

# Experimental Studies of Overconfidence in Financial Markets

Inaugural-Dissertation  
zur Erlangung des akademischen Grades eines Doktors  
der Wirtschafts- und Sozialwissenschaften  
der Wirtschafts- und Sozialwissenschaftlichen Fakultät  
der Christian-Albrechts-Universität zu Kiel

vorgelegt von  
MA, Julija Michailova  
aus Vilnius, Litauen

Kiel, 2010

Gedruckt mit Genehmigung der  
Wirtschafts- und Sozialwissenschaftlichen Fakultät  
der Christian-Albrechts-Universität zu Kiel

Dekanin: Prof. Dr. Birgit Friedl

Erstberichterstattender: Prof. Dr. Ulrich Schmidt

Zweitberichterstattender: Prof. Dr. Till Requate

Tag der Abgabe der Arbeit: 13.10.2010

Tag der mündlichen Prüfung: 21.12.2010

To my Family

Моей семье

Mano šeimai

Mojej rodzinie

## ACKNOWLEDGEMENTS

It is a pleasure to thank all those who made this thesis possible!

First of all I acknowledge the Stadt-Kiel Scholarship and the German Academic Exchange Office (DAAD) scholarship. I would like to thank the staff of DAAD, especially Ludmila Winter Souhradová, Maria-Luise Nünning and Hans Golombek.

I am indebted to my first supervisor Prof. Dr. Ulrich Schmidt for accepting me as a doctoral student, his continuous support and encouragement. I greatly appreciate helpful comments and guidance of my second supervisor Prof. Dr. Till Requate. Their experience and well-wishing attitude enabled me to conduct this research successfully.

I want to express gratitude to Prof. Arlington W. Williams and Dr. Briony D. Pulford for a fruitful cooperation, and to thank Prof. Dr. Jürgen Golz for valuable comments. I am grateful to Prof. Dr. Urs Fischbacher for prompt answers to questions about the software z-Tree. It is a pleasure to thank my colleague Israel Waichman for sharing his knowledge of experimental processes. I also acknowledge help of the student research assistants at the chair of “Innovation, Competition Policy and New Institutional Economics”: Julia Schirmacher and Dennis Nissen.

I can't help but thank my dear friends and colleagues who made my life in Kiel interesting and full of joy: Emily Sek, Aidan Islyami, Zhanna Kapsalyamova, Inna Melnykovska, Natalia Maksajda, Maciej Sierżant, Katarzyna Karłowska, Milda Stanionytė, Miriam Schneider, Jaba Ghonghadze, Patrick Weber, Justyna Mysliwy. A special thank to Christian Radden, who helped in translating experimental instructions of my first experiment. Not to forget Indrė Morkunaitė, a close friend who supported me all the way from home. My life is happier when there are you – my friends!

I owe my deepest gratitude to my parents Natalija Michailova and Dr. Andrejus Michailovas for their unconditional love, which only parents feel for their children. My parents have worked hard to give me and my brother the best they could, and I want to thank them for their sacrifice and for being an example for us to follow. I am indebted to you for supporting my initiatives and helping me become the person I am. I am grateful to my little brother Kirilas Michailovas for being my best friend and a trip companion! It is a pleasure to thank my loving grandmothers Liudmila Michailova and Tamara Macneva who were my biggest fans through all these years! I want to thank my aunt Natalja Michailova who has always been an example of the independent career woman to me. Thank you all for the spent hours on the phone listening to my endless stories with patience and interest!

I would like to show my gratitude to Piotr Jamróg who turned my life in Kiel to one of the happiest experiences of my life! The rainy days in Kiel were lighted up by his kind and loving heart! He was the greatest helper, supporter and an inspiration to me!

I want to thank my family for being there for me through good times and bad times, and dedicate my thesis to them!

## TABLE OF CONTENTS

Introduction	1
References	4
<b>Development of the Overconfidence measurement instrument for the Economic Experiment</b>	<b>6</b>
1.1 Introduction	7
1.2 Overconfidence in Psychological Research	9
1.2.1 Definition of the Overconfidence Bias	10
1.2.2 Measurement of Overconfidence	11
1.2.3 Factors Influencing the Degree of Overconfidence	13
1.3 Overconfidence in Financial Literature	14
1.3.1 Findings from the Research	14
1.3.2 Measurement of Overconfidence in Empirical and Experimental Studies	15
1.3.3 Problem Statement	19
1.4 Method	21
1.5 Results	23
1.5.1 Statistical Tests	25
1.6 Second Experiment	28
1.7 Conclusions	31
References	32
Appendix 1.A	40
Appendix 1.B	41
Appendix 1.C	42
Appendix 1.D	43
Appendix 1.E	43
Appendix 1.F	44
Appendix 1.G	46
Appendix 1.H	47
<b>Overconfidence and Bubbles in Experimental Asset Markets</b>	<b>50</b>
2.1 Introduction	51
2.2 Overconfidence	54
2.2.1 Overconfidence in Psychological Research	54
2.2.2 Overconfidence in Financial Research and Contributions	55
2.3 Hypotheses	59
2.4 Pre-Experimental Overconfidence Measurement	60
2.5 Experimental Design	62
2.5.1 Participants	62
2.5.2 Experimental Procedure and the Rules of the Game	63

2.5.3	Incentives	64
2.6	Experimental Results	65
2.6.1	Numerical Characteristics of the Two Types of the Market	65
2.6.2	Overconfidence Measure from the Forecasting Task	69
2.6.3	Forecasting Precision	71
2.6.4	Bubble Measures	72
2.7	Conclusions	78
	References	80
	Appendix 2.A	85
	Appendix 2.B	86
	Appendix 2.C	90
	Appendix 2.D	91
	Appendix 2.E	95
	<b>Overconfidence, Risk Aversion and (Economic) Behavior of Individual Traders in Experimental Asset Markets</b>	<b>96</b>
3.1	Introduction	97
3.2	Psychological Determinants of Trading Activity and Performance	100
3.2.1	Overconfidence	100
3.2.2	Risk aversion	102
3.2.3	Similar Research and Value Added	104
3.3	Hypotheses	108
3.4	Experimental Procedure and the Rules of the Game	109
3.4.1	Asset Market Experiment	110
3.4.2	Risk aversion Experiment	111
3.5	Experimental Results	114
3.5.1	Univariate and Bivariate Analysis	114
3.5.2	Multivariate Analysis	120
3.6	Risk Aversion Analysis: Experimental Results	129
	Conclusions	131
	References	134
	Appendix 3.A	139
	Appendix 3.B	140
	Appendix 3.C	141
	Appendix 3.D	142
	Appendix 3.E	144
	Appendix 3.F	145

## LIST OF FIGURES

1.1	Distribution of accuracy (a) and overconfidence per question (b) in <i>test-50</i>	24
1.2	Distribution of the bias score of participants per test: a. <i>test-50</i> and b. <i>test-18</i>	25
2.1	Average trade prices in both types of markets	65
2.2	Development of the average market price	66
2.3	Volatility of asset prices in both types of markets	67
2.4	Average trading activity (turnover) in both types of markets	68
2.5	Average overconfidence in both types of markets	69
2.6	Comparisons of BS(2-8) and BS(9-15): a. rational market; b. overconfident market	70
2.7	Average forecast and market price: a. rational market; b. overconfident market	72
2.8	Normalized price deviations from FV by trading period and overconfidence level: a. rational market, b. overconfident market	74
2.9	Normalized absolute price deviations from FV by trading period: a. rational market, b. overconfident market	75
3.1	Normalized profits per participant	114
3.2	Distribution of individual profits across trading activity quintiles	115
3.3	The cross-sectional distribution of relative profits by treatment	117
3.4	Distribution of safe choices in the group	128

## LIST OF TABLES

1.1	Average confidence, accuracy and bias score for the three levels of question difficulty of the final overconfidence test ( <i>test-18</i> )	23
1.2	Characteristics of the three levels of question difficulty of the test-18 from the experiment on 14.06.09	29
2.1	Bubble measures in each session	76
2.2	Average values of some of the bubble measures from previous studies	77
3.1	The ten paired lottery-choice decisions	111
3.2	Risk aversion classification based on the number of safe lottery choices	112
3.3	Profit and trading activity in the lowest and highest trading activity quartiles: a. The whole sample; b. Sample without outliers	116
3.4	Trading activity (all errors are heteroskedasticity corrected)	120
3.5	Gains from Trade (all errors are heteroskedasticity corrected)	125



## INTRODUCTION

Behavioral finance is a relatively new field in economics, which uses findings from psychology and sociology for explaining investor behavior or market anomalies when rational models provide no sufficient explanations (Glaser et al., 2003). Traditional decision making models in economics rest on the premise that the world is populated by rational, calculating, and unemotional utility maximisers, called *homo economicus*. However numerous empirical and experimental studies demonstrated that investors' behavior does not fit into this rationality framework.

Overconfidence is one of the psychological biases, which can cause deviations from rational behavior. The concept of overconfidence is based on the large body of evidence from cognitive psychological research. Studies conducted by psychologists demonstrate that most people are overconfident about their abilities, knowledge, precision of their personal information, and are unreasonably optimistic about their future. Interest of economists in the consequences of investor and market overconfidence has generated a large body of literature. The research evidence suggests that overconfidence in financial markets can manifest itself through high volumes of trade, excessive price volatility, dissemination of speculative bubbles, etc. On the individual level overconfidence results in higher trade aggressiveness, portfolio undiversification, risk underestimation and suboptimal performance.

This thesis takes a closer look at the imperfections of human nature and their impact on behavior in economic situations, by conducting a laboratory experiment in a financial market setting. The aim of this research was twofold: first, to investigate the role of market overconfidence in the occurrence of bubbles in asset prices and the emergence of other stylized facts of financial markets (excessive trade, excessive price volatility); second, to investigate the influence of behavioral factors, namely the degree of overconfidence and risk aversion, on financial decision making of economic subjects. This is a challenging research topic, especially in the face of devastating economic consequences of the recent real estate bubble. It is important to improve the knowledge of the causes lying behind the formation of bubbles and their subsequent burst, and get a better understanding of the motivation and reactions of market participants. Ability to grasp the role of psychological biases on the functioning of financial markets could result in the development of the mechanisms and implementation of policies that could, if not hinder the occurrence of the speculative price bubbles, at least enable control of their magnitudes.

In real-world situations no direct way to assess the degree of market overconfidence exists. It is also a troublesome task to guess who of the investors is overconfident and how strong her degree of overconfidence is. Not to mention their informational, financial, and preferential heterogeneity. Necessity to control for too many factors in the field data was the reason to choose laboratory experiments for the purpose of data collection. Experimental simulations of financial markets are of high value to research, as they allow isolating the effect of a particular bias on individual behavior. Moreover according to Friedman and Sunder (1994) experimental data are relatively easy to interpret. Biais et al. (2005) point out that study of controlled environments allows more confident inferences about cause and effect relations. Laboratory market setting also enables overcoming several problems, existing in determination of the fundamental asset value, namely estimation of expected dividends on the asset through the time period, determination of the terminal asset value and discount rates for calculation of the present value. These components can be controlled in the laboratory asset market, and thus favorable conditions for observing subjects' behavior are created.

This thesis consists of three papers. In Chapter 1 an instrument is developed that is later used in economic experiments to measure subjects' overconfidence. Chapter 2 analyzes the connection between overconfidence, stock market bubbles and other stylized facts (high trade and volatility) from an aggregated market perspective. In Chapter 3 economic consequences of overconfidence and risk aversion are being assessed at the individual level. Each chapter begins with motivation of the research issue and review of the related literature. After that description of the experimental design is presented and research objectives are formulated. Then data analysis follows. Finally conclusions are drawn. Experimental instructions are provided in the Appendixes to the papers.

Prior to running the economic experiments, described in Chapters 2 and 3, subjects' degree of overconfidence had to be measured. However in economic experiments no conventional method of overconfidence measurement exists. Various proxies, tests and tasks applied for this purpose do not always offer measures that are satisfactory. These considerations were the basis upon which the first chapter of this thesis was built. This chapter is aimed at the development of an instrument (test) that would enable the construction of a comprehensive measure of individual overconfidence for the use in economic experiments. The developed instrument should enable assessment of the differences among subjects with respect to their degree of overconfidence and minimize the measurement error. The final test, consisting of 18 general knowledge questions, was obtained in a two-stage procedure. In the first experimental phase a pilot test, consisting of fifty general knowledge questions of unknown

difficulty, was conducted. It enabled grouping of items into three difficulty categories: hard, average-difficulty and easy questions. The second phase was aimed at verification of the results' replicability. Statistical tests supported the existence of the hard-easy effect, verified the success of categorization of questions into three levels of difficulty, and showed that gender was not associated with overconfidence in the developed instrument. Compared to the tests used in previous studies, the obtained test is believed to result in improved measurement quality of overconfidence.

In Chapter 2 relationship between market overconfidence and occurrence of the stock prices' bubbles is investigated. Stock price bubbles are defined as deviations from asset's fundamental value, which equals the present value of the dividends' stream that an owner expects to receive. An interesting question is: Why paying a price for the asset which exceeds the expected value of its dividends? Scheinkman and Xiong (2003) state that overconfidence is the main factor that makes people pay higher prices than the underlying fundamental value of an asset. In this experiment, sixty participants traded in ten computerized asset markets of two types: rational and overconfident. Markets were constructed on the basis of subjects' overconfidence, measured in the administered pre-experimental psychological test sessions. The most overconfident subjects formed overconfident markets, and the least overconfident rational markets. The empirical evidence presented in the paper refined differences between market outcomes in the experimental treatments and suggested the connection between market overconfidence and market outcomes. Prices in rational markets tended to track the fundamental asset value more accurately than prices in overconfident markets, and were significantly lower and less volatile than the average overconfident prices. Strong positive correlation between market outcomes and overconfidence measures drew to conclusion, that increase in market overconfidence was associated with increase in average price and trading activity. Large and significant correlation between bubble measures and measures of overconfidence provided additional evidence that overconfidence had significant effect on price and trading behavior in experimental asset markets.

Baker and Nofsinger (2002) point out that, investors can do serious harm to their wealth, if they allow psychological biases and emotions to affect their investment decisions. While overconfidence is one of the biases, which can induce deviation from rational economic behavior, degree of individual risk aversion is another personality trait, determining investment choices and strategies. In Chapter 3 influence of these behavioral factors (overconfidence and risk aversion) on financial decision making of economic subjects is analyzed. In order to achieve this purpose two kinds of experiments were conducted: the first one was the asset

market experiment, whereas the second was aimed at the individual risk aversion measurement. In previous studies overconfident and rational traders interacted in the same asset market. In contrast, in present experiment these two types of subjects were separated from each other by assigning them to the different types of markets: rational and overconfident. In each market session six subjects of the same type (rational or overconfident) took part. Data, collected from ten experimental sessions, revealed that individual performance and trading activity were overconfidence dependent. Even small variations in miscalibration among players of the same type, comprising each of the asset markets, were sufficient to cause this effect. In the second part of experiment, post hoc assessment of risk aversion was implemented in a sample of former participants of the asset market experiment (32 persons). The presented evidence suggests that risk aversion was not among the factors that had influence on individual engagement in trading activity or performance. It is concluded that in the sample, for which risk aversion measurements were obtained, experimental market outcomes were overconfidence and not risk aversion driven.

I hope that this work and its conclusions could be of benefit to the policy makers. According to Gervais, Heaton, and Odean (2003) most economic theories assume that people behave rationally and optimally. However evidence, presented in this thesis, indicates that psychological bias of overconfidence has important impact on asset market prices, their volatility and bubble formation, and can be harming investors' wealth. Thus danger arises from adopting assumption of subjects' rationality into policy implications. In this context, I would like to emphasize the role of investment education in mitigating the negative effects of subjects' overconfidence. Equipping individuals with better investment knowledge could be one of the ways to correct their market behavior. Keeping in mind, that "financial intermediaries profit from exploitation of investor biases" (Daniel, Hirshleifer, and Teoh, 2002), additionally individuals could receive investment advice from financially disinterested governmental agencies.

## REFERENCES

- Baker, K. H., Nofsinger, J. R., (2002), Psychological biases of investors. *Financial Services Review*, Vol. 11(2), p. 97-116.
- Biais, B., Hilton, D., Mazurier, K., Pouget, S., (2005), Judgmental overconfidence, self-monitoring and trading performance in an experimental financial market. *Review of economic studies*, Vol. 72(2), p. 287-312.

Daniel, K., Hirshleifer, D., Teoh, S. H., (2002), Investor Psychology in Capital Markets: Evidence and Policy Implications. *Journal of Monetary Economics*, Vol. 49(1), p. 139-209.

Friedman, D., Sunder, S., (1994), *Experimental Methods, A Primer For Economists*. New York: Cambridge University Press, 248 p.

Gervais, S., Heaton, J. B., Odean, T., (2003), Overconfidence, Investment Policy, and Executive Stock Options. Rodney L. White Center for Financial Research, Working Paper No. 15-02.

Glaser, M., Nöth, M., Weber, M., (2003), Behavioral finance. In Koehler, D. J., and Harvey N. (Eds.), "Blackwell Handbook of Judgment and Decision Making", p. 527-546.

Scheinkman, J. A., Xiong, W., (2003), Overconfidence and Speculative Bubbles. *Journal of Political Economy*, Vol. 111, p. 1183-1219.

