return expectations compared to younger people.

Model R_H1+H2 (Annex 6) is comprised of a constant, four socio-economic variables (age, gender, education and savings), optimism, which serves as control variable for dependent variable returns, and main independent variables of interest – condition and past experiences. The models adjusted R^2 is much higher than for the basic model and is at the level of 0.124 (or 12,4%). The specification also improved in terms of statistical significance, which is now at 0.01. Variable constant remained also significant on the level of 0.01 and likewise variable age remained significant at 0.1. This is a good evidence that age plays a significant role in an Estonian investor's return preference. Although the variable *all_experience* has a rather low p-value of 0.170, it is not statistically significant. Most importantly this model provides strong evidence that the null hypothesis for H2 can be rejected and we can assume, that conditioning has played a significant role to investors expected return. A coefficient of -1.79 implies that the investor who was pessimistically conditioned had an expected return on average 1,79% lower compared to the investor with pessimistic condition.

Model R_H1+H2+OC (Annex 7) is the third and final model with dependent variable returns. Optional control variables of representativeness and availability heuristics, stocks and crypto currencies and financial knowledge are added to the specification. The model itself slightly improves once again, as the adjusted R^2 raises slightly to 0.129 (or 12,9%). Model itself remains significant on the level of 0.01 as the p-value actually decreases further. The constant and variable age remain significant, on the level of 0.01 and 0.1 respectively. Likewise, the variable condition remains statically important on the level of 0.01 which allows us to conclude that the null hypothesis for H2 can be rejected and state that randomly conditioning the respondents with an optimistic or pessimistic past performance of S&P500 has a direct effect on their expected market returns for the future. Also, one of the optional control variables, financial knowledge, becomes statistically significant on the level of 0.1. We might conclude that more sophisticated people will have lower expectations for future returns.

Although the first three models with dependent variable returns manage to provide evidence to support rejecting the null hypothesis of H2, they do not provide evidence about H1 as general negative financial experience remains insignificant in both models R_H1+H2 and R_H1+H2+OC.

Model V_Basic (Annex 8) is comprised of a constant, four socio-economic variables (age, gender, education and savings) and optimism which serves as control variable for dependent variable returns. The models adjusted R^2 is at an extremely low level of 0.006 which indicates that the basic specification for dependent variable returns is very weak to explain returns variability. Although this specification has three statistically significant variables – constant at 0.01, general overconfidence at 0.05 and education at 0.1, we are in no position to draw conclusions from this model, as the model itself is statistically insignificant.

Model V_H1+H2 (Annex 9) is comprised of a constant, four socio-economic variables (age, gender, education and savings), general overconfidence, which serves as control variable for dependent variable volatility, and main independent variables of interest – condition and past experiences. The models adjusted R^2 is higher than for the basic model and is at the level of 0.019 (or 1,9%). More importantly, the specification improved in general and became statistically significant on the level of 0.1. Constant has a coefficient of 12.11 and a standard deviation of 2.17 and is significant of the level of 0.01. General overconfidence has a coefficient of 0.887, a standard error of 0.441 and with p-value of 0.045 is statistically significant on the level of 0.05. Variable condition has a coefficient of 1.168, a standard deviation of 0.614 and p-value of 0.058 which makes is significant on the level of 0.1 Unfortunately negative financial experience variable has a p-value of 0.123 and is not significant. As the specification's F-statistic is statistically significant we can conclude that this model also supports rejecting H2 as conditioning seems to make a difference of nearly 1,2% in volatility expectations. Although not in either hypothesis, we can also conclude that more overconfident people expect higher volatility – this may be true if we would assume that overconfident people make riskier investments and are therefore expecting higher volatility.

Model V_H1+H2+OC (Annex 10) is the third model with dependent variable volatility. The model is comprised of a constant, four socio-economic variables (age, gender, education and savings), general overconfidence, which serves as control variable for dependent variable volatility, main independent variables of interest – condition and past experiences and also a set of control variables, such as representativeness and availability heuristics, stocks and crypto currencies and financial knowledge. The model itself worsens, as the adjusted R^2 raises decreases to 0.015 (or 1,5%). Also the model itself becomes statistically insignificant because the p-value of F-Statistic is 0.147. However, we can see from this model that constant has a coefficient of 13.331, standard error 2.577 and is itself still significant on the level of 0.01. Likewise, variables general overconfidence and condition remain significant and in addition our main variable of interest becomes statically significant on the level of 0.1 with p-value of 0.092. All other control variables like representativeness and availability heuristics, stocks and cryptocurrencies are insignificant with very high p-values and financial knowledge is insignificant with p-value of 0.132. This specification itself is statistically insignificant, which does not allow the author to make conclusions based on this model. However, as the main variable of interest (all_experience) becomes significant, the author constructs one additional model (V_H1+H2+OC_2.0) for volatility to see, if there might be a model where experience matters statistically.