# CHAPTER 1

## DEVELOPMENT OF THE OVERCONFIDENCE MEASUREMENT INSTRUMENT FOR THE ECONOMIC EXPERIMENT

## 1.1 INTRODUCTION

A large body of economic literature presents results of experiments on overconfidence. The concept of overconfidence is based on the evidence from cognitive psychological research, which suggests that human-beings overestimate their knowledge, abilities and precision of their information. As example Bar-Hillel (2001) points out that, when subjects are  $P\%$  sure that they have answered a question correctly in fact they are right on average less than  $P\%$  of the time. There is plenty of evidence for people to be in general overconfident, and phenomenon of overconfidence has been found in many different samples of the population, e.g. students (Fischhoff, Slovic and Lichtenstein, 1977; Koriat et al., 1980; Zakay and Glicksohn, 1992), members of the armed forces (Hazard and Peterson, 1973), CIA analysts (Cambridge and Shreckengost, 1978), entrepreneurs (Baron, 2000), clinical psychologists (Oskamp, 1962), bankers (Staël von Holstein, 1972), executives (Moore, 1977), negotiators (Neale and Bazerman, 1990), managers (Russo and Schoemaker, 1992), lawyers (Wagenaar and Keren, 1986), and civil engineers (Hynes and Vanmarcke, 1976). Overconfidence is already present in children (see Powel and Bolich, 1993; Allwood, Granhag, and Jonsson, 2006), and boys are found to be more overconfident than girls (e.g. Sieber, 1979; Newman, 1984; Allwood et al., 2006). However, in adult samples no differences between both genders in overconfidence<sup>1</sup> are observed (e.g. Lichtenstein and Fischhoff, 1981; Gigerenzer et al., 1991).

In economic experiments, there is no conventional method of measurement of the inborn level of subjects' overconfidence. For this purpose various proxies, tests and tasks are used, that not always offer a satisfactory measure of individual overconfidence. The need for this research is stimulated by the fact that previous experiments have drawbacks in the way they measure overconfidence, and thus overconfidence might have been caused (to some extent) by other reasons than the imperfection of human nature, but rather by the mistakes in the tests'/ tasks' construction. Findings from psychological research indicate that the observable biases in judgment are often result of the inappropriateness of the task, e.g. a task is unclear to subjects, one gender finds task more difficult than the other, or there is not enough motivation for active participation. Thus development of the overconfidence test was implemented with the following assumptions in mind. First of all, most of the foregoing researchers followed the famous work by Russo and Schoemaker (1992) and used interval elicitation tasks to assess overconfidence. However, these tasks are prone to produce extreme overconfidence (see

---

<sup>1</sup> Assessed via sets of general knowledge questions.

Klayman et al., 1999). Second, previous authors used tests that were neither balanced to the hard-easy effect nor country or gender balanced. Yet, unbalanced tests can artificially create high levels of under- or overconfidence either in the whole group, or in parts of it. Third, overconfidence was often assessed based on the insufficient number of assignments or test items; psychological studies of overconfidence use amounts of items that are much higher. And last but not least, many of the tasks and tests were either not administered, or were not (financially) rewarded.

This paper presents the results of the two experiments, aimed at the development of the instrument (test) that would enable construction of a comprehensive measure of individual overconfidence for the use in economic overconfidence experiments dealing with: 1) the role of overconfidence in occurrence of stock-prices' bubbles, and 2) impact of overconfidence and risk aversion on economic behavior of individual traders. Test is intended for the detection of potential experimental subjects with high and low degrees of overconfidence and their subsequent grouping into two types of asset markets: rational and overconfident. Hence, a well-designed instrument should allow assessment of differences between the subjects with respect to their overconfidence and minimize the measurement error.

The developed test differs from those used in prior economic experiments in some important respects. First, another test format was chosen, namely multiple choice discrete propositions' task format, which is clearer to subjects and is not inherently prone to production of extreme overconfidence levels. Second, test was balanced to the hard-easy effect, by the inclusion of an equal number of questions of three difficulty levels (hard, medium-difficulty and easy). Third, in construction of the test it was controlled for the possible country and gender bias, e.g. no inclusion of questions that might be easier to one gender. And finally, compared to some studies, the test is expanded to include more items. Instrument was obtained in a two-stage procedure in which a pilot test was used to assess questions' difficulty, based on the group accuracy in answering every item of the 50 initial. Then six questions of each of the three difficulty types were chosen for the inclusion in the final test. The second experimental phase was aimed at verification of replicability of results, namely of the average degree of group overconfidence, the obtained categorization into three difficulty levels and of controlling for the gender bias. Both experiments were conducted with the students enrolled into different disciplines of social sciences. Experimental sessions were administered and subjects were offered a reward, on the basis of competition in test accuracy. The final instrument consists of 18 general knowledge questions unrelated to economics, financial markets or experiments. Questions are not connected to economics, as otherwise they could



cause biased results if the same test is used with a heterogeneous pool of subjects<sup>2</sup>. Evidence was found for the significant effect of the question difficulty on the overconfidence measure and existence of the gender bias. Compared to medium and easy questions, which resulted in under-confidence, hard questions produced significantly higher levels of overconfidence. The three types of questions also significantly varied from each other in terms of the produced confidence and accuracy. This result verified the success of categorization of questions into 3 levels of difficulty in the created overconfidence measurement test. In the initial instrument as much as 16% of variance in accuracy and 7% of variance in confidence was explained by gender. In the final test gender is not associated with overconfidence, and there is almost no variance in confidence and accuracy that is gender dependent.

Paper proceeds as follows. In Section 1.2 a review of the findings of psychological overconfidence literature are presented. In Section 1.3 findings from the theoretical and empirical research on overconfidence in finance are introduced, and ideas, on how overconfidence was measured in the previous experimental research, are presented; in closing of this Section a problem statement with the ideas about research improvement are provided. In Section 1.4 methodology of the test construction is described. In Section 1.5 statistical data analysis is presented. In Section 1.6 findings from the experiment with the final overconfidence measurement instrument are analyzed, and, finally Section 1.7 concludes.

## 1.2 OVERCONFIDENCE IN PSYCHOLOGICAL RESEARCH

Our life is full of uncertainty, and many decisions are based on beliefs concerning the likelihood of uncertain events (Tversky and Kahneman, 1982). These beliefs can be expressed in numerical form as subjective probabilities. Question of generation of these probabilities is one of the most important topics in the area of cognitive psychology (Bar-Hillel, 2001). Bar-Hillel (2001) suggests that subjective probabilities are not just imperfect or inaccurate versions of objective probabilities, but rather are governed by cognitive principles of their own. To generate subjective probabilities, people rely on a limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations. In general, these heuristics are quite useful, but sometimes they lead to severe and systematic errors (Tversky and Kahneman, 1982), and thus to non-optimal judgments (see Russo and Schoemaker, 1992, for the description of the heuristics employed by people for assessment of probabilities). Use of heuristics for generation of

---

<sup>2</sup> Deaves et al. (2004) also motivate their choice of non-economic questions by the attempt “to avoid giving either group of participants a relative advantage because of subject content”.

subjective probabilities is a cognitive cause of overconfidence. Overconfidence is characterized by the tendency to overestimate one's skills, prospects for success, the probability of positive outcomes or the accuracy of one's knowledge, and arises from not knowing the limits of one's knowledge (Conger and Wolstein, 2004).

### 1.2.1 DEFINITION OF THE OVERCONFIDENCE BIAS

In financial literature there are several findings that are often summarized under the concept of overconfidence: miscalibration, the better than average effect, illusion of control, and unrealistic optimism. The issue of whether these notions are related is mainly unexplored<sup>3</sup> (Glaser and Weber, 2007). In psychological research, however, only miscalibration is defined as overconfidence.

#### *Miscalibration*

Miscalibration is a cognitive bias that rests on the fact that people tend to overestimate the precision of their knowledge. In the experiments on calibration, participants answer a series of (general knowledge) questions and stipulate their confidence of being correct for each answer. Calibration is tested by comparing the percentage of questions that a participant has answered correctly with the participant's average confidence in the answers to these questions. Individuals are considered to be well calibrated if the following condition is satisfied: over the long run of those responses made with confidence  $P$ , about  $P\%$  should be correct (Adams, 1957). However most of the people are not well-calibrated and demonstrate overconfidence (miscalibration), which manifests itself through a systematic deviation from perfect calibration and is defined as an "unwarranted belief in the correctness of one's answer" (Lichtenstein, Fischhoff, and Phillips, 1977). Typically, for all questions the proportion of correct answers is lower than the assigned probability (Lichtenstein, Fischhoff, and Phillips, 1982). In a second strand of psychological literature, where overconfidence is measured by asking subjects to state for a series of questions with unknown numerical answer an upper and lower limit such that a subject is  $X\%$  sure that the real answer would fall into that interval, the usual finding is that subjects' probability distributions are too tight (Lichtenstein, Fischhoff, and Phillips, 1982). E.g. when subjects are asked to state 90% confidence intervals for some uncertain quantity, the percentage of true values that fall outside the interval, is higher than 10% (the percentage of surprises of a perfectly calibrated person). In the study of Alpert and

---

<sup>3</sup> Oberlechner and Osler (2003), Glaser, Langer, and Weber (2005), Glaser and Weber (2007) found no significant correlation between miscalibration and better than average effect measures.

Raiffa (1982) 50% intervals included the true quantity only about 30% of the time; 98% intervals included the true quantity only about 60% of the time.

#### *Better than average effect*

Inclination of people to exaggeration of their talents embodies itself in the better than average effect. Taylor and Brown (1988) document in their survey that people have unrealistically positive views of the self, i.e. they think about themselves as possessing above the average abilities (e.g. with regard to skills or positive personal traits) compared to other people. One of the most cited works by Svenson (1981) states that 82 percent of a group of students rank themselves among the 30 percent of drivers with the highest driving safety. Sümer et al. (2006) also found better than average effect among drivers in their sample, resulting from exaggerated ratings of self-reported driving skills.

#### *Illusion of control and unrealistic optimism*

Illusion of control is linked to the exaggeration of the degree to which one can control one's fate. Subjects prone to the illusion of control, tend to underestimate the role of chance in human affairs and to misperceive games of chance as games of skill (Kahneman and Riepe, 1998). Langer (1975) in her pioneering work defines this phenomenon as “an expectancy of a personal success probability inappropriately higher than the objective probability would warrant”. The phenomenon of unrealistic optimism about future life events is a cognitive bias that is strongly related to the illusion of control (Weinstein, 1980). Johnson, McDermott, Barrett, Cowden, Wrangham, McIntyre, and Rosen (2006) point out that numerous empirical findings confirm “that mentally healthy people tend to exhibit psychological biases that encourage optimism, collectively known as “positive illusions”. According to Kahneman and Riepe (1998) “most people's beliefs are biased in the direction of optimism”. Griffin and Brenner (2005) note, that “optimistic overconfidence” represents overestimation of the probabilities of the events that are advantageous to the subject. Probabilities of the unfavorable events are underestimated by optimists; even in cases when they have no control over them, e.g. Kahneman and Riepe (1998) note “most undergraduates believe that they are less likely than their roommates to develop cancer or to have a heart attack before the age of fifty”.

### 1.2.2 MEASUREMENT OF OVERCONFIDENCE

There are two types of calibration assessment techniques used in the psychological experiments: making probability judgments about discrete propositions, and the calibration of probability density functions assessed for uncertain numerical quantities (the fractile method).

### *Measurement of Calibration with Discrete Propositions' Task*

To measure overconfidence with the discrete propositions, subjects are suggested to answer a series of questions and state their confidence for each question that their answer was correct. Discrete propositions can give no alternatives for an answer, or suggest one, two, or multiple answer choices.

Calibration can be expressed through several various measures (e.g. Calibration curve, Brier score). However a convenient measure, enabling discrimination between under- and overconfidence, is the bias score. The bias score is calculated as the difference between the average confidence level across all questions and the proportion of correct answers. A positive bias score represents overconfidence, and a negative bias score represents underconfidence. A bias score of zero indicates an accurately calibrated (neutral) person.

$$\text{bias score} = \text{average \% confidence} - \text{average \% correct} \quad (1.1)$$

Or as in Pulford (1996):

$$\text{over / underconfidence} = \frac{1}{N} \sum_{i=1}^T n_i (r_i - c_i) \quad (1.2)$$

Here,  $T$  is the total number of response categories used,  $n_i$  is the number of times the response  $r_i$  was used and  $c_i$  is the proportion correct for all items assigned probability  $r_i$ .

### *Measurement of Calibration with the Fractile Method*

Lichtenstein, Fischhoff, and Phillips (1982) note that uncertainty about the value of an unknown continuous quantity (e.g. what is the air distance from London to Tokyo) “may be expressed as probability density function across the possible values of that quantity”. The assessor has to state values of the uncertain quantity that are associated with a small number of predetermined fractiles (quantiles) of the distribution, i.e., as mentioned before, to state, for a series of questions with unknown numerical answer, upper and lower limits such that she is  $X\%$  sure that the real answer would fall into that interval. There are two calibration measures for continuous items: interquartile index and surprise index. Interquartile index is the percentage of items for which the true value falls inside the interquartile range (i.e., between the 0.25 and 0.75 fractiles) (Lichtenstein et al., 1982). Interquartile index of a perfectly calibrated person is 0.5. Surprise index is the percentage of true values that falls outside the most extreme fractiles assessed. Lichtenstein et al. (1982) write that when the most extreme fractiles are assessed as 0.01 and 0.99, the surprise index of the perfectly calibrated person should be 2. Large surprise index shows the inability of the assessor to state confidence

bounds wide enough to include as much as possible of the true values. This indicates overconfidence.

### 1.2.3 FACTORS INFLUENCING THE DEGREE OF OVERCONFIDENCE

Keasey and Watson (1989) identified four factors that have an impact on the accuracy-confidence relationship: task complexity, amount of feedback, subjects' level of motivation, and their skills.

#### *Hard-Easy Effect*

The degree of overconfidence is connected to the complexity of the task. This is called the "had-easy" effect. This effect occurs when the degree of overconfidence increases with the increase in the difficulty of the questions, where difficulty is measured as the percentage of correct answers (Gigerenzer, Hoffrage, and Kleinbölting, 1991). Lichtenstein et al. (1982) suggest that "the most pervasive finding in recent research is that people are overconfident with general-knowledge items of moderate or extreme difficulty". Many studies have supported this conclusion, e.g. Fischhoff, Slovic and Lichtenstein (1977), Koriat, Lichtenstein and Fischhoff (1980), Nickerson and McGoldrick (1965), Cambridge and Shreckengost (1978), Arkes, Christensen, Lai, and Blumer, (1987), Ronis and Yates, (1987), Sniezek, Paese, and Switzer, (1990) and etc. The degree of overconfidence is the highest with the tasks of high difficulty (e.g. Clarke, 1960; and Pitz, 1974); as tasks get easier, overconfidence is reduced (Lichtenstein et al., 1982).

#### *Motivation and Feedback*

The two ways of increasing subjects' calibration are: motivation through reward for their assessment to be more precise, and outcome feedback (Lichtenstein, Fischhoff, and Phillips, 1982).

Motivation, according to Bohner et al. (1998) is one of the factors that encourage people to abandon the use of effort minimizing heuristics in favor of more effortful probabilities' estimation strategies; thus it has an impact on the accuracy-confidence relationship. Motivation through reward is named to be a tool, helping to improve subjects' calibration, by Lichtenstein et al. (1982). This finding is supported by the paper of Hoelzl et al. (2005) who have discovered a significant change in overconfidence pattern depending on the existence or nonexistence of the monetary reward.

Lichtenstein, Fischhoff, and Phillips (1982) state, that receiving outcome feedback after every assessment is the best condition for successful training to be better calibrated. Adams and

Adams (1958) have found modest improvement in subjects' calibration after five training sessions where they were given feedback on their performance. Fischhoff (1982) reports some successful training exercises, mostly using large amounts of well-structured feedback. Lichtenstein and Fischhoff (1980) studied the impact of extensive, personalized calibration feedback on two groups of subjects. Perceptible improvement in calibration was reported, however no improvement was found in probabilities' assessment by fractile method. In general, improvement in the accuracy of estimates is difficult to achieve (see e.g. Ferrell and McGoey, 1980; Lichtenstein and Fischhoff, 1980; Koriat et al., 1980), and there are reasons to be pessimistic about how well training transfers across time or tasks (Camerer, 1995).

## 1.3 OVERCONFIDENCE IN FINANCIAL LITERATURE

### 1.3.1 FINDINGS FROM THE RESEARCH

Following psychological research in overconfidence, interest in the consequences of economic subjects' overconfidence on financial decision making, functioning of markets and economic outcomes has occurred in behavioral economics. Findings of behavioral finance have an important value in understanding various anomalies and stylized facts found for example in the stock market. Overconfidence research in economics is developing in two directions: theoretical modeling and empirical testing of these models.

Theoretical models about the impact of overconfidence on the processes in financial markets and the behavior of investors are based on the initial assumption of traders' overconfidence, whose decision-making is modeled according to this premise. Behavioral finance models predict that overconfidence causes excess trading volume (De Bondt and Thaler, 1984; Shiller, 2000; Benos, 1998; Caballé and Sákovics, 2003), and excess price volatility (Scheinkman and Xiong, 2003; Benos, 1998, Daniel et al., 1998); it induces occurrence of the speculative price bubbles (Scheinkman and Xiong, 2003) and increases the market depth (Odean, 1999; Kyle and Wang, 1997; Benos, 1998); it makes markets underreact to abstract, statistical, and highly relevant information and overreact to salient, but less relevant information (Odean, 1998); it makes returns of financial assets predictable (Daniel et al., 1998, 2001; Scheinkman and Xiong, 2003). Overconfidence increases investors' tendency to herd (Hirshleifer, Subrahmanyam and Titman, 1994) and makes them choose riskier and undiversified portfolios (Odean, 1998, 1999; Lakonishok, Shleifer and Vishny, 1992), overconfident investors trade more aggressively, i.e. their trading activity is too high (Odean, 1999; Gervais and Odean, 2001) and their expected utility is reduced (De Long et al., 1991; Odean, 1998).

There are not so many empirical and experimental studies testing the assumptions of the abovementioned theories of the impact of overconfidence on financial decision making, functioning of the markets and economic outcomes. Empirical findings support the premise of theoretical models, that overconfidence results in high trading volume in the market (Statman, Thorley, and Vorkink, 2006; Kim and Nofsinger, 2003); it also increases the probability of bubbles' occurrence<sup>4</sup> (Oechssler, Schmidt and Schnedler, 2007). A higher degree of overconfidence reduces traders' performance/ welfare (Fenton-O'Creevy et al., 2003; Biais, Hilton, Mazurier, and Pouget, 2005; Odean, 1999; Barber and Odean, 2002; Nöth and Weber, 2003), and causes mistakes in financial decision making (Biais et al., 2005); unrealistically positive self-evaluation increases trading volume (Glaser and Weber, 2007). There is no clear conclusion about how overconfidence may influence markets' reaction to new information: e.g. Loughran and Ritter (1995) found that markets overreact to new information, and studies by Bernard and Thomas (1989, 1990), Michaely, Thaler and Womack (1995) detected underreaction. There is no clear relationship between the degree of overconfidence, and the degree of professionalism: in the studies by Kirchler and Maciejovsky (2002) and Glaser and Weber (2007) overconfidence increases with experience, whereas studies of Menkhoff, Schmeling, and Schmidt (2006) and Biais et al. (2005) find the reverse dependence.

### 1.3.2 MEASUREMENT OF OVERCONFIDENCE IN EMPIRICAL AND EXPERIMENTAL STUDIES

As mentioned above, there are a few empirical and experimental studies designed to test the impact of overconfidence on financial decisions, market outcomes and subjects' performance. Some of them present only an indirect evidence of such impact, as they measure overconfidence via various proxies and it is not always clear who of the subjects and how strong are overconfident. Other studies measure the inborn level of subjects' overconfidence via the different tasks and tests, related or non-related to economics and finance. Such tests usually enable construction of the overconfidence measure for each individual. Most often these tasks are related to confidence intervals' estimations in the spirit of the work by Russo and Shoemaker (1992).

#### *Proxies for Overconfidence*

Papers that use proxies for overconfidence do not allow for the numerical measurement of the degree of overconfidence. E.g. Statman, Thorley, and Vorkink (2006) test the hypothesis of

---

<sup>4</sup> Top-rank belief variable "has a positive and significant effect on the probability of bubbles" (Oechssler et al., 2007).

interdependence of overconfidence and high trading volume for the USA stock market. As a proxy for the degree of overconfidence authors suggest using high past returns. They argue that after high past returns posterior volume of trade will be higher, as successful investment increases the degree of overconfidence. The same proxy for overconfidence (i.e. high past returns) was utilized by Kim and Nofsinger (2003) for the Japanese stock market.

Barber and Odean (2001) use gender of the trader as a proxy for overconfidence. Their assumption is that, based on the psychological literature, women are less overconfident than men, thus they are going to trade less than men. In their study men are actually found to trade more than women. In another paper Barber and Odean (2002) employ as a proxy of overconfidence changes in the trading patterns and performance of the 1607 investors who switched from the phone-based trading to online trading between 1992 and 1995. They present evidence that these investors traded more actively and speculatively, and performed subpar.

Blavatsky (2008) measures overconfidence by the taken choice in a simple task: subjects can either bet on their knowledge or on the equivalent lottery. Those who choose an option to bet on their own knowledge are classified to be overconfident (others are underconfident). Under this measurement procedure, subjects, on average, exhibit underconfidence about their own knowledge, and their confidence does not depend on their attitude towards risk/ ambiguity.

#### *Overconfidence Measured via Tests and Tasks*

In comparison to the studies that use various proxies to measure overconfidence, questionnaire studies enable direct assessment of each subject's under- or overconfidence.

Kirchler and Maciejovsky (2002) investigate overconfidence within the context of an experimental asset market. Miscalibration of subjects is measured before each trading period, with the help of the two price prediction tasks: point prediction with the confidence in forecast, and 98% confidence interval prediction. Results presented in their paper indicate that in some periods participants demonstrate overconfidence and in others underconfidence, thus they are not generally prone to overconfidence. Kirchler and Maciejovsky (2002) also show that higher degree of overconfidence is negatively correlated with the earnings of the participants of their experiment.

Fenton-O'creevy, Nicholson, Soane, and Willman (2003) examine the impact of illusion of control on the performance of traders in four investment banks. They use a computer-based task measurement of the illusion of control to execute measurement of overconfidence: 107 participants had to raise an index on the computer screen by pressing keyboard-buttons and



rate their success in doing so from 1 (not at all successful) to 100 (very successful). The index in reality was modeled as random walk process with an upward trend, and thus the button had no influence on its development.

Biais, Hilton, Mazurier, and Pouget (2005) conduct an experiment to check if overconfidence has impact on subjects' trading performance (trading activity and profits). They use the scale adapted from Russo and Schoemaker (1992) to measure the degree of overconfidence in a group of 245 students. Their test consisted of 10 general-knowledge questions with known numerical answers for which subjects had to state 90% confidence intervals. Several weeks later, after the students' overconfidence was measured, they participated in the experimental asset market. Questions that were used to measure subject's miscalibration had nothing to do with financial markets, yet they affected strategies and performance in the experimental market; this points at the robustness of the psychological construct independent of the context in which the questions are asked (Biais et al., 2005).

In their stock market experiment Deaves, Lüders, and Luo (2004) are testing for premises that overconfidence leads to an increase in trading activity, and that gender influences trading activity through differences in overconfidence. They measure overconfidence of their subjects using a calibration based approach prior to conducting the experiment. Compared to the tests used in the other studies, their test contains more (up to 20) items. Each of the general knowledge questions in their test had a known numerical answer for which subjects had to state upper and lower bounds of 90% confidence interval in which the real answer would fall. Their choice of the non-economic questions is motivated by the attempt to avoid giving either group of participants a relative advantage because of subject content (Deaves et al., 2009).

Stotz and von Nitzsch (2005) in their paper investigate the extent of analyst overconfidence in their abilities to forecast prices and earnings. 112 bank analysts had to answer two questions: one asked them to rank their skills with regard to their price or earnings estimates in comparison to their colleagues, and another asked to estimate what percentage of analysts produce work superior/ inferior to them. Two types of coefficients, measuring subjects' overconfidence, were then calculated: overconfidence coefficients for earnings and in price targets. Results presented in their paper suggest, that overconfidence increases with an increasing perception of control.

Glaser and Weber (2007) asked a sample of approximately 3000 individual investors with online broker accounts to answer an online test, which enabled the authors to measure several manifestations of overconfidence: miscalibration, better than average effect, illusion of control, and unrealistic optimism. To measure miscalibration they asked subjects to state

upper and lower bounds of 90% confidence interval to the five economy-related questions<sup>5</sup> and five stock price predictions. Only 114 investors answered all their economy-related questions, and 165 – stock price prediction questions. By correlation the obtained measures of overconfidence and trading volume Glaser and Weber (2007) explored the connection between them.

Glaser, Langer, and Weber (2005) surveyed 123 professional traders and investment bankers, and compared results to a student control group in order to analyze whether professionals are prone to judgmental biases to the same degree as lay men. They measured overconfidence of their subjects by the means of four tasks: 1) subjects stated 90% confidence intervals for 20 knowledge questions (ten general knowledge and ten economics and finance knowledge); 2) subjects had to assess their performance in the knowledge task (how many right answers?) and assess own performance compared to others (how many right answers compared to the others?); 3) make 15 stock market forecasts by stating 90% confidence intervals, and 4) predict a trend in stock prices forecasting via confidence intervals. In most tasks the degree of overconfidence of professionals was significantly higher than of the student group.

To analyze the effect of professionalism on investment decisions Menkhoff, Schmeling, and Schmidt (2006) conducted a survey of approximately 500 subjects, consisting of professionals and lay men. Alongside with other aspects in their survey, they measured overconfidence via two questions on the “appropriate self-evaluation” (in other words better than average effect) in which subjects had to estimate their performance and information compared to the other investors. They find that, among other control variables, portfolio turnover is related to lower risk aversion and higher overconfidence.

Menkhoff, Schmidt and Brozynski (2006) surveyed 117 fund managers in order to detect an impact of experience on overconfidence, risk taking, and herding behavior. Their survey measures overconfidence via three tasks that enable assessment of the three manifestations of overconfidence: 1) evaluation of the own performance compared to the other fund managers (better than average effect), 2) 90% confidence estimation of the DAX index forecast (miscalibration), and 3) a third task is aimed at measurement subjects’ illusion of control (subjects are asked to rate the statement: economic news are not surprising to me). They find that experienced fund managers tend to exhibit herding behavior to a lesser extent than inexperienced ones; while evidence concerning the impact of experience on risk taking and

---

<sup>5</sup> However they call them “general knowledge” questions.

overconfidence is mixed: positive-self evaluation and illusion of control are increasing with experience, whereas miscalibration on the contrary decreases.

In their paper Oechssler, Schmidt, and Schnedler (2007) study whether bubbles can occur in the experimental markets that pay no dividends on assets. To measure overconfidence in their experiment they asked subjects, prior to each round, to rank themselves among the 60 subjects of a treatment in terms of payoff of that round. For each period the percentage of subjects who ranked themselves to be better than median (rank 30 or higher) was compared to the expected number of 50%. Overconfidence in their experiment was modest as merely 54% of the subjects thought to be better than the median. They have also constructed a second variable – “top-rank belief” – that measured, for each round, the number of subjects who thought they would be the best in terms of payoff. This construct has positive and significant effect on the probability of bubbles’ occurrence in their experimental market.

### 1.3.3 PROBLEM STATEMENT

This paper is aimed at the development of the instrument (test) that would enable construction of the comprehensive measure of individual overconfidence for the use in two economic experiments dealing with: 1) the role of overconfidence in occurrence of stock-prices’ bubbles, and 2) impact of overconfidence and risk aversion on economic behavior of individual traders. A well-designed instrument will allow assessing differences of the subjects with respect to their overconfidence and minimize the measurement error.

In my opinion previous works have drawbacks in the way they measure overconfidence, and thus in prior experiments overconfidence might have been caused (to some extent) by other reasons than the imperfection of human nature, i.e. by the mistakes in the tests’/ tasks’ construction<sup>6</sup>. Thus development of the overconfidence test was conducted with the following assumptions in mind:

From the review above one can see that overconfidence in financial settings is estimated either with the help of some assignments (e.g. estimate what percentage of analysts produce work superior to you?) or by the means of interval elicitation tests. However, overconfidence is often assessed based on the insufficient number of assignments or test items. Thus it raises doubts that these instruments actually offer a comprehensive measure of overconfidence of an individual. This fact is mentioned in the work of Menkhoff et al. (2006), who measured overconfidence with three assignments; Barber and Odean (2002) use only two assignments.

---

<sup>6</sup> Not to mention the studies in which overconfidence was never measured directly. For a review see Glaser and Weber (2007).

In comparison, the psychological studies of overconfidence use the amount of items that is much higher, and the minimum number of items for a reliable test is ten (Kline, 1993).

Most of the foregoing researchers followed the famous work by Russo and Schoemaker (1992) and used interval elicitation tests to assess overconfidence. However, interval estimation tasks are prone to produce extreme overconfidence (see e.g. Klayman et al., 1999). One reason to that is that subjects do not really understand the nature of these intervals and “what they are being asked to come up with” (Deaves et al, 2004). Also use of these instruments to measure the improvements in calibration, when the test is conducted before and after the experiment, is useless as this method does not allow for the improvement in calibration after training sessions; on the other hand subjective probability elicitation for the discrete items, combined with financial reward, can be improved (Lichtenstein and Fischhoff, 1980).

Findings from psychological research show that overconfidence is the most pronounced for the hard questions (few people know the right answer) and the least pronounced for the easy ones (most of the people know the correct answer). However, the abovementioned papers did not make use of the balanced to hard-easy effect tests. This could have artificially created high levels of under- or overconfidence, e.g. in the experiment of Deaves et al. (2009) none of the subjects got close to the perfect calibration measure, and even the best calibrated participants exhibited rather high degrees of overconfidence.

Connected to the hard-easy effect are country and gender biases. Country bias rests on the fact, that some questions might be easy in one country, but in another one they might be hard. Gender bias is produced by the choice of questions for the test that could be easier for men than women (e.g. sports, masculine hobbies) and vice versa. This could result in the inappropriate levels of under- or overconfidence for one gender compared to the other, or in one country compared to the other. Nevertheless, previous authors used tests that were not country or gender balanced, e.g. Deaves et al. (2009), used the same test in several locations. Finally, many of the tasks and tests discussed above were either not administered (e.g. Glaser and Weber (2007) conducted their survey via internet, and subjects might have used other sources than their own knowledge for answering the test), or were not (financially) rewarded.

Based on the abovementioned analysis, the developed instrument for measurement of overconfidence in the planned stock market experiments will differ in some important respects. First, another test format is chosen, namely multiple choice discrete propositions’ task format, which, due to its simplicity, is clearer to subjects and not inherently prone to production of extreme overconfidence levels. Second, a pilot test is conducted to assess questions’ difficulty and to single out easy, medium and hard questions. Then an equal

number of questions of the three difficulty levels are included in the final test. Third, in construction of the test it is controlled for possible country and gender biases, e.g. I have tried to avoid questions that might be easier to one gender than the other. Fourth, to check if the categorization into three difficulty levels and controlling for gender bias is successful, final instrument is pre-tested with the target group of students, namely those who are enrolled in different disciplines of the social sciences. Fifth, overconfidence measurement phase of the experiment is administered and financially rewarded. Moreover it is rewarded on the competition in the test accuracy basis, which should discourage sharing the results among students and thus increase the reliability of the measurement. And finally, compared to some of the authors, my test is expanded to include more questions.

## 1.4 METHOD

### *Procedure and Subjects*

A pilot test, whose purpose was to select questions for the final questionnaire, was conducted on the 19<sup>th</sup> May, 2008 at Christian-Albrechts University of Kiel. Subjects were given approximately 30 minutes time to fill in the 50 questions test at the end of the lecture on Social Politics. Three monetary prizes were offered for those participants who got the most questions right. A reward on the basis of competition in test accuracy was chosen in order to decrease the desire of subjects to share answers, and thus increase reliability of the obtained individual bias scores. Only fully completed tests were considered for the prize. A total of 96 tests were completed, of them 44 by males, and 52 by females. Most of the students were German (91 subjects). After the initial analysis 12 partially incomplete tests were not included in the further analysis. From the remaining 84 tests 50 were chosen randomly – 25 of men, and 25 of women. Participants of the test aged from 20 to 29 years ( $M = 24.32$ ,  $SD = 0.31$ ), and have studied from 3 to 11 semesters ( $M = 6.98$ ,  $SD = 2.11$ ). All participants were students of social sciences; of them 40% studied management, 38% were economics students, and 22% subjects were enrolled into other social studies. Average age of male subjects was 24.48 years ( $SD = 2.43$ ), and their average duration of study was 7 semesters ( $SD = 2.27$ ). Average age of female subjects was 24.16 years ( $SD = 1.97$ ), and their average duration of study was 6.96 semesters ( $SD = 1.99$ ). For information about participants' age and duration of studies refer to Appendix 1.D.

### *Design and Materials*

For the pilot test 50 general knowledge questions were selected from the German quiz webpage <http://wissen.de>. Questions on this web page have four short (one or two-word) multiple

exclusive answers. In the test, only three possible answers to each question were left, as one of the choices would usually be clearly incorrect. In choosing test questions I have tried to avoid the gender bias, which could result in inappropriate levels of under- or overconfidence for one gender, i.e. no questions that could be easier for men than women (e.g. sports, masculine hobbies) and vice versa were chosen. In the test students were asked to answer each of the 50 questions, and state their level of confidence in the correctness of their answer. Any number between 33% and 100% could be used to express subjects' confidence, where 33% meant that subjects did not know the correct answer, and were guessing, and 100% corresponded to being absolutely certain that the answer was correct.

In addition to measuring how well the subjects were calibrated, some personal data were collected: name, age, educational background, duration of studies, and nationality. In the final test students could also mark if they wanted to take part in the further experiments and, if answer was positive, submit their email. At the beginning of the pilot participants were informed that their personal data would be treated confidentially, and their identities would be used by the experimenter only for the purposes of determining the three winners. Thus, subjects' identity was revealed to other students only in the case of being one of the winners of the quiz, which was an honor to students. Test's instructions and design are based on the samples that were obtained from Dr. Briony Pulford (University of Leicester, School of Psychology) and Dr. Sabina Kleitman (University of Sidney, School of Psychology).