Model V_H1+H2+OC_2.0 (Annex 11) plays an important role for this study. The model is comprised of a constant, four socio-economic variables (age, gender, education and savings), general overconfidence, which serves as control variable for dependent variable volatility, and main independent variables of interest – condition and past experiences and optional control variables from previous model (V_H1+H2+OC). However, compared to the previous model, two optional control variables with highest p-value are omitted from the specification. These omitted variables are representativeness and availability heuristics, with p-values of 0.786 and 0.779 respectively. The models adjusted R^2 is higher than for any models with dependent variable volatility and is at the level of 0.021 (or 2,1%). More importantly, the specification improved in general and became statistically significant on the level of 0.1 (p-value is 0.075). Constant has a coefficient of 13.367 and a standard deviation of 2.379 and is significant of the level of 0.01. The statistical significance of the constant remains on the same level of 0.01. General overconfidence has a coefficient of 1.014, a standard error of 0.447 and with p-value of 0.024 is statistically significant on the level of 0.05. Variable condition has a coefficient of 1.062, a standard deviation of 0.621 and p-value of 0.088 which makes is significant on the level of 0.1. Most important finding of this specification is however main independent variable of interest, negative financial experience, has a coefficient of 1.515, a standard error of 0.881 and has a p-value of 0.087 which makes is statistically important on the level of 0.1. as the specification itself becomes significant we can conclude that the null hypothesis for H2 remains rejected and the conditioning has an effect also on respondent's expectations about volatility. But more importantly it is now possible to think about partially rejecting the null hypothesis for H1, as variable all_experiences became significant. Based on this last model we can conclude, that optimistic past scenario increases respondent's volatility expectations by about 1,06%. It is also possible that more overconfident people expect higher volatility (as they might be invested in riskier assets as proposed with model V_H1+H2). And finally, there might be a chance to conclude that people with negative financial experiences do expect more turbulence in the future and even disregard the optimistic past (as the coefficient for all_experience is almost 50% higher).

3.2. Discussion

The survey data analysis has shown, that indirect and direct exposure and negative financial experiences do not affect Estonian investors' expectations about future returns and volatility as specific variables. Thus a general negative financial experience variable (all_experience) was created.

Initially, the data analysis creates three models for each of the dependent variables. First model is a basic specification with a total of 6 variables: constant, 4 socio-economic variables and a control variable for the representative dependent variable. In case of expected returns, the control variable was optimism and in case of volatility the control variable was general overconfidence.

Both basic specifications have a very low adjusted R^2 and only basic specification for return is statistically significant. Thus the basic specifications only show that age might have some effect on the expected returns.

In the second stage, two main independent variables of interest were added to the basic specifications. Variable condition indicated if the respondent was conditioned to an optimistic or a pessimistic past and is part of hypothesis H1. General negative financial experience variable indicated if the respondent had suffered any negative financial experience and is part of hypothesis H1.

Both models improve significantly (adjusted R^2 for returns jumped to 12,8%) and are statistically important. More importantly both models provide initial evidence that the null hypothesis of H2 can be rejected and that conditioning the respondents to a certain scenario of the past might have an effect on their future expectations. Also we have first evidence that the older the Estonian investors get, the lower are their expectations for future returns. General overconfidence (which becomes significant in the volatility model) may also lead to higher expectations about future volatility – the author offers a possible explanation stating that overconfident people take higher risks, invest in riskier assets and thus are ready for higher volatility.

In the third stage, optional control variables (Representativeness and availability, investments in stocks and crypto currencies and financial knowledge) are added to the models as developed in the previous stage. This improves the model for returns, as adjusted R^2 improves slightly, keeping the overall specification statistically significant and also keeping the same variables significant as in the specification of the second stage. Additionally, financial knowledge becomes statistically important with a negative coefficient. A movement in similar direction with age might imply that more sophisticated people might also expect lower returns.

Although adding optional control variables improves specification for returns, it worsens the specification for volatility. As the adjusted R^2 decreased and the specification's F-statistic became insignificant, it is not possible to make conclusions based on this specification. Therefore, the author proceeds by creating an additional specification for volatility, excluding two of the optional control variables with the highest p-values: representativeness and availability heuristics. This improves the models adjusted R^2 to the highest level seen for dependent variable volatility and also makes the specification statistically significant. More importantly evidence is provided that in addition to conditioning and overconfidence which were already important in the second stage model, general negative financial experience becomes statistically significant. As it is only significant on the level of 0.1 and the model's adjusted R^2 is very low (2,1%) this provides only evidence to assume, but not to conclude, that negative financial experience has a certain effect on expected future volatility. If one would want to make a conclusion based on this data, then this conclusion would state that Estonian investors who have had a direct or indirect exposure or financial loss during the major financial crises that have affected Estonian economy in the last 20 years, have higher expectations for future volatility or rephrasing Andersen (2017): once bitten, twice fearful.

CONCLUSION

During the past three decades, Estonia has evolved from a developing post-soviet era country into a country with significantly higher GDP and average salaries. As the local economy has grown, so have the locals become more interested in financial independence and investing in general. Various financial groups with thousands of members have evolved on Facebook, but also a dozen or so investment blogs attract the attention of locals interested in investing. The buzz about investing might seem similar to older investors who were active during the previous major financial crisis, nearly 10 years ago, which was one of the three major crises affecting Estonian investors in the past three decades.