Based on the analysis of the pilot-test outcomes, a final test (*test-18*) was constructed from the 18 questions of the three difficulty levels: six hard, six medium difficulty, and six easy questions. Items were differentiated according to their difficulty on the basis of the number of correct answers to each of them from the whole group that participated in the pilot study. This methodology is described in the article of Pulford and Colman (1997), who suggest assigning questions to three difficulty categories, based on the total accuracy of the group in answering each question: 0-33% accuracy - hard questions, 34-66% - moderate difficulty, 67-100% easy questions. After the initial division, four questions have fallen in the category of hard questions (average accuracy 17.5%), 10 questions into category of medium difficulty questions (average accuracy 55.2%), and 36 of 50 questions turned to be easy (average accuracy 88.5%). As the category of hard-questions had not enough items, based on the idea that overconfidence is the most pronounced for hard questions (see Clarke, 1960; and Pitz, 1974), average overconfidence ratio over each of the medium difficulty questions was calculated and the two, having the highest overconfidence coefficient, were chosen to be included into hard-questions category. Thus six hard questions rather than four were obtained.

Characteristics of the final test in terms of the confidence, accuracy and the bias score are presented in the Table 1.1. Translation of *test-18* and instructions are included in Appendix 1.H.

Table 1.1: Average confidence, accuracy and bias score for the three levels of question difficulty of the final overconfidence test (*test-18*)

	Hard		Medium		Easy	
	M	SD.	M	SD	M	SD
Confidence	67.90	6.64	65.01	9.01	97.43	2.12
Accuracy	26.00	16.00	62.33	2.34	100.00	0.00
Overconfidence	41.90	18.24	2.68	7.48	-2.57	2.12

## 1.5 RESULTS

Consistent with previous research, on average, subjects have proved to be overconfident: the bias score of the group on the *test-50* pointed at slight overconfidence ( $M = 4.47$ ,  $SD = 7.34$ ); recalculation of the bias score for the *test-18*<sup>7</sup> has increased the average overconfidence measure ( $M = 14.11$ ,  $SD = 10.63$ ). Appendix 1.F presents data on the bias score and accuracy of all participants who took part in the pilot test for both *test-50* and *test-18*, and men and women separately. Average overconfidence of men for *test-50* is 3.33 ( $SD = 5.96$ ), and for *test-18* it is 14.11 ( $SD = 10.70$ ). Average overconfidence of women for *test-50* is 5.63 ( $SD = 8.47$ ), and for *test-18* it is 14.12 ( $SD = 10.79$ ). Noteworthy is the fact, that whereas for the complete *test-50* average overconfidence of men was slightly lower than that of women, after recalculating the overconfidence ratio for the questions chosen to comprise the final test (*test-18*), average bias score for both groups practically equalized. For the *test-50* correlation between accuracy and the bias score is found to be strong and significant, pointing at the decrease in overconfidence with the increase in accuracy (Pearson correlation (48) = -0.629,  $p < 0.01$ , one-sided); for the *test-18* this relationship is even stronger (Pearson correlation (48) = -0.823,  $p < 0.01$ , one-sided).

### *Overconfidence and experience*

After obtaining the bias score for each individual participant of the pilot test, a check of the proposition that overconfidence of subjects changes with experience was conducted. The two

---

<sup>7</sup> Same subjects.

variables that were used as a proxy of subjects' experience are age and duration of study at the university, measured in semesters. From the graphical analysis of the scatterplots (Appendix 1.B) no conclusions about linear relationship between the measures of experience and the bias score could be drawn for both *test-50* and *test-18*. Pearson's correlation analysis also has not detected any significant linear relationship between the variables of interest (see Appendix 1.C). Based on these findings, I conclude that students of different age groups and being at different levels of progress with their studies can be recruited for the participation at the planned economic experiment.

*Test-50 vs. Test-18: Accuracy and Confidence*

Analysis of the accuracy of the group for *test-50* revealed that even 72% of the questions have fallen in the category of easy questions (67-100% accuracy). See Figure 1.1(a). This test is distinguished by high precision, and inadequate to that precision confidence, consequently 58% of questions resulted in average underconfidence (see Figure 1.1(b). Appendix 1.A (a) confirms, that the distribution of accuracy per question for the *test-50* has more mass on the right tail (skewness = -1.31), and the distribution of overconfidence per question is left-skewed (skewness = 1.86). This example illustrates the dangers of using the unbalanced to hard-easy effect test in economic research: by using *test-50* one can artificially create high levels of underconfidence in ones subjects<sup>8</sup>.

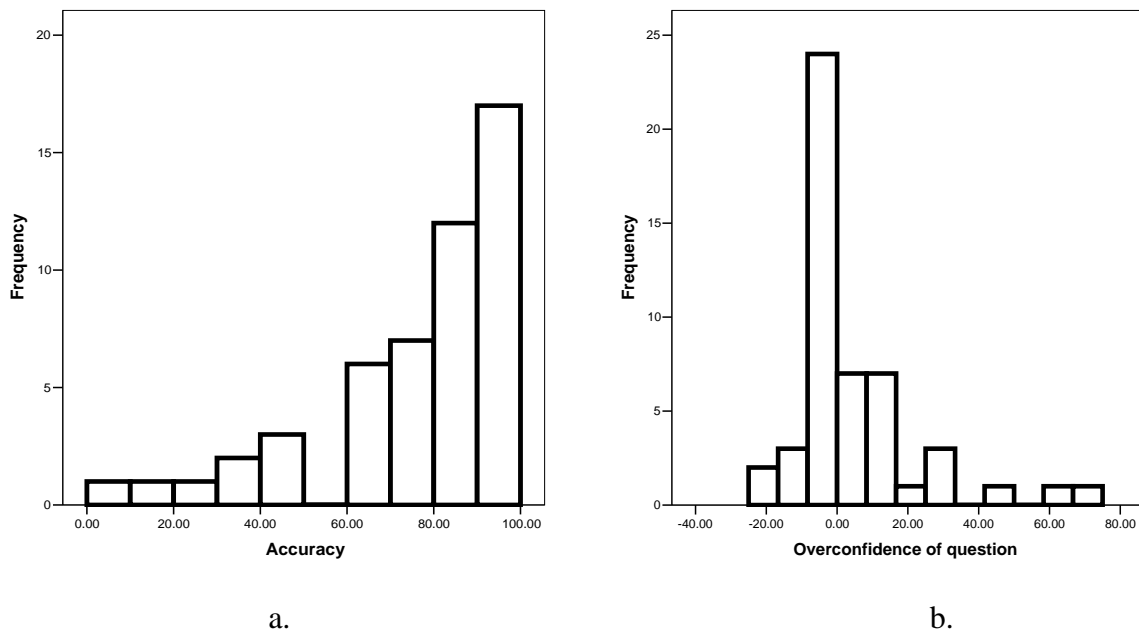


Figure 1.1: Distribution of accuracy (a) and overconfidence per question (b) in *test-50*

<sup>8</sup> Tests skewed in the direction of hard questions, can artificially create group overconfidence.



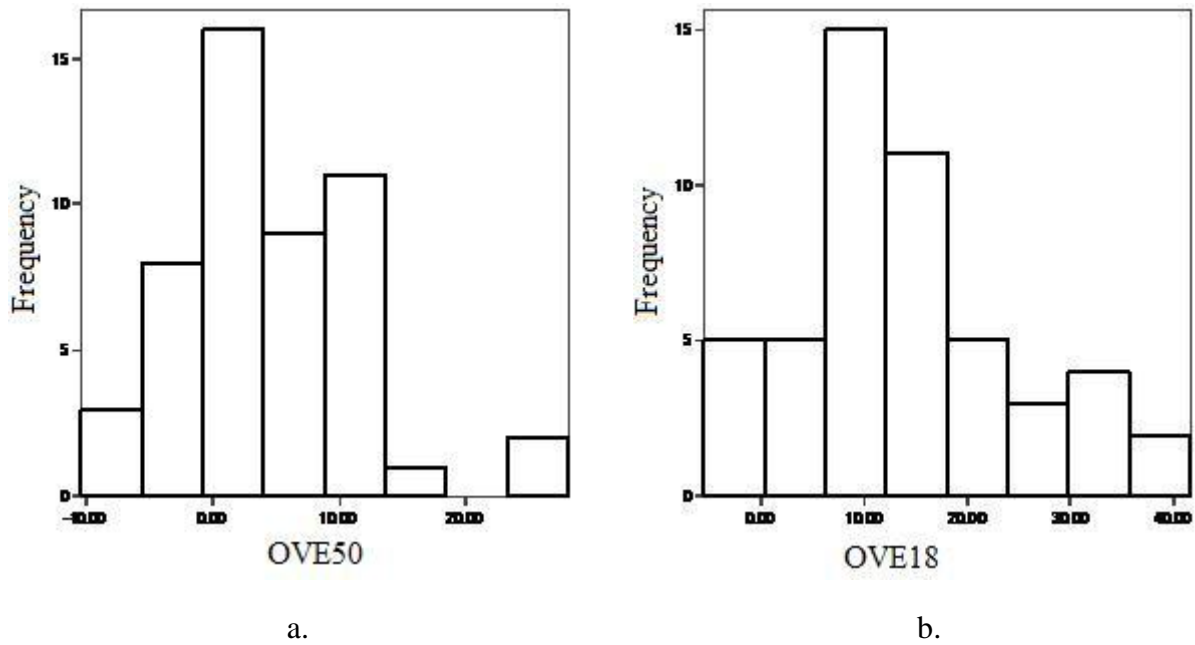


Figure 1.2: Distribution of the bias score of participants per test: a. *test-50* and b. *test-18*

24% of subjects who completed *test-50* were found to be underconfident (see Figure 1.2(a); for *test-18* this number decreases to 8% (see Figure 1.2(b). Alongside with the decrease in the percentage of underconfident subjects, an increase in the range of the bias score of the participants is observed (from 38.60 to 47.23). There is also improvement in the symmetry of the distribution of the bias score (*test-50*: skewness = 0.73; *test-18*: skewness = 0.53). See Appendix 1.A (b). The increase in bias score range is important for the future experiments as it leaves more room for finding subjects whose degree of overconfidence differs significantly.

### 1.5.1 STATISTICAL TESTS

In this section results of the statistical tests are presented that verify the success of categorization of the questions into three levels of difficulty for the *test-18*, and provide a sufficient basis to conclude that overconfidence is a robust phenomenon and not an artifact (Bar-Tal et al., 2001).

#### *Confidence*

I start by analyzing differences in the confidence levels of the subjects for the three difficulty levels of questions. On average subjects had the highest confidence for answering easy questions – 97.43%; the average confidence level for the hard questions was 67.90%, and for the medium questions – 65.01%. The performed Kruskal-Wallis H Test showed that the three levels of question difficulty resulted in significantly different from each other confidence levels (Chi-square(2) = 11.617,  $p < 0.01$ ), pointing out, that at least two of the three difficulty

levels were characterized by unequal confidence levels. Effect size is  $\eta^2 = 0.856$ , which means that 86% of the variance in the confidence assessments is due to the difficulty of the questions. Pairwise Mann-Whitney U tests revealed that there is no significant difference in the confidence for the medium and hard questions ( $U = 13.500$ ,  $p = 0.470$ , two-sided); confidence in answering easy questions is significantly higher than in answering medium ( $U = 0.00$ ,  $p < 0.01$ , one-sided) and hard questions ( $U = 0.00$ ,  $p < 0.01$ , one-sided).

### *Accuracy*

Average accuracy levels for the *test-18* were: 26% for hard questions, 62.33% for medium, and 100% for easy questions. Kruskal-Wallis H Test indicates that the difficulty level of questions significantly affected accuracy of the answers (Chi-Square (2) = 15.760,  $p < 0.001$ , two-sided); effect size is  $\eta^2 = 0.926$ . A series of Mann Whitney U tests were carried out. These tests show that there is a significant difference in accuracy for answering three categories of questions: medium questions tend to outperform in accuracy hard questions ( $U = 0.00$ ,  $p < 0.01$ , one-sided); accuracy for answering easy questions significantly exceeds the accuracy of medium ( $U = 0.00$ ,  $p < 0.01$ , one-sided), and hard questions ( $U = 0.00$ ,  $p < 0.01$ , one-sided). These results prove that the division of questions into three difficulty levels is successful.

### *Overconfidence*

Table 1.1 demonstrates that participants exhibit overconfidence for two levels of question difficulty (hard questions: BS = 41.90; medium questions: BS = 2.68) and underconfidence for the third one (easy questions: BS = -2.57). This is in line with the previous research that found hard questions to be the most prone to overconfidence, and easy questions to be often subject to underconfidence. The bias scores for easy and hard questions differ significantly from zero (easy questions: Wilcoxon signed rank test  $T = 2.097$ ,  $p < 0.05$ , two-sided; hard questions: Wilcoxon  $T = 2.097$ ,  $p < 0.05$ , two-sided). However, for the medium difficulty questions the null hypothesis of the equality of the bias score to zero cannot be rejected (Wilcoxon  $T = 0.419$ ,  $p = 0.675$ , two-sided). It can be concluded that the medium difficulty questions produced on average the bias score which was the most indistinguishable from the perfect calibration score of zero. To examine the existence of the hard-easy effect I first test a joint hypothesis of the equality of the levels of overconfidence generated by three levels of questions' difficulty versus the alternative, that some difficulty levels produced more overconfidence than the others. The null hypothesis is rejected at a high level of significance (Chi-Square(2) = 12.117,  $p < 0.01$ ). Effect size is  $\eta^2 = 0.783$ . Mann-Whitney U test, performed on each pair of the three levels of the bias score, confirmed the existence of the hard-easy effect. Subjects showed significantly higher

overconfidence for the hard questions than for the medium (Mann-Whitney  $U = 0.00$ ,  $p < 0.01$ , one-sided) and easy ones (Mann-Whitney  $U = 0.00$ ,  $p < 0.01$ , one-sided); overconfidence for the medium questions was slightly higher than for the easy questions (Mann-Whitney  $U = 9.50$ ,  $p < 0.1$ , one-sided).

### *Gender Differences*

#### *Test-50:*

Males were slightly less overconfident than females for the *test-50* (men BS:  $M = 3.33$ ,  $SD = 5.96$ ; women BS:  $M = 5.63$ ,  $SD = 8.47$ ), however this difference was not significant ( $t(48) = -1.109$ ,  $p = 0.27$ , two-sided; effect size  $\eta^2 = 0.025$ ) (see Appendix 1.F). Male subjects achieved higher accuracy for the *test-50* than female subjects (men:  $M = 78.80$ ,  $SD = 5.45$ ; women:  $M = 73.52$ ,  $SD = 6.72$ ), and this difference is significant ( $t(48) = 3.053$ ,  $p < 0.01$ , two-sided). Effect size is  $\eta^2 = 0.163$ , which points out that 16.3% of the variance in accuracy was gender dependent. Male subjects have also shown higher confidence in answering questions of *test-50*, than female subjects (men:  $M = 82.13$ ,  $SD = 4.93$ ; women:  $M = 79.07$ ,  $SD = 6.70$ ), this difference is found to be significant ( $t(48) = 1.840$ ,  $p < 0.05$ , one-sided); effect size  $\eta^2 = 0.069$ . The fact that about 16% of variation in accuracy and 7% in confidence is gender dependent is not satisfactory because there is more gender bias in the overconfidence test than it was expected. Correlation between overconfidence and accuracy is strong and significant for both genders (men: Pearson's Correlation (23) =  $-0.630$ ,  $p < 0.01$ , one-sided; women: Pearson's Correlation (23) =  $-0.625$ ,  $p < 0.01$ , one-sided).

#### *Test-18:*

Both genders have shown almost equal overconfidence for the *test-18* (male BS:  $M = 14.11$ ,  $SD = 10.70$ ; female BS:  $M = 14.12$ ,  $SD = 10.79$ ;  $t(48) = -0.002$ ,  $p = 0.998$ , two-sided; effect size  $\eta^2 = 0.00$ ) (see Appendix 1.F). Overconfidence for the *test-50* was significantly lower than for the *test-18*. Male subjects have slightly higher accomplishments in terms of accuracy than female subjects (men:  $M = 63.78$ ,  $SD = 9.64$ ; women:  $M = 61.78$ ,  $SD = 10.43$ ), although this difference is found to be insignificant ( $t(48) = 0.704$ ,  $p = 0.485$ , two-sided); effect size  $\eta^2 = 0.010$ . Male subjects were slightly more confident in answering questions of *test-18* (men:  $M = 77.40$ ,  $SD = 5.21$ ; women:  $M = 74.87$ ,  $SD = 5.20$ ), however this difference is insignificant ( $t(48) = 1.37$ ,  $p = 0.176$ , two-sided); effect size  $\eta^2 = 0.037$ . Compared to *test-50*, *test-18* has very low amount of variation in confidence and accuracy that is gender dependent. Test of the difference in overconfidence between men and women for the three levels of question difficulty has shown that both groups have expressed similar biases in answering the test and that the encountered differences were not significant (hard questions:  $t(48) = 0.085$ ,  $p =$

0.933, two-sided; medium questions:  $t(48) = 0.354$ ,  $p = 0.725$ , two-sided; easy questions:  $t(48) = 0.737$ ,  $p = 0.465$ , two-sided). Correlation between overconfidence and accuracy is strong and significant for both genders (men: Pearson's Correlation (23) = -0.847,  $p < 0.01$ , one-sided; women: Pearson's Correlation (23) = -0.810,  $p < 0.01$ , one-sided).

## 1.6 SECOND EXPERIMENT

To check if the results obtained by using *test-18* were replicable, namely the average group degree of overconfidence, the obtained categorization into three difficulty levels and controlling for gender bias, the experiment was repeated with the students of the target group: those enrolled into different disciplines of social sciences. In this subsection I will also estimate the reliability of my scale.

A second experiment was conducted on the 14<sup>th</sup> June, 2008 at Christian-Albrechts University of Kiel. Subjects were given approximately 15 minutes time to fill in the final, 18 questions, overconfidence test (*test-18*) at the end of the lecture on Economics of Risk and Uncertainty. As in the pilot test, three monetary prizes were offered for the participants who got the most questions right. A total of 37 tests were completed, of them 3 had no personal information and were not included in the further analysis. Participants of the test aged from 22 to 31 years ( $M = 26.06$ ,  $SD = 2.62$ ), and have studied on average 9.10 semesters ( $SD = 2.60$ ). Of the 34 participants 21 were males (age:  $M = 25.95$ ,  $SD = 2.64$ ), and 13 were females (age:  $M = 26.23$ ,  $SD = 2.68$ ). The majority of the subjects were Germans (86%). All participants were students of social sciences, of them 26 studied economics, seven studied management, and one other social sciences. For information about subjects' age and duration of studies refer to Appendix 1.E. Consistent with previous research, on average, subjects were prone to overconfidence ( $M = 10.41$ ,  $SD = 9.26$ ). Average group overconfidence on test-18 obtained from the experiment on the 6<sup>th</sup>, June and on the 19<sup>th</sup>, May did not significantly differ from each other ( $t(82) = 1.649$ ,  $p = 0.103$ , two-sided; size effect  $\eta^2 = 0.032$ ). Men on average were slightly more overconfident ( $M = 10.68$ ,  $SD = 9.81$ ) than women ( $M = 9.98$ ,  $SD = 8.68$ ), however this difference was found to be insignificant. Appendix 1.G presents data on the bias score of all participants who took part in the pilot, and men and women separately. Just as in the pilot test, correlation coefficient between age and overconfidence (Pearson coefficient (32) = 0.189,  $p = 0.142$ , one-sided), and semester and overconfidence (Pearson coefficient (32) = -0.054,  $p = 0.388$ , one-sided) is small and insignificant. Correlation between the accuracy and the bias score is strong and significant, pointing at the decrease in overconfidence with the increase in accuracy (Pearson correlation (332) = -0.731,  $p < 0.01$ ).

After division of questions into three difficulty levels characteristics of each category, in terms of confidence, accuracy and the bias score, were calculated (see Table 1.2).

Table 1.2: Characteristics of the three levels of question difficulty of the *test-18* from the experiment on 14.06.09

	Hard		Medium		Easy	
	M	SD	M	SD	M	SD
Accuracy	22.35	12.06	52.94	9.49	95.38	7.76
Confidence	55.75	7.67	53.95	11.67	93.52	8.74
BS	33.40	19.24	1.01	11.38	-1.86	2.42

### *Accuracy*

Subjects' average accuracy level for answering hard questions was 22.35%, 52.94% for medium and 95.38% for easy questions. Kruskal-Wallis H Test shows that the difficulty level of questions had significant impact on the accuracy of answers (Chi-Square (2) = 15.065,  $p < 0.01$ , one-sided); effect size  $\eta^2 = 0.920$ . Pairwise comparisons, performed using the Mann-Whitney U test, revealed that the accuracy for answering the medium difficulty questions significantly exceeds the accuracy for answering hard questions (U = 0.50,  $p < 0.01$ , one-sided); accuracy for answering easy questions significantly exceeds accuracy for medium questions (U = 0.00,  $p < 0.01$ , one-sided) and hard questions (U = 0.00,  $p < 0.01$ , one-sided).

### *Confidence*

On average subjects had the most confidence for answering easy questions 93.52% (SD = 8.74); confidence levels for the hard and medium difficulty questions were correspondingly 55.75% (SD = 7.67) and 53.95% (SD = 11.67). The Kruskal-Wallis H Test demonstrated that the three difficulty levels differed significantly from each other with regard to confidence (Chi-Square (2) = 12.158,  $p < 0.01$ , one-sided) (see Appendix 1.G). Effect size is  $\eta^2 = 0.824$ , which means that 82% of the variance in confidence assessments is due to the difficulty of questions. Pairwise comparison revealed that the confidence for easy questions was significantly higher than the confidence for medium and hard questions (both: Mann-Whitney U = 0.00,  $p < 0.01$ , one-sided); however there was no significant difference in the confidence levels for the hard and medium questions (Mann-Whitney U = 15.00,  $p = 1.00$ , two-sided).

### *Overconfidence*

The overconfidence level for the hard questions was the highest ( $M = 33.40$ ,  $SD = 19.24$ ), the medium difficulty questions produced almost no overconfidence ( $M = 1.01$ ,  $SD = 11.38$ ), whereas easy questions resulted on average in underconfidence ( $M = -1.86$ ,  $SD = 2.42$ ). The three difficulty levels of questions differed significantly in terms of the produced bias score (Chi-Square (2) = 9.079,  $p < 0.01$ , one-sided), size effect  $\eta^2 = 0.659$ . Mann-Whitney U analysis showed subjects showed significantly higher overconfidence for the hard questions than for the medium (Mann-Whitney U = 2.00,  $p < 0.01$ ) and easy questions (Mann-Whitney U = 0.00,  $p < 0.01$ ), which is in line with the previous research. Overconfidence levels for medium and easy questions, on average, were not significantly different from each other (Mann-Whitney U = 18.00,  $p = 0.334$ , one-sided). The bias score for the hard questions was significantly higher than zero (Wilcoxon Signed Rank T = 2.097,  $p < 0.05$ , one-sided); for the easy and medium difficulty questions the null hypothesis of the equality of the bias score to zero cannot be rejected (easy questions: Wilcoxon Signed Rank T = 1.606,  $p = 0.108$ , two-sided; medium questions: Wilcoxon Signed Rank T = 0.00,  $p = 1.00$ , two-sided). For this group of the participants, easy and medium difficulty questions produced on average the bias score which was the most indistinguishable from the perfect calibration score of zero.

### *Gender differences*

No significant difference between male and female participants in terms of overconfidence was found ( $t(32) = 0.211$ ,  $p = 0.834$ , two-sided; effect size  $\eta^2 = 0.001$ ) (see Appendix 1.G). Although men, on average, were less accurate than women this difference is not significant ( $t(32) = -0.524$ ,  $p = 0.604$ , two-sided; effect size  $\eta^2 = 0.009$ ). The difference in average confidence across all items of the test between male and female participants is insignificant ( $t(32) = -0.53$ ,  $p = 0.600$ , two-sided; effect size is  $\eta^2 = 0.009$ ). No significant difference in overconfidence is found between male and female subjects for the three levels of question difficulty (hard questions:  $t(32) = 0.042$ ,  $p = 0.967$ , two-sided; medium questions:  $t(32) = -0.357$ ,  $p = 0.723$ , two-sided; easy questions:  $t(32) = 1.468$ ,  $p = 0.152$ , two-sided). Correlation between overconfidence and accuracy is strong and significant for both genders (men: Pearson's Correlation (19) = -0.653,  $p < 0.01$ , one-sided; women: Pearson's Correlation (11) = -0.883,  $p < 0.01$ , one-sided).

### *Reliability*

According to DeCoster (2000), a scale can be called reliable (possess internal consistency) “if repeated measurements under the same circumstances tend to produce the same results”. A common way to estimate reliability of an instrument is to calculate Cronbach's alpha. Moss et

al (1993) state, that a generally acceptable value of coefficient alpha equals 0.6; however the more recognized threshold is 0.7. These values of alpha are considered to be optimal for the use in social research. For my instrument three values of alpha were estimated: alpha for the test confidence equaled 0.79, alpha for the test accuracy – 0.54, and alpha for the bias score – 0.68. Values of the calculated alphas were either close or exceeded the threshold values. A somewhat lower degree of alpha for the accuracy dimension resulted from low variance in answering easy questions. Easy questions cannot be removed from the test, in the desire to improve its reliability, as a good instrument should not only have a reasonable internal consistency (reliability) but also a “meaningful content coverage” (Schmitt, 1996). Based, on the calculated values of Cronbach’s alpha, it can be concluded that the developed instrument possesses good internal consistency (reliability).

## 1.7 CONCLUSIONS

In this paper results of the two experiments, aimed at the development of the instrument (test) that would enable the construction of the comprehensive measure of individual overconfidence, are presented. Desired instrument, which is to be used in economic overconfidence experiments, should allow assessment of the differences between the subjects with respect to their degree of overconfidence and minimize the measurement error.

After carrying out the analysis of some of the instruments used in foregoing experimental research, there were good reasons to suspect that overconfidence was previously measured inadequately. The principal steps needed to improve the instrument (test) were: 1) choice of another test-format (multiple choice discrete propositions’ tasks instead of confidence intervals estimation), 2) balancing the test for the hard-easy effect, and 3) controlling for gender and country bias. Instrument was obtained in a two-stage procedure in which a pilot test was used to assess questions’ difficulty, based on the groups’ accuracy in answering each of the initial 50 items. Then six questions of the three difficulty types (hard, medium, and easy) were selected for the final test. The second experimental phase was aimed at verification of replicability of results, namely of the average degree of group overconfidence, the obtained categorization into three difficulty levels and of controlling for the gender bias. Both experiments were conducted with the students enrolled into different disciplines of social sciences. The two experimental sessions were administered and subjects were offered a reward, on the basis of competition in test accuracy. As in previous experimental work, subjects on average proved to be overconfident.

Evidence was found for the significant effect of the question difficulty on the overconfidence measure and for the existence of the gender bias. Hard questions produced significantly higher levels of overconfidence than medium-difficulty and easy questions, which in turn resulted in underconfidence. Analysis of the groups' accuracy on answering initial test (*test-50*) revealed that even 72 percent of the questions fell in the category of easy questions. Thus, by using initial *test-50* to measure subjects' overconfidence, one would artificially create high levels of underconfidence in ones subjects. Statistical analysis confirmed that in both experimental sessions the three types of questions, that comprised the final test, significantly differed from each other in terms of the produced confidence, accuracy and overconfidence. This result verified the success of categorization of questions into three levels of difficulty in the overconfidence measurement instrument. Average group overconfidence measures on *test-18*, obtained from both experimental sessions, did not differ significantly from each other. Instrument's internal consistency (reliability), assessed as the value of the Cronbach's alpha, was found to be good and acceptable for the use in social research.

Combining all levels of questions' difficulty, both genders expressed overconfidence that did not differ significantly from each other. It can be concluded, that for the created instrument (*test-18*), gender is not associated with overconfidence: first, there were no significant differences between male and female subjects' bias scores and, second, no significant difference in overconfidence was found between male and female subjects for the three levels of question difficulty. There was also almost no variance in confidence and accuracy that was gender dependent. By contrast, for the initial instrument (*test-50*) as much as 16 percent of variance in accuracy and 7 percent of variance in confidence was explained by gender.

Based on the analysis of the data obtained from both phases of the instrument construction, and in the light of the importance of employment of a reliable measure to assess subjects' overconfidence for the validity of the results of economic experiments, it can be concluded that a better instrument was developed for the use in planned experiments, suitable for evaluation of individual differences in terms of the degree of overconfidence.

## REFERENCES

- Adams, J. K. (1957). A confidence scale defined in terms of expected percentages. *The American Journal of Psychology*, Vol. 70(3), p. 432-436.
- Adams, P. A., Adams, J. K., (1958), Training in confidence-judgments. *The American Journal of Psychology*, Vol. 71(4), p. 747-751.



- Allwood, C.M., Granhag, P. A., Jonsson, A. C., (2006), Child witnesses' metamemory realism. *Scandinavian Journal of Psychology*, Vol. 47(6), p 461-447.
- Alpert, M., Raiffa, H., (1982), A progress report on the training of probability assessors. In Kahneman, D., Slovic, P., Tversky, A. (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, p. 294-305.
- Arkes, H. R., Christensen, C., Lai, C., Blumer, C., (1987), Two methods of reducing overconfidence. *Organizational Behavior and Human Decision Processes*, Vol. 39, p. 133-144.
- Barber, B., Odean, T., (2001), Boys will be boys: Gender, overconfidence and common stocks investments. *Quarterly Journal of Economics*, Vol. 116, p. 261-292.
- Barber, B. M., Odean, T., (2002), Online investors: do the slow die first? *Review of Financial Studies*, Vol. 15(2), p. 455-487.
- Bar-Hillel, M., (2001), Subjective probability judgments. In Smelser, N.J. and Baltes, D.B. (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*, Amsterdam: Elsevier Science Ltd., p. 15247-15251.
- Bar-Tal Y., Sarid A., Kishon-Rabin L., (2001), A test of overconfidence phenomenon using audio signals. *The Journal of General Psychology*, Vol. 128(1), p. 76-80.
- Baron, R. A., (2000), Psychological Perspectives on Entrepreneurship: Cognitive and Social Factors in Entrepreneurs' Success. *Current Directions in Psychological Science*, Vol. 9(1), p. 15-18.
- Benos, A., (1998), Aggressiveness and Survival of Overconfident Traders. *Journal of Financial Markets*, Vol. 1, p. 353-383.
- Bernard V. L., Thomas J., (1989), Post-earnings-announcement drift: delayed price response or risk premium. *Journal of Accounting Research*, Vol. 27, p. 1-36.
- Bernard, V. L., Thomas J., (1990), Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics*, Vol. 13(4), p. 305-340.
- Biais, B., Hilton, D., Mazurier, K., Pouget, S., (2005), Judgmental overconfidence, self-monitoring and trading performance in an experimental financial market. *Review of economic studies*, Vol. 72(2), p. 287-312.
- Blavatsky, P., (2009), Betting on own knowledge: Experimental test of overconfidence. *Journal of Risk and Uncertainty*, Vol. 38, p. 39-49.

- Bohner, G., Rank, S., Reinhard, M. A., Einwiller, S., Erb, H. P., (1998), Motivational determinants of systematic processing: expectancy moderates effects of desired confidence on processing effort. *European Journal of Social Psychology*, Vol. 28, p. 185-206
- Caballé, J., Sákovics J., (2003), Speculating against an overconfident market. *Journal of Financial Markets*, Vol. 6, p. 199-225.
- Cambridge, R. M., Shreckengost, R. C., (1978), Are you sure? The subjective probability assessment test. Langley, VA: Office of Training, Central Intelligence Agency.
- Camerer, C., (1995), Individual decision making. In Kagel, J. H., Roth, A. (Eds.), *Handbook of experimental economics*, Princeton University press, p. 587-703.
- Clarke, F. R., (1960), Confidence Ratings, Second-Choice Responses, and Confusion Matrices in Intelligibility Tests. *Journal of the Acoustical Society of America*, Vol. 32(1), p. 35-46.
- Conger, R. F., Wolstein, Ch. R., (2004), Managing overconfidence in pricing. *Emphasis*, Vol. 2, p. 10-13.
- Daniel, K., Hirshleifer D., Subrahmanyam A., (1998), Investor psychology and security market under-and overreactions. *Journal of Finance*, Vol. 53(6), p 1839-1885.
- Daniel, K. D, Hirshleifer, D., Subrahmanyam, A., (2001), Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance*, Vol. 56 (3), p. 921–965.
- Deaves, R., Lüders, E., Luo, G. Y., (2009), An Experimental Test of the Impact of Overconfidence and Gender on Trading activity. *Review of finance*, Vol. 13(3), p. 555–575.
- DeCoster, J., (2005), Scale Construction Notes. Retrieved on 04.30.2010 from <http://www.stat-help.com/notes.html>
- De Bondt, W. F. M., Thaler R., (1984), Does the Stock Market Overreact? *Journal of Finance*, Vol. 40(3), p. 793-805.
- De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., (1991), The Survival of Noise Traders in Financial Markets. *The Journal of Business*, Vol. 64(1), p. 1-19.
- Fenton-O’Creevy, M., Nicholson, N., Soane, E., Willman, P., (2003), Trading on Illusions: Unrealistic Perceptions of Control and Trading Performance. *Journal of Occupational and Organizational Psychology*, Vol. 76, p. 53–68.

Fischhoff, B., Slovic, P., Lichtenstein, S., (1977), Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 3, p. 552-564.

Fischhoff, B., (1982), Debiasing. In Kahneman, D., Slovic, P., Tversky, A. (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, p. 422-444.

Gervais, S., Odean, T., (2001), Learning to be Overconfident. *Review of Financial Studies*, Oxford University Press for Society for Financial Studies, Vol. 14(1), p. 1-27.

Gigerenzer, G., Hoffrage, U., Kleinbölting, H., (1991), Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, Vol. 98, p. 506-528.

Glaser, M., Langer, T., Weber, M., (2005), Overconfidence of Professionals and Lay Men: Individual Differences within and between Tasks? Working Paper, Mannheim.

Glaser, M., Weber, M., (2007), Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, Vol. 32(1), p. 1-36.

Hazard, T. H., Peterson, C. R., (1973), Odds versus probabilities for categorical events. Tech. Rep. No. 73-2, McLean, VA: Decisions and Designs, Inc.

Hirshleifer D., Subrahmanyam, A., Titman S., (1994), Security analysis and trading patterns when some investors receive information before others. *Journal of Finance*, Vol. 49(5), p. 1665-1698.

Hoelzl, E., Rustichini, A., (2005), Overconfident: do you put your money on it? *Economic Journal*, Vol. 115, p. 305-318.