This master thesis sets two main objectives:

1. To identify if past negative financial experiences, such as job loss, stock market loss, or simply exposure to financial crises of Estonian investors have an effect on their expectations about future market returns and volatility.
2. To identify if randomly conditioning the Estonian investors to an optimistic or a pessimistic past performance scenario affects their expectations about future market returns and volatility.

Based on the main objectives two hypothesis are constructed in the first part of this thesis and a literature review of selected variables is presented.

Going further, the second part of this thesis is dedicated to data collection and analysis. For this a survey is conducted among Estonian investors over three days in April 2018. The author suggests that the population size of this study is in the range of 30 000 – 60 000 people. The survey collected 366 answers with a completed response rate of 65%, thus providing sufficient data to make conclusions about the given sample size on a 95% confidence level with a 5,40% confidence interval. The survey was divided into 5 major parts which measured for:

1. Independent variables of interest which were represented by the direct or indirect past exposure and negative financial experiences during three major crises' that have affected Estonia in the past three decades.
2. Dependent variables which were represented by expected returns and volatility of the S&P500 in the following 7 years, conditioned to an optimistic or a pessimistic representation of the past.
3. Socio-economic variables, such as age, gender, education, income, savings and experience.
4. Control variables for returns (optimism) and volatility (overconfidence).
5. Optional control variables such as heuristics, financial knowledge, risk preferences and other.

Majority of the questions were self-assessment questions, however in some cases, questions from previous studies have been used. Optimism was assessed with the traditional Revised Life Orientation Test (LOT-R).

The collected data is cleaned by extracting incomplete answers and consolidating some answers where the scale was too wide and did not generate sufficient answers. Also as the first data analysis showed, the financial experience measure was too specific and thus a general variable of negative financial exposure was created. After the initial processing a total of 336 set of answers are used for the analysis.

In the next stage, a regression analysis is conducted in statistical program Gretl using the ordinary least square method (OLS). A total of seven models is constructed: three for dependent variable of expected returns and four for dependent variables of volatility.

The results of data analysis allow us to come to the following definite conclusion regarding hypothesis H2:

1. When subjects are manipulated by conditioning them to an optimistic or pessimistic background information their level of optimism and overconfidence will be influenced by the proposed condition. Investors in optimistic condition report higher level of optimism (higher expected returns) and higher level of overconfidence (higher level of expected volatility).

The results of data analysis allow us to come to the following potential assumption regarding hypothesis H1:

1. General negative financial experiences seem to have some impact on the expected future volatility. As the variable was significant only in one of the models, we can only assume a link between the two, stating that if such link is present, then the negative experiences from the past would increase the expected volatility.

Therefore, this study is rejecting the null hypothesis and providing strong evidence regarding the hypothesis which H2 which implies that conditioning does affect future return and volatility expectations. However, due to marginal evidence regarding the hypothesis H1, related to effect of negative financial experiences on future return and volatility expectations, this study has to accept the null hypothesis stating that there is no strong evidence providing a link between these variables.

The results of the data analysis allow us to propose additional ideas:

1. Age has a negative effect on the future expected returns, as older people expect lower returns.
2. General overconfidence might have an upward effect on future expected volatility, as more overconfident people expect higher volatility.
3. More sophisticated people expect lower future returns of the S&P500 in the upcoming 7 years.

As this thesis is one of the few works that has aimed to study Estonian investors from a behavioral finance perspective, the author encourages future research in this field. Also as Estonians are generally very tech-savvy people the author encourages to finding links between behavioral economics and new financial instruments such as peer-to-peer lending, and cryptocurrencies, something that this thesis fails to provide.

KOKKUVÕTE

NEGATIIVSETE KOGEMUSTE MÕJU OODATAVALE TOOTLUSELE JA VOLATIILSUSELE EESTI INVESTORITE NÄITEL

Jegor Mitrofanski

Viimase kolme aastakümne jooksul on Eesti muutunud arenevast post-sovetlikust riigist oluliselt kõrgema SKP ja keskmise palgaga riigiks. Majanduse kasvades on kohalikud tundmas järjest enam huvi finantsvabaduse ja investeerimise vastu üldiselt. Facebookis on moodustunud tuhandete jälgijatega finantsteemalised grupid ja paar tosinat investeerimisblogi püüavad kohalike tähelepanu sagedaste postitustega. Melu investeerimise ümber võib tunduda tuttav investoritele, kes olid aktiivsed juba 10-aasta taguse majanduskriisi ajal, mis oli vaid üks kolmest suuremast kriisist, mis viimase kolme aastakümne jooksul Eesti majandust mõjutanud .

Käesoleva magistritöö kaks põhilist eesmärki on:

1. Identifitseerida, kas mineviku negatiivsed finantskogemused nagu kaotus aktsiaturul, töökaotus või üleüldine negatiivne finantskogemus mõjutab Eesti investorite ootusi tootlusele ja volatiilsusele.
2. Identifitseerida, kas juhuslikkuse alusel investoritele esitatav optimistlik või pessimistlik informatsiooni mineviku tootluse ja volatiilsuse kohta mõjutab nende ootusi tuleviku tootlustele ja volatiilsusele.

Lähtuvalt magistritöö eesmärkidest formuleeritakse töö esimeses osas hüpoteesid ja antakse ülevaade antud teemaga seotud kirjandusest.