Hynes, M., Vanmarcke, E., (1976), Reliability of embankment performance predictions. *Proceedings of the ASCE Engineering Mechanics Division Specialty Conference*, Waterloo, Ontario, Canada: University of Waterloo Press.

Johnson, D. D. P., McDermott, R., Barrett E. S., Cowden, J., Wrangham R., McIntyre, M.H., Rosen, S. P., (2006), Overconfidence in wargames: experimental evidence on expectations, aggression, gender and testosterone. *Proceedings: Biological Sciences*, Vol. 273(1600), p. 2513-2520.

Kahneman D., Riepe, M. W., (1998), Aspects of Investor Psychology. *Journal of Portfolio Management*, Vol. 24(4), p. 52-65.

Keasey, K., Watson, R., (1989), Consensus and accuracy in accounting studies of decision making: A note on a new measure of consensus. *Accounting, Organizations and Society*, Vol. 14, p. 337-345.

- Kim, A. K., Nofsinger, J. R., (2003), The behavior and performance of individual investors in Japan. Working paper.
- Kirchler, E., Maciejovsky, B., (2002), Simultaneous over- and underconfidence: evidence from experimental asset markets. *Journal of Risk and Uncertainty*, Springer, Vol. 25(1), p. 65-85.
- Klayman, J., Soll, J. B., Gonzáles-Vallejo, C., Barlas, S., (1999), Overconfidence: it depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, Vol. 79(3), p. 216-247.
- Koriat A, Lichtenstein S, Fischhoff B., (1980), Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, Vol. 6, p. 107-18.
- Kyle, A., Wang, F. A., (1997), Speculation duopoly with agreement to disagree: can overconfidence survive the market test? *Journal of Finance*, Vol. 52(5), p. 2073-2090.
- Lakonishok, J., Shleifer, A., Vishny, R. W., (1992), The impact of institutional trading on stock prices. *Journal of Financial Economics*, Vol. 32(1), p. 23-43.
- Langer, E. J., (1982), The illusion of control. In: Kahneman, D., Slovic, P., Tversky A. (Eds.), *Judgment under uncertainty: heuristics and biases*, Cambridge University Press, p. 231-338.
- Loughran, T., Ritter, J. R., (1995), The new issues puzzle. *Journal of Finance*, Vol. 50(1), p. 23-51.
- Lichtenstein, S., Fischhoff, B., (1980), Training for calibration. *Organizational Behavior and Human Performance*, Vol. 26, p. 149-171.
- Lichtenstein, S., Fischhoff, B., (1981), The effects of gender and instructions on calibration (Decision Research Report 81-5). Eugene OR: Decision Research.
- Lichtenstein, S., Fischhoff, B., and Phillips, L. D., (1977), Calibration of probabilities: The state of the art. In Jungermann, H., deZeeuw, G. (Eds.), *Decision making and change in human affairs*, Amsterdam: D. Reidel, p. 275-324.
- Lichtenstein, S., Fischhoff, B., Phillips, L. D., (1982), Calibration of probabilities: the state of the art to 1980. *Judgment under Uncertainty: Heuristics and Biases*. In Kahneman, D., Slovic, P., Tversky, A. (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, p. 306-334.
- Menkhoff, L., Schmidt, U., Brozynski, T., (2006), The impact of experience on risk taking, overconfidence, and herding of fund managers: Complementary survey evidence. *European Economic Review*, Vol. 50(7), p. 1753-1766

- Menkhoff, L., Schmeling, M., Schmidt, U., (2006), Does Professionalism Consistently Affect Portfolio Biases? Working Paper.
- Michaely, R., Thaler, R. H., Womack, K. L., (1995), Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift? *Journal of Finance*, Vol. 50(2), p. 573-608.
- Moore, P. G., (1977), The manager's struggle with uncertainty. *Journal of the Royal Statistical Society*, Vol. 140, p. 129-165.
- Moss, S., Patel, P., Prosser, H., Goldber, D., Simpson, N., Rowe, S., Lucchino, R., (1993), Psychiatric morbidity in older people with moderate and severe learning disability. I: Development and reliability of the patient interview (PAS-ADD). *British Journal of Psychiatry*, Vol. 163, p. 471-480.
- Neale, M. A., Bazerman, M. H., (1990), *Cognition and rationality in negotiation*. New York: The Free Press, p. 240.
- Nickerson, R. S., McGoldrick, C. C., (1965), Confidence ratings and level of performance on a judgmental task. *Perceptual & Motor Skills*, Vol. 20, p. 311-316.
- Nöth, M., Weber, M., (2003), Information Aggregation with Random Ordering: Cascades and Overconfidence. *The Economic Journal*, Vol. 113, p. 166-189.
- Oberlechner, T., Osler, C.L., (2003), Overconfidence in currency markets. Working Paper.
- Odean, T., (1998), Volume, volatility, price and profit when all traders are above average. *Journal of Finance*, Vol. 53(6), p. 1887 -1934.
- Odean, T., (1999), Do investors trade too much? *American Economic Review*, Vol. 89(5), p. 1278-1298.
- Oechssler, J., Schmidt, C., Schnedler, W., (2007), Asset bubbles without dividends - an experiment. Working Paper Nr. 07-01, University of Mannheim.
- Oskamp, S., (1962), The relationship of clinical experience and training methods to several criteria of clinical prediction. *Psychological Monographs*, Vol. 76(547), p. 1-28.
- Pitz, G.F., (1974), Subjective probability distributions for imperfectly known quantities. In Gregg, L. W. (Ed.), *Knowledge and Cognition*. New York: Wiley, p. 29-41.
- Powel, W. D., Bolich, C., (1993), Changing predictions based on knowledge of past performance in preschool and young grades-school. *Child Study Journal*, Vol. 23(3), p. 209.
- Pulford, D. B., (1996), *Overconfidence in Human Judgment*, PhD thesis. Department of Psychology, University of Leicester.

- Pulford, B. D., Colman, A. M., (1997), Overconfidence: Feedback and Item Difficulty Effects. *Personality and Individual Differences*, Vol. 23(1), p. 125-133.
- Ronis, D. L., Yates, J. F., (1987), Components of probability judgment accuracy: Individual consistency and effects of subject matter and assessment method. *Organizational Behavior and Human Decision Processes*, Vol. 40, p. 193-218.
- Russo, J. E., Schoemaker, P. J., (1992), Managing overconfidence. *Sloan Management Review*, Vol. 33, p. 7-17.
- Scheinkman, J. A., Xiong, W., (2003), Overconfidence and speculative bubbles. *Journal of Political Economy*, Vol. 111, p. 1183-1219.
- Shiller, R. J., (2000), *Irrational Exuberance*. Princeton, N.J.: Princeton Univ. Press, 344 p.
- Sieber, J. E., (1979), Confidence estimates on the correctness of constructed and multiple-choice responses. *Contemporary Educational Psychology*, Vol. 4, p. 272-287.
- Snizek, J. A., Paese, P. W., Switzer, F. S., (1990), The effect of choosing on confidence in choice. *Organizational Behavior and Human Decision Processes*, Vol. 46, p. 264-282.
- Staël von Holstein, C. A. S., (1972), Probabilistic forecasting: An experiment related to the stock market. *Organizational Behavior and Human Performance*, Vol. 8, p. 139-158.
- Statman, M., Thorley, S., Vorkink, K., (2006), Investor overconfidence and trading volume. *Review of Financial Studies*, Vol. 19, p. 1531–1565.
- Stotz, O., von Nitzsch, R., (2005), The perception of control and the levels of overconfidence: evidence from analyst earnings estimates and prices targets. *The Journal of Behavioral Finance*, Vol. 6(3), p. 121-128.
- Sümer, N., Özkan, T., Lajunen, T., (2006), Asymmetric relationship between driving and safety skills. *Accident Analysis and Prevention*, Vol. 38(4), p. 703-711.
- Svenson, O., (1981), Are we all less risky and more skilful than our fellow drivers? *Acta Psychologica*, Vol. 47(2), p. 143-148.
- Taylor, S. E., and Brown, J. D., (1988), Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, Vol. 103, p. 193-210.
- Tversky, A., Kahneman, D., (1982), Judgment under Uncertainty: Heuristics and Biases. In Kahneman, D., Slovic, P., Tversky, A. (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, p. 3-20.

Wagenaar, W. A., Keren, G., (1986), Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Maneini, and D. D. Woods (Eds.), *Intelligent decision support in process environments*, Berlin: Springer, p. 87-107.

Weinstein, N. D., (1980), Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, Vol. 39, p. 806-820.

Zakay, D., Glicksohn, J., (1992), Overconfidence in a multiple-choice test and its relationship to achievement. *Psychological Record*, Vol. 42(4), p 519-524.

APPENDIX 1.A:

a: Skewness of the accuracy parameters and overconfidence scores per question for the *test-50*

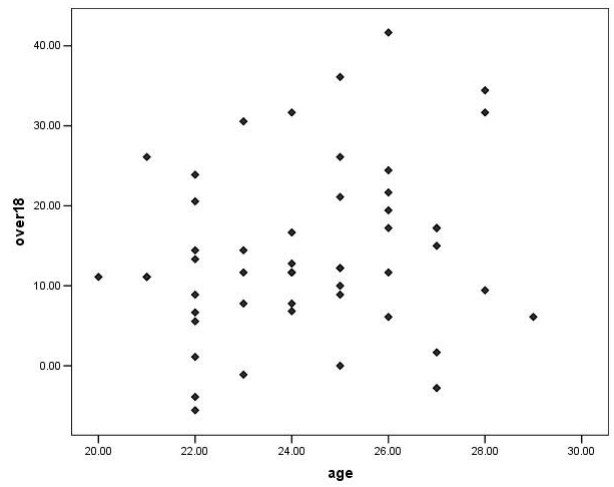
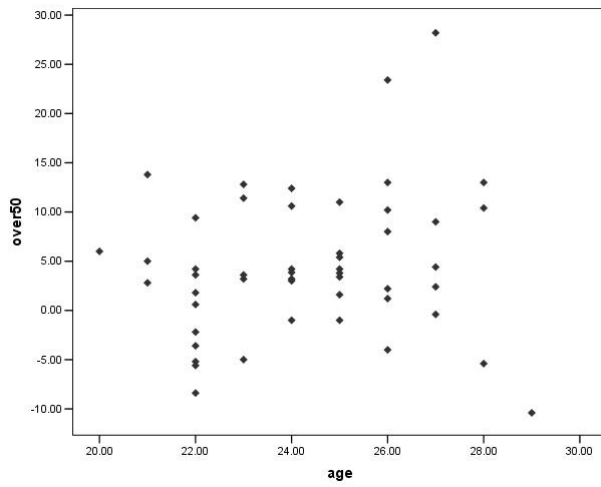
		Accuracy	BS
N		50	50
Skewness		-1.310	1.855
Std. err. of skewness		0.337	0.337
Range		92.00	91.40
Min		8.00	-22.80
Max		100.00	68.60
Percentiles:	25	63.50	-4.60
	50	84.00	-1.10
	75	94.50	11.60

b: Comparison of *test-50* and *test-18* in terms of overconfidence

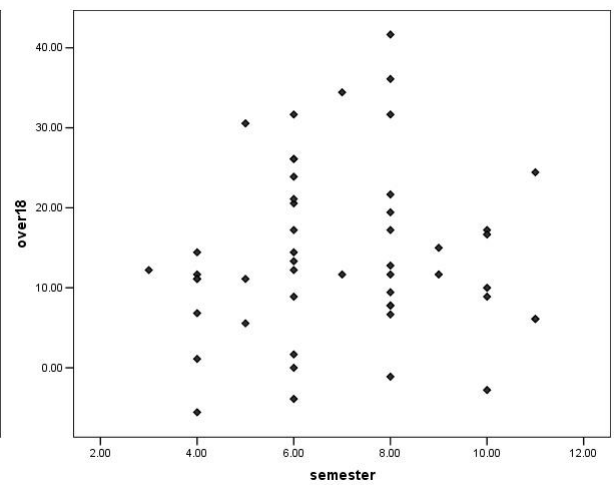
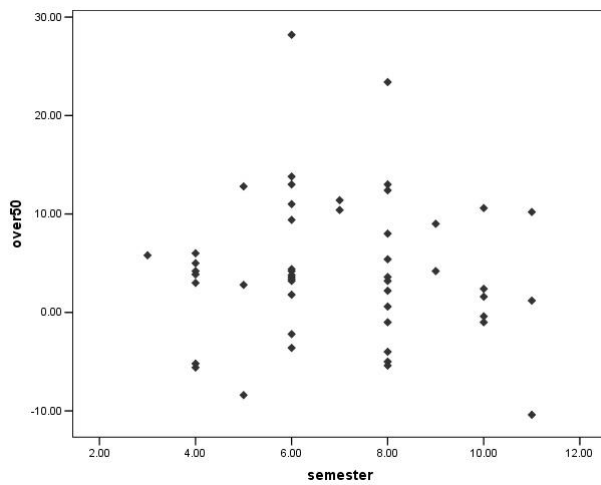
		BS50	BS18
N		50	50
Mean		4.47	14.11
SD		7.34	10.63
Skewness		0.726	0.525
Std. err. of skewness		0.337	0.337
Range		38.60	47.23
Min		-10.40	-5.56
Max		28.20	41.67
Percentiles:	25	0.35	7.54
	50	3.70	11.95
	75	9.60	20.70



APPENDIX 1.B: SCATTERGRAMS OF THE EXPERIENCE MEASURES RELATIONSHIP TO OVERCONFIDENCE (a. *test-50* and *test-18* age vs. bias score, and b. *test-50* and *test-18* semester vs. bias score).



a.



b.

APPENDIX 1.C: PEARSON'S TEST (DF. = 48) FOR CORRELATION RESULTS

		Semester	Age
OVE50	Correlation Coefficient	-0.045	0.148
	Sig. (one-sided)	0.377	0.152
OVE18	Correlation Coefficient	0.078	0.194
	Sig. (one-sided)	0.312	0.088

APPENDIX 1.D: AGE AND STUDY DURATION INFORMATION OF THE PILOT TEST ON 19.05.2008

	N	M	SD	Min	Max
Age	50	24.32	2.20	20	29
Male age	25	24.48	2.43	20	29
Female age	25	24.16	1.97	21	27
Semester	50	6.98	2.11	3	11
Male semester	25	7.00	2.27	3	11
Female semester	25	6.96	1.99	4	11

APPENDIX 1.E: AGE AND STUDY DURATION INFORMATION OF THE PILOT TEST ON 14.06.2008

	N	M	SD	Min	Max
Age	34	26.06	2.62	22	31
Male age	21	25.95	2.64	22	31
Female age	13	26.23	2.68	22	30
Semester	30	9.10	2.60	4	15
Male semester	20	9.10	2.97	4	15
Female semester	10	9.10	1.79	6	11

APPENDIX 1.F: INFORMATION ON OVERCONFIDENCE AND ACCURACY OF THE PARTICIPANTS OF  
PILOT ON 19.05.2008

*Overconfidence*

Pilot Test 50					
OBS	Group	Mean	SD	Mini	Max
50	All	4.48	7.34	-10.40	28.20
25	Female	5.63	8.47	-8.40	28.20
25	Male	3.33	5.96	-10.40	13.00
	Male vs. female diff.	-2.298 (0.273)			
Pilot Test 18					
OBS	Group	Mean	SD	Min	Max
50	All	14.11	10.63	-5.56	41.67
25	Female	14.12	10.79	-5.56	41.67
25	Male	14.11	10.70	-3.89	36.11
	Male vs. female diff.	-0.007 (0.998)			

APPENDIX 1.F - CONTINUATION:

*Accuracy*

Pilot Test 50					
OBS	Group	Mean	SD	Min	Max
50	All	76.16	6.61	58	90
25	Female	73.52	6.72	58	84
25	Male	78.80	5.45	66	90
	Male vs. female diff.	5.28 (0.004)			
Pilot Test 18					
OBS	Group	Mean	SD	Min	Max
50	All	62.78	9.99	38.89	83.33
25	Female	61.78	10.43	38.89	77.78
25	Male	63.78	9.64	44.44	83.33
	Male vs. female diff.	2.00 (0.485)			

APPENDIX 1.G: INFORMATION ON OVERCONFIDENCE AND ACCURACY OF THE PARTICIPANTS OF PILOT ON 14.06.2008

*Overconfidence*

OBS	Group	M	SD	Min	Max
34	All	10.41	9.26	-6.28	30.00
13	Female	9.98	8.68	-3.44	28.94
21	Male	10.68	9.81	-6.28	30.00
	Male vs. female diff.	0.700 (0.604)			

*Accuracy*

OBS	Group	M	SD	Min	Max
34	All	60.46	9.35	38.89	77.78
13	Female	61.54	9.48	38.89	77.78
21	Male	59.79	9.45	38.89	77.78
	Male vs. female diff.	-1.750 (0.834)			

### General Knowledge Questionnaire

Below you will be presented with some general knowledge questions. Imagine that you are taking part in a game, like “Trivial Pursuit” or “Who wants to be a Millionaire?”, and you have to choose the correct answer from the three given alternatives. A person who answers the most questions right will get a 30 EUR prize. The second place will be awarded by the 20 EUR prize, and the third place by 10 EUR. You will be paid next week!

1) Please circle **ONLY ONE** of three given answers. Only one of them is correct.