Magistritöö teine osa on pühendatud andmete kogumisele ja analüüsile. Aprillis 2018 viiakse Eesti investorite seas läbi küsitlus, mida jagatakse erinevates kanalites alates sotsiaalmeediast, lõpetades finantsteemaliste blogide ja autori sotsiaalvõrgustikuga. Autori kalkuleeritud populatsioon, mille kohta analüüsi koostatakse on vahemikus 30 000 – 60 000 inimest. Küsitlusele kogutakse 65%-lise täitmismäära juures kokku 366 vastust, mis võimaldab antud populatsiooni osas teha järeldusi 95% tõenäosusega 5,40%-lise usaldusintervalliga. Küsitlus jagunes viieks suuremaks osaks:

6. Sõltumatud muutujad, milleks on otsesed ja kaudsed negatiivsed finantsilised kokkupuuted ja kogemused kolme suurima kriisi jooksul, mis on mõjutanud Eesti majandust viimase kolme aastakümne jooksul.
7. Sõltuvad muutujad, milleks on küsitletavate investorite ootused S&P 500 aktsiaindeksi järgmise seitsme aasta tootlusele ja volatiilsusele pärast tutvumist juhuslikkuse alusel esitletud optimistliku või pessimistliku mineviku stsenaariumiga.
8. Sotsiaalmajanduslikud muutujad, milleks antud küsitluses on vanus, sugu, haridustase, sissetulek, säästud ja investeerimiskogemus.
9. Sõltuvate muutujate kontrollmuutujad, milleks tootluse puhul on optimism ja volatiilsuse puhul liigne enesekindlus.
10. Valikulised kontrollmuutujad nagu heuristikud, finantsteadmised, riski eelistused ja muud.

Suur osa küsitluse küsimustest olid enesehindamise küsimused, kuid oli ka integreeritud test, milleks oli üleüldise optimismi test ehk LOT-R test.

Kogutud andmed puhastati jättes kõrvale poolikute ja selgelt ekstreemsed vastustega ankeedid. Liiga laia vastuste spektriga küsimuste tulemused ühendati suurematesse kategooriatesse. Pärast esmaseid andmete analüüse pidi autor toimima sarnaselt ka negatiivsete finantskogemuste andmete osas, sest küsitluste tulemused olid kohati liiga spetsiifilised. Seetõttu moodustati üldine negatiivse finantsilise kogemuse näitaja, mis sisaldab endas nii kaudseid kui ka otseseid kogemusi. Andmepuhastuse järgselt kasutatakse selles töös 336 kogutud vastuste komplekti.

Järgmise etapina viiakse läbi regressioonanalüüs statistika programmis Gretl kasutades harilikku vähimruutude meetodit. Kokku koostatakse seitse ökonomeetrilist mudelit: kolm mudelit oodatava tootluse muutujaga ja neli mudelit oodatava volatiilsuse muutujaga.

Andmeanalüüsi tulemused võimaldavad käesoleva magistritöö raames anda selge seisukoht püstitatud hüpoteesi H2 osas:

1. Kui küsitletavale esitatakse optimistlik või pessimistlik taustinformatsioon, on sellel otsene mõju vastaja optimismile ja liigsele enesekindlusele. Optimistliku taustinformatsiooni saanud investorite oli kõrgem nii optimism (oodati kõrgemat tootlust) kui ka liigne enesekindlus (oodati kõrgemat volatiilsust).

Andmeanalüüsi tulemused võimaldavad käesoleva magistritöö raames anda potentsiaalsed oletused püstitatud hüpoteesi H1 osas:

1. Üldistel negatiivsetel finantskogemustel tundub olevat teatav mõju tulevikus oodatavale volatiilsusele. Kuna sõltumatu muutuja oli statistiliselt oluline ainult ühes ökonomeetrilises mudelis võime vaid eeldada et nende muutujate vahel esineb seos. Sellise seose esinemisel, suurendavad negatiivsed finantskogemused tõenäoliselt investori oodatavat volatiilsust.

Eelpool mainitu tõttu lükkab antud töö ümber tõstatatud hüpoteesi H2 null-hüpoteesi ja tõendab ära sisuka hüpoteesi kehtivuse näidates, et optimistlik või pessimistlik taustinformatsioon mõjutavad investor tootluse ja volatiilsuse ootusi. Kuid kuna tõstatud hüpoteesi H1 osas leidub vaid marginaalseid näiteid muutujate vahelistest seostest ei saa antud magistritöö ümber lükata hüpoteesi H1 null-hüpoteesi ja peab tõdema, et negatiivsete finantskogemustel ei ole selget mõju oodatavale tootlusele ja volatiilsusele.

Andmeanalüüsi tulemused võimaldavad pakkuda järgmised töö käigus leitud ideed:

1. Vanusel on negatiivne mõju tootlusootustele, sest vanemate inimeste tootlusootused olid noorematest madalamad.
2. Liigne enesekindlus võib endaga kaasa tuua kõrgemaid ootust volatiilsusele.
3. Paremate finantsteadmistega inimesed ootavad järgmise seitsme aasta jooksul madalamat S&P500 aktsiaindeksi tootlust.

Kuna antud magistritöö on üks väheseid Eesti investoreid käitumusliku rahanduse vaatenurgast uuriv töö, julgustab autor läbi viima täiendavaid uurimusi antud valdkonnas. Võttes arvesse eestlaste keskmisest kõrgemaid IT-oskusi, julgustab autor otsima ühenduslülisid käitumusliku rahanduse ja uute finantsinstrumentide, nagu näiteks ühisrahastuse ja krüptovaluutade, vahel, milleni antud magistritöö ei jõudnud.

LIST OF REFERENCES

Alicke, M. D., Klotz, M., Breitenbecher, D. L., & Yurak, T. J. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*(68), 804-825.

Andersen, S. ((fourthcoming)). Once Bitten, Twice Shy: The Power of Personal Experiences in Risk Taking. *Journal of Financial Economics*, fourthcoming.

Barber, B. M., & Odean, T. (2008). All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. *Review of Financial Studies*(21), 785-818.

Barber, B., & Odean, T. (2002). Online investors: Do the slow die first? *Review of Financial Studies, 15*(2), 455-488.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. . *Journal of Financial Economics, 49*, 1-53.

Barone-Adesi, G., Mancini, L., & Shefrin, H. (2013). *A Tale of Two Investors: Estimating Optimism and Overconfidence.* Lugano: Swiss Finance Institute Research Paper.

Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review, 30*, 879-889.

Carver, C., Segerstrom, S., & Bridges, M. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*(67), 1063-1078.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. (2011). Individual risk attitudes: Measure, determinants, and behavioral consequences. *Journal of the European Economic Association*(9), 522-550.

Easton, P. D., & Sommers, G. A. (2007). Effect of Analysts' Optimism onEstimates of the Expected Rateof Return Implied by EarningsForecasts. *Journal of Accounting Research, 45*, 983-1015.

Fabre, B., & Heude, A.-F. (2009). Optimism And Overconfidence Investor's Biases: A Methodological Note. *Finance, 30*, 79-119.

Heinonen, K., Räikkonen, K., Matthews, K. A., Scheier, M. F., Raitakari, O. T., Pulkki, L., & et al. (2006). Socioeconomic status in childhood and adulthood: Associations with dispositional optimism and pessimism over a 21-year follow-up. *Journal of Personality, 74*, 1111-1126.

Horcher, K. A. (2005). *ESSENTIALS of Financial Risk Management.* Hoboken, New Jersey: John Wiley & Sons, Inc.

Jlassi, M., Naoui, K., & Mansour, W. (2014). Overconfidence behavior and dynamic market volatility: evidence from international data. *Procedia Economics and Finance*, 128-142.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430-454.

Kaustia, M., & Nnüpfer, S. (2008, Dec). Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. *The Journal of Finance, 63*, 2679-2702.

Khan, H. H., Naz, I., Qureshi, F., & Ghafoor, A. (2017). Heuristics and stock buying decision: Evidence from Malaysian and Pakistani stock markets. *Borsa Istanbul Review*, 97-110.

Kliger, D., & Kudryavtsev, A. (2010). The Availability Heuristic and Investors Reaction to Company Specific Events.pdf. *The Journal Of Behavaioral Finance*(11), 50-65.

Larwood, L., & Whittaker, W. (1977). Managerial myopia: Self-serving biases in organizational planning. *Journal of Applied Psychology*(62), 94-198.

Lee, B., O'Brien, J., & Sivaramakrishnan, K. (2008). An Analysis of Financial Analysts' Optimism in Long-term Growth Forecasts. *The Journal of Behavioral Finance, 9*(3), 171-184.

Malmendier, U., & Nagel, S. (2011). Depression babies: do macroeconomic experiences affect risk taking. *The Quarterly Journal of Economics*, 373–416.

Malmendier, U., & Tate, G. (2005). CEO Overconfidence and Corporate Investment. *Journal of Finance, 60*, 2661-2700.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77-91.

Miller , D., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin*(82), 213-225.

Moore, D., & Healy, P. J. (2008). The Trouble with Overconfidence. *Psychological Review, 115*, 502-517.

Nguyen, L. V. (2017). Investment risk profile of the millenial generation. Aalto: Aalto University.

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*(86), 71-99.

Rei, T. (2009). *Kriisist kriisini ehk Eesti praegu ja 10 aastat tagasi.* Tallinn: Statistics of Estonia.

Rei, T. (2009). *Kriisist kriisini ehk Eesti praegu ja 10 aastat tagasi.* Tallinn: Eesti Statistikaamet.

Scheier, M. F., & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research, 16*, 201-228.

Segestrom, S. C. (2007). Optimism and resources: Effects on each other and on health over 10 years. *Journal of Research in Personality, 41*, 772-786.

Solberg, N. L., Evans, D. R., & Segerstrom, S. C. (2009). Optimism and college retention: Mediation by motivation, performance, and adjustment. *Journal of Applied Social Psychology, 39*, 1887-1912.