2) When you have made your choice and have circled your answer, we would like to know how sure/confident you are that your answer is correct. Since there are three alternative answers and only one of them is correct you have a 33% chance of giving a correct answer. Therefore 33% means that you are guessing and do not know the correct answer, and 100% corresponds to absolute certainty.

You can use any number between 33% and 100% to indicate your confidence that your answer is correct.

Enter your confidence for every answer in the gap in the question after every test item:

How confident are you that your answer is correct? \_\_\_\_\_ %

Please answer all questions. Even if you have to guess everything, you could answer 33% correct by chance. You are not allowed to consult anyone else, or copy the answers from somebody.

NOTE: Please answer all questions, one after another in order in which they are presented in the questionnaire. Guess any answers you do not know. Do not jump around the questions, and do not return to already answered questions to change your answers; we are interested in your first answer.

You will be paid the money only if you have filled in the **WHOLE** questionnaire! Don't leave unanswered questions or unfilled gaps!

Please ask questions if something is unclear to you.

Thank you for your patience in completing this questionnaire.

---

Your personal data will be treated confidentially.

Surname, Name: _____
Gender: _____
Age: _____
Nationality: _____
Field of Study: _____
Semester: _____

Would you like to participate in another experiment, in which you can also win money?

Yes <input type="checkbox"/>	No <input type="checkbox"/>
E-Mail: _____	

1.	How does one still call an instant camera? (circle one) Canon camera                      Polaroid camera                      Minolta camera
	How confident are you that your answer is correct? _____ %

2.	Where do flounders mostly live? (circle one) in coral reef                      dug on the ground                      in the reed
	How confident are you that your answer is correct? _____ %

3.	What does the <i>rollmops</i> consist of? (circle one) herring filet                      pork                      salmon filet
	How confident are you that your answer is correct? _____ %

4.	Which land does the Nobel Prize winner in Literature Gabriel García Márquez come from? (circle one) Colombia                      Spain                      Venezuela
	How confident are you that your answer is correct? _____ %

5.	Which style movement does anacreontics belong to? (circle one) Rococo                      Romanticism                      Realism
	How confident are you that your answer is correct? _____ %

6.	What is a hot chili sauce? (circle one) Tabasco                      Curacao                      Macao
	How confident are you that your answer is correct? _____ %

7.	How many letters does the Russian alphabet consist of? (circle one) 40                      33                      26
	How confident are you that your answer is correct? _____ %

8.	"Tosca" is an opera from ...? (circle one) G. Puccini                      G. Verdi                      A. Vivaldi
	How confident are you that your answer is correct? _____ %

9.	What s the name of the Greek Goddess of wisdom? (circle one) Pallas Athena                      Nike                      Penelope
	How confident are you that your answer is correct? _____ %



10.	Which is the most abundant metal on the Earth? (circle one) iron    aluminum    copper
	How confident are you that your answer is correct? _____ %

11.	How does one call an unknowing person? (circle one) Ignatius    ignorant    ideologue
	How confident are you that your answer is correct? _____ %

12.	Who flew for the first time with an airship around the Eiffel Tower? (circle one) Santos-Dumont    count Zeppelin    Saint-Exupéry
	How confident are you that your answer is correct? _____ %

13.	How is the snow shelter of Eskimos called? (circle one) wigwam    igloo    tipi
	How confident are you that your answer is correct? _____ %

14.	Which enterprise does Bill Gates belong to? (circle one) Intel    Microsoft    Dell Computers
	How confident are you that your answer is correct? _____ %

15.	How is the fasting month in Islam called? (circle one) Sharia    Ramadan    Imam
	How confident are you that your answer is correct? _____ %

16.	Which language does the concept "Fata Morgana" come from? (circle one) Italian    Arabic    Swahili
	How confident are you that your answer is correct? _____ %

17.	How many days does a hen need to incubate an egg? (circle one) 21 day    14 days    28 days
	How confident are you that your answer is correct? _____ %

18.	What is ascorbic acid? (circle one) apple vinegar    vitamin C    vitamin A
	How confident are you that your answer is correct? _____ %

## CHAPTER 2

# OVERCONFIDENCE AND BUBBLES IN EXPERIMENTAL ASSET MARKETS

## 2.1 INTRODUCTION

Many different factors are continuously contributing to the changes in stock prices. As a consequence stock-prices' bubbles might occur. Although different definitions of the stock price bubble notion exist, one thing is common to all of them: bubbles are deviations from the fundamental value of an asset. Fundamental asset value equals the present value of the stream of dividends that owner expects to receive, and therefore dividend is the only driving force of the asset prices. There exist several problems in determining the fundamental value of an asset, namely estimation of dividends on the asset through the time period, determination of the terminal asset value and discount rates for calculation of the present value. All these components can be controlled in the laboratory asset market.

A question arises, why people pay for an asset a price that differs from its fundamental value? According to Scheinkman and Xiong (2003) overconfidence is the main factor which makes people pay higher prices, than the underlying fundamental value of an asset. Overconfidence is one of the psychological characteristics, stipulating deviations from rational behavior. The concept of overconfidence is based on the large body of evidence from cognitive psychological research, which suggests that human-beings overestimate their own knowledge, abilities and precision of their personal information. Although the beginning of overconfidence research lies in psychological works, the effect of overconfidence on financial decision making, functioning of financial markets and economic outcomes is a widely researched topic in behavioral economics.

Most of the theoretical overconfidence papers are based on the initial assumption of traders' overconfidence, which is modelled as overestimation of the precision of private information that manifests itself via underestimation of the variance of the private signal that subjects get. Theoretical models of overconfidence predict that overconfidence causes excess trading volume and excess price volatility, as well it induces occurrence of the speculative price bubbles. There are a few empirical and experimental studies designed to test whether cognitive bias of overconfidence affects financial decisions, market outcomes and subjects' performance. Market experiments which are the closest in spirit to mine are by Biais, Hilton, Mazurier, and Pouget (2005), Kirchler and Maciejovsky (2002), Deaves et al. (2009). All these experiments analyzed relation between measures of overconfidence and trading behaviour, however only Deaves et al. (2009) explore the impact of overconfidence on the market-level. Kirchler and Maciejovsky (2002) run a multi-period experimental market and analyze development of overconfidence of the participants in the course of the experiment.

Their results indicate that participants of the experiment were well-calibrated in certain periods, and under- or overconfident in other periods. Biais et al (2005), use psychological questionnaire to measure the degree of overconfidence via interval estimation tasks in a group of 245 students. The main conclusion of the authors is that miscalibration does not lead to an increase in trading activity. On the contrary, Deaves et al. (2009) in their paper report that greater overconfidence leads to higher trading volume. They found no evidence that overconfidence and trading activity are gendered.

My experiment was constructed with the following assumptions in mind. First, previous experiments were not aimed at discovering the connection between the phenomenon of overconfidence and occurrence of stock-prices' bubbles. Second, there were no papers that previously used the suggested procedure of markets' formation, based on the participants' inborn level of overconfidence, and have managed directly connect changes in markets' overconfidence to the experimental outcomes. Third, previous experiments provided participants by private information with differences in signal quality, which itself creates potential for trade; in my experiment all subjects are given the same information. Fourth, to measure subjects' overconfidence I use a specially tailored test, weighted for the inclusion of easy, hard and medium difficulty questions, which is also gender-balanced; none of the previous experiments makes use of such test. However, unbalanced to hard-easy effect tests might artificially create high levels of overconfidence; the same is valid for gender bias. Fifth, I use the second construct to measure markets' overconfidence: a price-prediction task (in each period subjects submit their forecast of the next period's average market price and their confidence in this prediction). This design also enables following the evolution of market's overconfidence in the course of experiment. Both pre-experimental test and price prediction assignment are financially rewarded.

In this paper results of the experiment, designed to investigate the role of market overconfidence in the occurrence of bubbles in the asset prices and in the emergence of other stylized facts of the financial market (excessive trade, excessive price volatility), are reported. Additional interest is paid to the examination of the extent to which such relationship exists, i.e. determination of the linear relationship between price bubbles and the prevailing degree of market overconfidence, measured as the bias score. The design of the experiment follows Smith, Suchanek and Williams (1988) and is extended by a new feature, in which markets are constructed on the basis of subjects' overconfidence, assessed in pre-experimental studies. For the participation in the experiment two types of subjects are invited: those who have low bias score (rational subjects) and those who have high bias score (overconfident subjects). Of them

in the experiment two types of markets are formed: rational and overconfident. When there are no asymmetries in information and all traders have identical assets' and currency endowments, and all of them are "homogenous ... with statistically rational dividend and price expectations" (Gillette et al., 1999) a theory predicts that either no trading should occur or some marginal trading at the prices around the fundamental value. I assume that overconfident traders overestimate the probability of the occurrence of the maximum dividend value, thus they erroneously perceive possible future dividend income and optimistically overestimate the probability of existence of other traders ("greater fools") ready to pay for the asset an even higher price. This results in that the participants are taking excessive risk and trade at prices above the fundamental asset value. Thus bubbles in the asset's price occur. These bubbles usually burst several periods before the end of the experiment; research on overconfidence showed that overconfidence is decreasing with the task repetitiveness. Thus my second focus is to investigate changes in markets' overconfidence towards the end of the game.

Main findings from my experiment can be summarized as follows. In the ten sessions of this experiment, it is observed that, higher market overconfidence is accompanied by the higher average market prices and larger deviations of the security prices from fundamental value. Prices in rational markets tend to track the fundamental asset value more accurately than the prices in the overconfident markets, and are significantly lower than the average overconfident prices. Moreover, bubble and burst pattern was observed in the aggregated overconfident market, whereas in the rational market no sudden drop of the aggregated market price to the fundamental value occurred. Volatility of the prices and trade volume proved to be significantly lower in the rational market, as it was hypothesized. Overconfidence measure of the first part of the experiment is, in most markets, lower than that of the second part and this difference is significant. This finding could serve as an explanation why bubbles burst close to the end (or in some cases middle) of the experiment. Analysis of the bubble measures revealed that in the markets formed of overconfident subjects bubbles are more likely to occur and that they are significantly larger in magnitude than in rational markets. Large and significant correlation between bubble measures and measures of overconfidence provide additional evidence that overconfidence has significant effect on price and trading behavior in experimental asset markets.

Paper proceeds as follows. In Section 2.2 a brief overview of the findings of psychological<sup>9</sup> and financial literature on overconfidence are given; along analysis of the similar work and

---

<sup>9</sup> A detailed discussion of the relevant literature is provided in the working paper "Development of the overconfidence measurement instrument for the economic experiment".

discussion of the paper's contributions is presented. In Section 2.3 the research hypotheses are listed. In Section 2.4 details of the pre-experimental overconfidence measurement are provided. Section 2.5 provides description of experimental design. In Section 2.6 data analysis is presented, and, finally Section 2.7 concludes.

## 2.2 OVERCONFIDENCE

### 2.2.1 OVERCONFIDENCE IN PSYCHOLOGICAL RESEARCH

The beginning of the overconfidence research in finance and economics lies in psychological works. In psychological research overconfidence is defined as a prevalent tendency to overestimate one's skills, prospects for success, the probability of positive outcomes or the accuracy of one's knowledge. Phenomenon of overconfidence has been found in many different samples of the population, e.g. students (Fischhoff et al., 1977; Koriat et al., 1980, Zakay and Glicksohn, 1992), members of the armed forces (Hazard and Peterson, 1973), CIA analysts (Cambridge and Shreckengost, 1978), entrepreneurs (Baron, 2000), clinical psychologists (Oskamp, 1962), bankers (Staël von Holstein, 1972), executives (Moore, 1977), negotiators (Neale and Bazerman, 1990), managers (Russo and Schoemaker, 1992), lawyers (Wagenaar and Keren, 1986), and civil engineers (Hynes and Vanmarcke, 1976); overconfidence is already present in children (see Powel and Bolich, 1993; Allwood, Granhag, and Jonsson, 2006).

*Confidence and uncertainty* In our life, many decisions are based on beliefs concerning the likelihood of uncertain events (Tversky and Kahneman, 1982). These beliefs can be expressed in numerical form as subjective probabilities. Subjective probabilities are the probabilities that people generate in their own minds to express their uncertainty about the possibility of the occurrence of various events or outcomes (Bar-Hillel, 2001). If over the long run, for all predictions made with some specific confidence, the actual proportion of correct outcomes equals the probability assigned, a person is considered to be well calibrated. Overconfidence, or miscalibration, concerns the fact that people overestimate how much they actually know: when they are  $P$ -percent sure that they have answered the question correctly or predicted (the outcome) correctly, they are in fact right on average less than  $P$ -percent of the time (Bar-Hillel, 2001). Optimistic overconfidence is a specific form of overprediction, based on overestimation of the probability of events thought to be beneficial to the judge (Griffin and Brenner, 2005). Most of the people are not well-calibrated and demonstrate overconfidence. Overconfidence can also be defined with respect to subjective confidence intervals (Kirchler and Maciejovsky, 2002). The assessor has to state values of the uncertain quantity that are

associated with a small number of predetermined fractiles of the distribution. The usual finding is that the subjects' probability distributions are too tight. In the study of Alpert and Raiffa (1982) fifty-percent intervals included the true quantity only about 30 percent of the time; 98 percent intervals, only 60 percent of the time.

The degree of overconfidence is connected to the complexity of the task, and is the highest with the tasks of high difficulty (e.g. Clarke, 1960; and Pitz, 1974). As tasks get easier, overconfidence is reduced (Lichtenstein et al., 1982). Russo and Schoemaker (1992) note that being well calibrated is a teachable, learnable skill, which is demonstrated by the example of weather forecasters, who significantly improved accuracy of their forecast predictions and became one of the best ever calibrated group of subjects. Lichtenstein et al., (1982) conclude that continuance, repetitiveness of the task and the fact that, the outcome feedback for weather forecasters is well defined and promptly received, have high impact on accuracy of their predictions. There are two ways to achieve better subjects' calibration, which according to Lichtenstein et al. (1982) are motivation through reward for their assessment to be more precise, and outcome feedback<sup>10</sup>.

### 2.2.2 OVERCONFIDENCE IN FINANCIAL RESEARCH AND CONTRIBUTIONS

Following the psychological research in overconfidence, interest in the consequences of economic subjects' overconfidence on financial decision making, functioning of markets and economic outcomes has occurred in behavioral economics. Theoretical models of overconfidence predict that overconfidence causes excess trading volume (De Bondt and Thaler, 1985; Shiller, 2000; Benos, 1998; Caballé and Sákovics, 2003), and excess price volatility (Scheinkman and Xiong, 2003; Benos, 1998, Daniel et al., 1998); it induces occurrence of the speculative price bubbles (Scheinkman and Xiong, 2003) and increases market depth (Odean, 1999; Kyle and Wang, 1997; Benos, 1998); it makes markets underreact to abstract, statistical, and highly relevant information and overreact to salient, but less relevant information (Odean, 1998); it makes returns of financial assets predictable (Daniel et al., 1998, 2001; Scheinkman and Xiong, 2003); overconfidence increases investors' tendency to herd (Hirshleifer, Subrahmanyam and Titman, 1994) and makes them choose riskier and undiversified portfolios (Odean, 1998, 1999; Lakonishok, Shleifer and Vishny, 1992), overconfident investors trade more aggressively, i.e. their trading activity is too high (Odean, 1999; Gervais and Odean, 2001) and their expected utility is reduced (De Long et al.,

---

<sup>10</sup> Moreover, receiving outcome feedback after every assessment is the best condition for successful training (Lichtenstein et al., 1982).

1991; Odean, 1998). Most of these papers are based on the initial assumption of traders' overconfidence, which is modelled as overestimation of the precision of private information that manifests itself via underestimation of the variance of the private signal that subjects get, or, in other words, too tight confidence intervals for the value of the risky asset (Glaser and Weber, 2007).

There are a few empirical and experimental studies designed to test the impact of overconfidence on financial decisions, market outcomes and subjects' performance. Some of them present only an indirect evidence of such impact, as they measure overconfidence via different proxies and it is not always clear who of the subjects and how strong are overconfident. For example Statman, Thorley, and Vorkink (2006) test the hypothesis of interdependence between overconfidence and high trading volume for the USA stock market. As a proxy for the degree of overconfidence authors suggest using the high past returns, i.e. they argue that after high past returns posterior volume of trade will be higher, as successful investment increases the degree of overconfidence. These conclusions are supported by Kim and Nofsinger (2003) for the Japanese stock market. Barber and Odean (2001) proxy overconfidence by the gender of the trader, i.e. their proposition is that, based on the psychological literature, women are less overconfident than men, thus they are going to trade less than men. In their study men were actually found to trade more than women.

A much clearer results are obtained through test-studies, enabling direct observation whether an examined person overestimate their knowledge, or underestimate variance of stock returns etc. For example, Menkhoff, Schmidt and Brozynski (2006) surveyed 117 fund managers in order to detect an impact of experience on overconfidence, risk taking, and herding behavior. However, only experiments enable a direct test of the hypothesis that a certain degree of overconfidence leads to a specific market outcome, expressed as some of the market parameters, e.g. average price, or trade volume. Market experiments which are the closest in spirit to mine were conducted by Biais, Hilton, Mazurier, and Pouget (2005), Kirchler and Maciejovsky (2002), Deaves et al. (2009). All these experiments analyzed relation between measures of overconfidence and trading behaviour.