Statisista. (2018, 04 30). *Facebook: global user age and gender distribution 2018*. Retrieved from Statista: https://www.statista.com/statistics/376128/facebook-global-user-age-distribution/

Statistcs Estonia. (2017, 04 09). *Statistics Estonia*. Retrieved from RAA0013: Gross domestic product and gross national income per capita (ESA 2010): http://andmebaas.stat.ee/Index.aspx?lang=et&DataSetCode=RAA0013

Svenson, O. (1981). Aren we all less risky and more skllful tha our fellow drivers? *Acta Psychologica*(47), 143-148.

Taylor, S., & Brown, J. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193-210.

Treibich, C. (2015). *Are Survey Risk Aversion Measurements Adequate in a Low Income Context?* Aix-Marseille: HAL.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*(5), 207-232.

Tversky, A., & Kahneman, D. (1983). Extensional Verss Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review, 90*(4), 293-315.

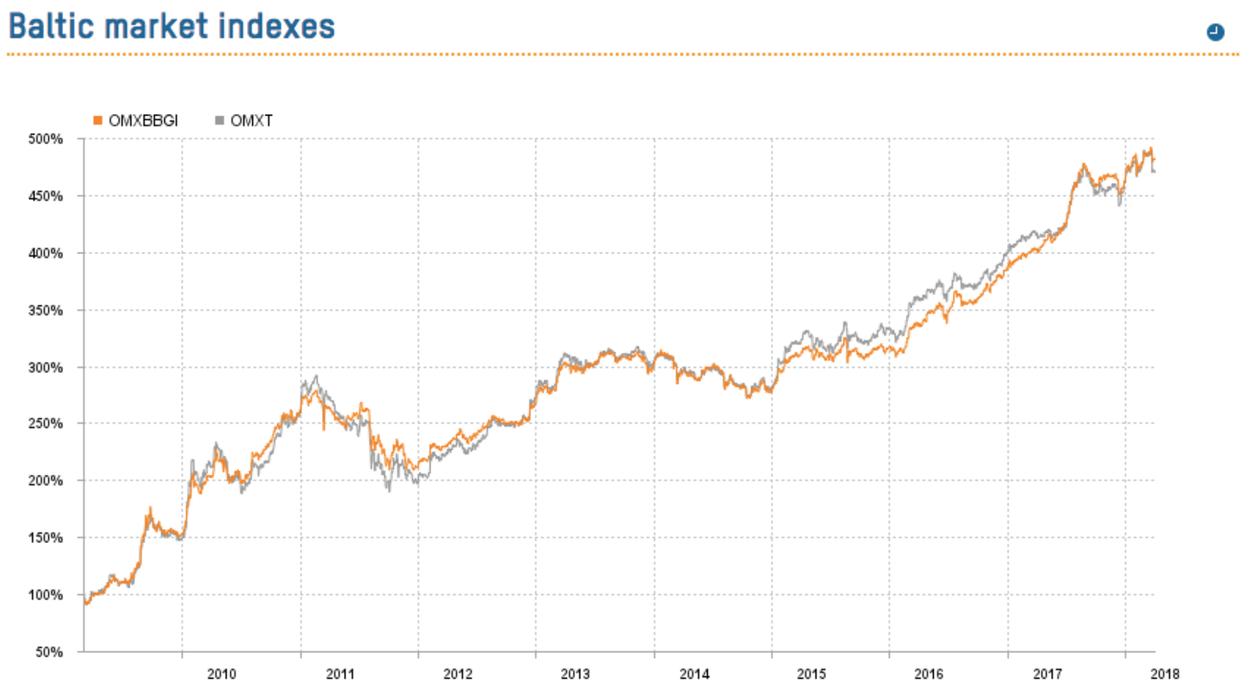
Vissing-Jorgensen, A. (2003). "Perspectives on Behavioral Finance: Does 'Irrational- ity' Disappear with Wealth? Evidence from Expectations and Actions,". *NBER Macroeconomics Annual*.

Weber, M., Weber, E. U., & Nosic, A. (2012). Who takes Risks When and Why: Determinants of Changes in Investor Risk Taking. *Review of Finance*, 1-37.

Zenger, T. (1992). Why do employers only reward extreme performance? Examining the relationships among performance, pay, and turnover. (37), 198-219.

APPENDICES

Appendix 1. Baltic market indexes 2009 – 2018



Index/Equity	01.03.2009	03.04.2018	+/-%
OMX Baltic Benchmark GI	204.22	984.7	382.18
OMX Tallinn	269.18	1,266.78	370.61

Source: http://www.nasdaqomxbaltic.com

Appendix 2. Euribor 6m rates 1993 – 2018

Source: http://www.euribor-rates.eu

Appendix 3. Life Orientation Test Revised

Scale:

Please be as honest and accurate as you can throughout. Try not to let your response to one statement influence your responses to other statements. There are no "correct" or "incorrect" answers. Answer according to your own feelings, rather than how you think "most people" would answer.

A = I agree a lot B = I agree a little C = I neither agree nor disagree D = I disagree a little E = I disagree a lot

1. In uncertain times, I usually expect the best.

2. It's easy for me to relax.

3. If something can go wrong for me, it will. (R)

4. I'm always optimistic about my future.

5. I enjoy my friends a lot.

6. It's important for me to keep busy.

7. I hardly ever expect things to go my way. (R)

8. I don't get upset too easily.

9. I rarely count on good things happening to me. (R)

10. Overall, I expect more good things to happen to me than bad.

Scoring:

Items 3, 7, and 9 are reverse scored (or scored separately as a pessimism measure). Items 2, 5, 6, and 8 are fillers and should not be scored. Scoring is kept continuous – there is no benchmark for being an optimist/pessimist.

Source: Carver et al. (1994)

Appendix 4. Summary statistics

Variable	Mean	Median	S.D.	Min	Max
id	1.86e+07	1.86e+07	1.70e+04	1.86e+07	1.87e+07
time	794.	634.	697.	258.	9.37e+03
returns	5.27	5.00	2.73	-5.00	12.0
volatility	17.2	17.0	5.64	5.00	45.0
returns2	0.0510	0.0488	0.0261	-0.0513	0.113
volatility2	-1.82	-1.77	0.339	-3.00	-0.799
condition	0.494	0.00	0.501	0.00	1.00
all_experience	0.848	1.00	0.359	0.00	1.00
all_exposure	0.494	0.00	0.501	0.00	1.00
dir_exposure	0.280	0.00	0.450	0.00	1.00
indir_exposure	0.381	0.00	0.486	0.00	1.00
all_losses	0.812	1.00	0.391	0.00	1.00
dir_losses	0.679	1.00	0.468	0.00	1.00
indir_losses	0.661	1.00	0.474	0.00	1.00
age	34.5	33.0	8.55	17.0	63.0
gender	0.202	0.00	0.402	0.00	1.00
education	2.23	2.00	0.677	1.00	3.00
income	1.76	2.00	0.702	1.00	3.00
experience	7.73	5.00	6.28	1.00	27.0
savings	2.15	2.00	0.692	1.00	3.00
fin_overconf	2.58	2.50	0.844	1.00	5.00
gen_overconf	2.91	3.00	0.705	1.00	4.75
risk	2.73	3.00	0.696	1.00	4.00
finlit	2.63	2.60	0.695	1.00	4.60
finlit2	0.652	1.00	0.477	0.00	1.00
linda	0.464	0.00	0.499	0.00	1.00
repr_heur	2.79	3.00	0.922	1.00	5.00
availab_heur	2.72	3.00	1.07	1.00	5.00

Source: author’s calculations

Appendix 5. Model R_Basic

R_Basic: OLS, using observations 1-336
Dependent variable: returns

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	6.24491	1.07121	5.830	<0.0001	***
age	−0.0343860	0.0176218	−1.951	0.0519	*
gender	0.203761	0.374082	0.5447	0.5863	
education	−0.0638037	0.222049	−0.2873	0.7740	
savings	−0.312886	0.219315	−1.427	0.1546	
optimism	0.0629567	0.0434878	1.448	0.1487	

Mean dependent var	5.266220		S.D. dependent var	2.731179
Sum squared resid	2423.889		S.E. of regression	2.710188
R-squared	0.030009		Adjusted R-squared	0.015313
F(5, 330)	2.041902		P-value(F)	0.072444
Log-likelihood	−808.7343		Akaike criterion	1629.469
Schwarz criterion	1652.371		Hannan-Quinn	1638.598

Source: author’s calculations

Appendix 6. Model R_H1+H2

R_H1+H2: OLS, using observations 1-336
Dependent variable: returns

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	6.96491	1.02924	6.767	<0.0001	***
age	−0.0301682	0.0169711	−1.778	0.0764	*
gender	0.151322	0.353053	0.4286	0.6685	
education	0.0423558	0.210486	0.2012	0.8406	
savings	−0.220591	0.207456	−1.063	0.2884	
optimism	0.0667063	0.0410753	1.624	0.1053	
condition	−1.78668	0.280916	−6.360	<0.0001	***
all_experience	−0.548984	0.399590	−1.374	0.1704	

Mean dependent var	5.266220		S.D. dependent var	2.731179
Sum squared resid	2142.837		S.E. of regression	2.555981
R-squared	0.142481		Adjusted R-squared	0.124180
F(7, 328)	7.785530		P-value(F)	1.01e-08
Log-likelihood	−788.0295		Akaike criterion	1592.059
Schwarz criterion	1622.596		Hannan-Quinn	1604.232

Source: author’s calculations

Appendix 7. Model R_H1+H2+OC

R_H1+H2+OC: OLS, using observations 1-336
Dependent variable: returns

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	8.14730	1.29216	6.305	<0.0001	***
age	−0.0318843	0.0177715	−1.794	0.0737	*
gender	0.0479609	0.357304	0.1342	0.8933	
education	0.0532311	0.212574	0.2504	0.8024	
savings	−0.115432	0.212071	−0.5443	0.5866	
optimism	0.0625373	0.0413651	1.512	0.1316	
condition	−1.87567	0.283878	−6.607	<0.0001	***
all_experience	−0.452165	0.403053	−1.122	0.2628	
repr_heur	−0.0578392	0.157706	−0.3668	0.7140	
availab_heur	0.0854122	0.135974	0.6282	0.5303	
finlit	−0.371554	0.213432	−1.741	0.0827	*
stocks	−0.391582	0.457564	−0.8558	0.3927	
crypto	−0.378403	0.373602	−1.013	0.3119	

Mean dependent var	5.266220		S.D. dependent var	2.731179
Sum squared resid	2101.318		S.E. of regression	2.550613
R-squared	0.159096		Adjusted R-squared	0.127855
F(12, 323)	5.092532		P-value(F)	8.52e-08
Log-likelihood	−784.7424		Akaike criterion	1595.485
Schwarz criterion	1645.107		Hannan-Quinn	1615.266

Source: author’s calculations

Appendix 8. Model V_Basic

V_Basic: OLS, using observations 1-336
Dependent variable: volatility

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	13.0721	2.14044	6.107	<0.0001	***
age	−0.00262638	0.0365055	−0.07194	0.9427	
gender	0.271600	0.774633	0.3506	0.7261	
education	0.780356	0.460306	1.695	0.0910	*
savings	−0.125253	0.454737	−0.2754	0.7832	
gen_overconf	0.921618	0.441564	2.087	0.0376	**

Mean dependent var	17.19077		S.D. dependent var	5.636417
Sum squared resid	10419.75		S.E. of regression	5.619164
R-squared	0.020947		Adjusted R-squared	0.006113
F(5, 330)	1.412073		P-value(F)	0.219339
Log-likelihood	−1053.734		Akaike criterion	2119.467
Schwarz criterion	2142.370		Hannan-Quinn	2128.597

Source: author’s calculations

Appendix 9. Model V_H1+H2

V_H1+H2: OLS, using observations 1-336
Dependent variable: volatility

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	12.1134	2.16560	5.594	<0.0001	***
age	−0.0140678	0.0370656	−0.3795	0.7045	
gender	0.316907	0.769918	0.4116	0.6809	
education	0.734603	0.459652	1.598	0.1110	
savings	−0.208513	0.453070	−0.4602	0.6457	
gen_overconf	0.887415	0.441396	2.010	0.0452	**
condition	1.16787	0.614248	1.901	0.0581	*
all_experience	1.35331	0.876095	1.545	0.1234	

Mean dependent var	17.19077		S.D. dependent var	5.636417
Sum squared resid	10225.01		S.E. of regression	5.583350
R-squared	0.039245		Adjusted R-squared	0.018741
F(7, 328)	1.914045		P-value(F)	0.066667
Log-likelihood	−1050.564		Akaike criterion	2117.128
Schwarz criterion	2147.665		Hannan-Quinn	2129.301

Source: author’s calculations

Appendix 10. Model V_H1+H2+OC

V_H1+H2+OC: OLS, using observations 1-336
Dependent variable: volatility

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	13.3312	2.57741	5.172	<0.0001	***
age	−0.0245016	0.0389378	−0.6293	0.5296	
gender	0.144893	0.780518	0.1856	0.8528	
education	0.713395	0.466117	1.531	0.1269	
savings	−0.138959	0.465049	−0.2988	0.7653	
gen_overconf	1.02239	0.454823	2.248	0.0253	**
condition	1.05914	0.622647	1.701	0.0899	*
all_experience	1.49514	0.885711	1.688	0.0924	*
repr_heur	0.0936793	0.344690	0.2718	0.7860	
availab_heur	−0.0844247	0.301062	−0.2804	0.7793	
crypto	−0.621682	0.819638	−0.7585	0.4487	
stocks	0.673769	1.00337	0.6715	0.5024	
finlit	−0.717562	0.474847	−1.511	0.1317	

Mean dependent var	17.19077		S.D. dependent var	5.636417
Sum squared resid	10102.96		S.E. of regression	5.592721
R-squared	0.050713		Adjusted R-squared	0.015445
F(12, 323)	1.437937		P-value(F)	0.147046
Log-likelihood	−1048.547		Akaike criterion	2123.094
Schwarz criterion	2172.716		Hannan-Quinn	2142.874

Source: author’s calculations

Appendix 11. Model V_H1+H2+OC_2.0

V_H1+H2+OC_2.0: OLS, using observations 1-336
Dependent variable: volatility

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	13.3672	2.37913	5.619	<0.0001	***
age	−0.0262871	0.0384764	−0.6832	0.4950	
gender	0.159816	0.775580	0.2061	0.8369	
education	0.703922	0.463922	1.517	0.1302	
savings	−0.120973	0.460771	−0.2625	0.7931	
gen_overconf	1.01389	0.447118	2.268	0.0240	**
condition	1.06242	0.620739	1.712	0.0879	*
all_experience	1.51480	0.881095	1.719	0.0865	*
crypto	−0.638660	0.811488	−0.7870	0.4318	
stocks	0.690032	0.999092	0.6907	0.4903	
finlit	−0.705416	0.467593	−1.509	0.1324	

Mean dependent var	17.19077		S.D. dependent var	5.636417
Sum squared resid	10106.91		S.E. of regression	5.576574
R-squared	0.050342		Adjusted R-squared	0.021122
F(10, 325)	1.722852		P-value(F)	0.074530
Log-likelihood	−1048.612		Akaike criterion	2119.225
Schwarz criterion	2161.213		Hannan-Quinn	2135.962

Source: author’s calculation