Kirchler and Maciejovsky (2002) run a multi-period experimental market and analyze development of overconfidence of the participants in the course of the experiment. Miscalibration of subjects was measured before each trading period, via the two price prediction tasks: point prediction and interval prediction. Their results indicate that participants of the experiment were well-calibrated in certain periods, and under- or



overconfident in other periods. They also find that higher degree of overconfidence is negatively correlated with the earnings of the participants of the experiment.

Biais et al (2005), use psychological questionnaire to measure, among other psychological traits, the degree of overconfidence via interval estimation tasks in a group of 245 students. Several weeks later after the students' overconfidence was measured they participated in an experimental asset market. The main conclusions of the authors are, that although miscalibration does not lead to an increase in trading activity it reduces trading performance of the subjects, and miscalibrated traders show "excessive confidence in their assessment of the value of asset", which eventually causes mistakes in financial decision making. Miscalibration reduces profits for men, whereas has no significant effect on women.

Deaves et al. (2009), conduct their experiment in order to test premises that overconfidence leads to an increase in trading activity, and that gender influences trading activity through differences in overconfidence. Compared to the two abovementioned experiments Deaves et al. (2009), instead of a multi-period experiment, conduct a battery of 12 single-period markets per experimental session and they use an increased up to 20 questions test consisting of the interval estimation tasks. To some of their sessions subjects were assigned based on their gender, and to some based on the overconfidence measure (OC). The values of OC measure used in the experiment of Deaves et al. (2009) show that all their subjects were extremely overconfident<sup>11</sup>. The main finding reported in their paper is that greater overconfidence leads to higher trading volume and leads to reduced earnings, but there is no evidence that overconfidence and trading activity are gendered.

My experiment was constructed with the following assumptions in mind:

First of all, most of the previous experiments concentrate on the connection between overconfidence and high market trade volume, and none of them was aimed at discovering the connection between the phenomenon of overconfidence and the occurrence of the bubbles in asset prices.

Second, there were no papers that previously used suggested procedure of markets formation, based on the participants' inborn level of overconfidence, and have managed directly connect changes in traders' psychological characteristics to the experimental market outcomes. Although Deaves et al. (2009), as mentioned above, run several sessions to which subjects

---

<sup>11</sup> In the experiment of Deaves et al. (Deaves et al. (2004) OC measure is constructed so as to vary in the interval [0, 1], where 1 points at extreme overconfidence. A well-calibrated person's OC score is 0.1, and values below point at underconfidence. However none of their subjects comes close to 0.1, the lowest OC being equal to 0.45.

were assigned by the degree of overconfidence, the issue of association of overconfidence with price-bubble was not in their focus, and therefore not explored. Not to mention, that they utilized a different overconfidence measurement methodology, and opted for different market structure (a battery of one-period markets per session vs. one multi-period market).

Third, previous experiments provided participants by private information with differences in signal quality, which according to Glaser et al. (2007) already creates a potential for trade<sup>12</sup>. E.g. in the experiment of Kirchler and Maciejovsky (2002) half of the participants had no information about the dividend distribution, and the other half had complete information. Experimental approach of Biais et al. (2005) relies on the asymmetric information trading game, where traders observe different private signals: bullish, bearish, and neutral. Deaves et al. (2009), also supply their subjects with different, in terms of quality, signals that depend on the results of the pre-experimental test. Moreover they try to manipulate the subjects' beliefs so that they think that their signals are more accurate. I do not create artificial belief in being better or possessing a piece of a more qualitative information. Instead all subjects are given the same information and I believe that only such approach enables the refinement of the pure differences between the two experimental groups.

Fourth, economic experiments on overconfidence measure the inborn level of subjects' overconfidence via the different tasks and tests, and in previous experiments overconfidence might have been caused (to some extent) by other reasons than the imperfection of human nature, namely by mistakes in the development of tests/tasks. E.g. findings from the psychological research show that overconfidence is the most pronounced for the hard questions (few people know the right answer) and the least for the easy (most of the people give a correct answer) questions. However, none of the abovementioned papers makes use of the balanced to hard-easy effect tests. This could have artificially created high levels of under- or overconfidence. For example in the experiment of Deaves et al. (2009) none of the subjects gets even close to the perfect calibration measure, and even the best calibrated participants exhibit rather high degree of overconfidence<sup>13</sup>. I created the specially tailored test, weighted for the inclusion of easy, hard and medium difficulty questions (also accounting for the possible gender bias) that was pre-tested and used with students enrolled in different disciplines of the social sciences. Compared to some of the authors, my test is expanded to

---

<sup>12</sup> If investors receive different pieces of private information about the uncertain value of the risky asset, there is heterogeneity between investors and thus a potential for trade (Glaser et al, 2003).

<sup>13</sup> This also raises doubts in the validity of their division of subjects in low and high overconfidence markets.

include more questions. Both overconfidence test and price prediction assignment are financially rewarded, which increases reliability of the overconfidence measurements.

And last but not least, I use two constructs to measure subjects' overconfidence: a general knowledge based, and based on the stock-price prediction task. Biais et al. (2003) and Deaves et al. (2009) use only general-knowledge tasks, where overconfidence is being estimated via the interval estimation tasks. In the experiment of Kirchler and Maciejovsky (2002) a pre-experimental overconfidence measurement did not occur, but rather, overconfidence was measured in the course of the experiment via the price prediction task. My design makes possible not only the evaluation of the students' pre-experimental degree of overconfidence, and based on that, division of students into two different types of market, but also the construction of the measure of the change in the markets' overconfidence from the first half of the experiment to the second. This enables more confident inference about the connection between the development of overconfidence and the bubble burst.

## 2.3 HYPOTHESES

Investment decisions in the experimental market are based on beliefs concerning the likelihood of the two kinds of independent uncertain events: 1) size of dividend at the end of the period and 2) probability to resell to the party willing to pay even more. I assume that subjective probabilities generated by overconfident traders make them overestimate the probability of the occurrence of the maximum dividend value, thus traders erroneously perceive possible future dividend income and optimistically overestimate the probability of existence of the irrational traders ("greater fools") ready to pay for the asset an even higher price. This results in that the participants are taking excessive risk and trade at prices above the fundamental asset value, and are even higher than the maximum possible dividend value. Both these reasons create a fertile field for the occurrence of the bubble in the experimental asset's price. Following this discussion the first hypotheses is formulated:

*H<sub>1</sub>. Trade in the two types of constructed markets will follow such patterns:*

### 1. Rational market:

- No trade or trade around the fundamental value (average expected dividends)
- Investors trade relatively infrequently (low trading volume)
- Prices are not too volatile relative to fundamentals
- No bubble-crash pattern observed

## 2. Overconfident (irrational market):

- Trade at prices around maximum possible dividend value and trade at irrationally high prices i.e. exceeding the maximum possible dividend value.
- Excessive trade volume.
- Observed bubble and burst pattern

The second hypothesis is based on the work of Smith, Suchanek, and Williams (1988) and findings from psychological literature. Experiments by SSW (1988) showed that bubble/burst pattern is persisting scenario in the markets with inexperienced agents. Usually bubbles burst several periods before the end of the trading game. Research on overconfidence showed that overconfidence is decreasing in experts or with the task repetitiveness (see Sieber, 1974; Pitz, 1974; Lichtenstein et. al., 1980; Russo and Schoemaker, 1992). Also optimism diminishes with experience (Fraser and Green, 2006). Thus a second hypothesis is postulated:

*H<sub>2</sub>: Reduction in overconfidence causes bubbles' crash. Overconfidence is reduced with experience.*

As mentioned earlier subjects can be trained to be better calibrated by motivating them financially to be more precise in their predictions, and by giving them feedback on their predictions' results. These both conditions are fulfilled in the experiment. Thus in the course of the trading game participants gain experience in it, and supported by market information about the results of their repeated actions turn into being "experts" of the game. Expertise should improve calibration of the subjects and bring about changes in their trade patterns (e.g. decrease in trading volume and price), causing stock price bubble's crash. Thus bubble bursts as participants turn being better calibrated, and correct their subjective probabilities downwards.

## 2.4 PRE-EXPERIMENTAL OVERCONFIDENCE MEASUREMENT

Pre-experimental psychological test sessions were conducted during several lectures on economics at the Christian-Albrechts University of Kiel. In each of the chosen classes, students were announced that they had an opportunity to take part in the short experiment on the voluntarily basis, for which a general knowledge quiz had to be filled out. For this activity 15 minutes were given. Participants of each pre-experimental session competed for the three prizes of 30, 20 and 10 EUR, which were awarded to those who answered the most questions right. Before the students started with the tests, a planned market experiment was advertised, and those subjects who were eager to take part in the economic experiment were encouraged

to mark their interest on the tests by ticking the “I’m interested in participation in further experiments” option and leaving their contact in the form of an e-mail address.

The pre-experimental quiz consisted of the 18 general knowledge questions unrelated to economics, financial markets or experiments<sup>14</sup>. Every question had three answer alternatives, only one of which was right. After answering each question participants had to state how confident they were that the answer was right. For this purpose they could use any number in the range from 33%, meaning complete uncertainty, to 100% - complete certainty.

The overconfidence (underconfidence) of each participant was measured as the bias score. The bias score of an individual was calculated as the difference between the mean confidence level across all questions and the proportion of correct answers:

$$\text{bias score} = \text{average \% confidence} - \text{average \% correct} \quad (2.1)$$

A positive bias score represented overconfidence, and a negative bias score represented underconfidence. A bias score of zero indicated accurately calibrated person (neutral person).

This pre-experimental procedure allowed the author to obtain a large pool of students with their estimated bias scores in her database, and to ensure that the two stages of the experiment were perceived by the students as two rather non-associated experiments. Based on the pre-experimental calibration test individuals were divided into two groups – the least and the most overconfident, which are further on called correspondingly rational and overconfident subjects. Students were addressed through the e-mails according to their overconfidence and invited to register for the suggested experimental sessions. All students of a specific type of the calibration were approached at the same time and were given several possible experiment days for their choice, thus subgroups participating in different experimental sessions differed in their average overconfidence within the two main groups (rational and overconfident).

More than two hundred students showed interest in the forthcoming economic experiment. A database of the interested persons included information on 222 students’ name, age, nationality, direction of studies, semester and overconfidence score. Potential experimental subjects were undergraduate and graduate students of economics, business administration and other social science disciplines, aged from 19 to 43 years ( $M = 22.95$ ,  $SD = 2.73$ ). Of those

---

<sup>14</sup> Questions were not connected to economics, as otherwise it could cause biased results if the same questionnaire was used with the heterogeneous pool of subjects the experimenter had in her disposition. Deaves et al., (2008) also motivate their choice of non-economic questions by the attempt “to avoid giving either group of participants a relative advantage because of subject content”.

only nine percent were of non-German nationality (19 non-German, and 203 German). Consistent with the previous research, subjects in the database on average were prone to overconfidence ( $M = 11.78$ ,  $SD = 10.58$ ).

Appendix 2.A presents data on the bias scores of the various (pre-)experimental subgroups: all participants who were in the database, all students who participated in the experimental sessions (a subsample of those in the database), and their subsamples – men, women, participants of rational, and overconfident markets. All groups seemed to be extremely overconfident, except for the participants of the rational market. A hypothesis of the equality of the average overconfidence of different subgroups was tested against the alternative that different subgroups varied by their overconfidence levels. The mean equality hypothesis is failed to be rejected for the difference between overconfidence of male versus female subjects both in the whole sample of pre-experimental test participants, as well as among all experiment participants and overconfident/ rational participants. The bias score of the participants of the overconfident markets is significantly higher than of those of the rational markets.

## 2.5 EXPERIMENTAL DESIGN

### 2.5.1 PARTICIPANTS

A set of ten experimental sessions was conducted at the Christian-Albrechts University of Kiel between November and May 2008-2009. For each session six participants were recruited from the undergraduate and graduate students in economics, business administration and other social science disciplines who had not previously participated in a similar asset market experiment<sup>15</sup>. Seventy four people took part in experimental sessions, of them 60 people actually traded in the experimental markets. Thirty five males and 25 females, aged 19 to 28 ( $M = 22.73$ ,  $SD = 2.06$ ) participated in the experimental sessions. 87% of the participants were of German nationality. Thirty five subjects studied economics, 18 – business management, and 7 were students of the other social science disciplines. Approximate time required to conduct the experiment was 1 hour and 40 minutes. Subjects earned on average 390.36 ECU (10.54 EUR) ( $SD = 197.89$ ) on the asset market (without the reward for the forecasting activity). Men earned on average more ECUs than women: women 335 ECU and

---

<sup>15</sup> Inexperienced subjects were chosen, because Smith, Suchanek, and Williams (1988) found that, when participants had little or no previous experience in asset markets the markets exhibited price bubbles and crashes rather than tracked the fundamental value.

men 447 ECU. This difference is significant (Mann-Whitney  $Z = -2.646$ ,  $p < 0.01$ , one-sided). Instructions familiarized participants with the rules of the experimental market. English translation of instructions is included in Appendix 2.B.

## 2.5.2 EXPERIMENTAL PROCEDURE AND THE RULES OF THE GAME

All experimental sessions were conducted in the computer lab. Six players participated in each of the experimental asset markets. Subjects could take part in only one experimental session and only in that type of the market (rational/overconfident) to which they were appointed based on the results of the psychological test. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007).

At the beginning of the typical session students were given time to read the detailed instructions and ask the questions. At the end of the time devoted for reading the instructions experimenter again repeated the most important information at which students should pay attention, to ensure that everyone understood the rules of the game. Two trial periods followed, during which students could familiarize themselves with the experimental software, and were allowed to ask questions if something was unclear to them. Both prior to the trial periods and after them subjects were informed that these periods had no impact on their results.

Experimental design followed the pioneering work of Smith, Suchanek, and Williams (1988) with slight changes in the price forecasting task, and was performed as a continuous anonymous double auction. Every experimental market consisted of the sequence of 15 trading periods lasting at most 180 seconds during which each trader could post her bid and ask price of the asset unit. Therefore each participant could purchase asset units for their inventory by spending an amount of their working capital equal to the purchase price, or sell the inventory units and increase their working capital by an amount equal to the unit's sale price. Prior to the start of the experiment each trader was endowed with an equal amount of experimental assets and cash: 300 units of experimental currency (ECU) and 3 units of the experimental asset. At the end of the trading period, each asset in the inventory of the participants paid a dividend with possible values of 0.0, 0.8, 2.8, or 6.0 ECU. Probability of each dividend value was 0.25. Thus on average, through many draws subjects could expect a 2.4 ECU value dividend. Fundamental value of the stock is found according to the formula  $n \times 2.4$  ECUs, where  $n$  stands for the number of periods remaining to the end of the session. Thus in round 1, the expected fundamental value from the dividend stream was  $15 \times 2.4 = 36$  ECUs per each share; in every subsequent period it fell by 2.4 ECUs.

As the trading period was over, participants were shown market summary information from the past trading periods, and were asked to predict the average market price for the next period as well as to state how confident they were that their price forecast was correct. To express their confidence subjects could use any value between 0% and 100%, where 0% meant complete disbelief that the forecast could be true, and 100% meant complete belief that the forecast was correct. Participants were paid for their predictions based on their accuracy. Each period subjects were given feedback on their accuracy and their reward for the price forecasting task. Point estimation form of price prediction task, e.g. used by SSW (1988), was chosen over price interval estimation form due to several reasons. First, overconfidence measure obtained through interval estimation in the article by Kirchler and Maciejovsky (2002) did not vary in time and remained in the area of overconfidence; however, their point-estimate measure varied in time and took values from overconfident, to well-calibrated, and underconfident. Second, this form of price prediction task enabled comparison between pre-experimental and post experimental overconfidence measures.

### 2.5.3 INCENTIVES

A typical experiment lasted 1 hour and 40 minutes, and at the end of it subjects were paid in cash the amount of money that was based on their final working capital converted at the predefined exchange rate to Euros. Final working capital (FWC) equalled:

$$FWC = (300 \text{ ECU starting capital}) + (\text{dividend earnings}) + (\text{stock sales revenue}) - (\text{stock purchase cost}) \quad (2.2)$$

In order to motivate students they were offered an hourly reward, which was comparable to what on average an hour of the “student-job” in Germany pays<sup>16</sup>, thus the asset market offered participants on average 7 EUR per hour of the experiment; for the whole experiment participants could expect to get on average 11 EUR.

Reward for the accuracy of predictions was constructed to be an additional income source in order to reduce mechanical participation and encourage conscious engagement into this activity. The closer the prediction was to the actual average market price, the higher was the reward. The reward scheme used in the experiment was similar to the suggested by Haruvy, Lahav, and Noussair (2007)<sup>17</sup>:

---

<sup>16</sup> To author’s knowledge in the job market students could get on average 7 to 8 EUR.

<sup>17</sup> This incentive scheme instead of a quadratic scoring rule was chosen for the sake of keeping the instructions simple (Haruvy et al., 2007).



Level of Accuracy	Reward
Within 10% of actual price	3 ECU
Within 25% of actual price	1 ECU
Within 50% of actual price	0.5 ECU

Both monetary reward and the feedback about their predictions' accuracy were used for improving the subjects' calibration in the price prediction task.

## 2.6 EXPERIMENTAL RESULTS

### 2.6.1 NUMERICAL CHARACTERISTICS OF THE TWO TYPES OF THE MARKET

In this section various numerical characteristics of the two types of the market are compared. Each session counts as one observation. Totally ten market sessions were conducted: five sessions for the overconfident market, and five sessions for the rational market. If not stated otherwise, all data for each type of the market are ranked from the lowest to the highest.

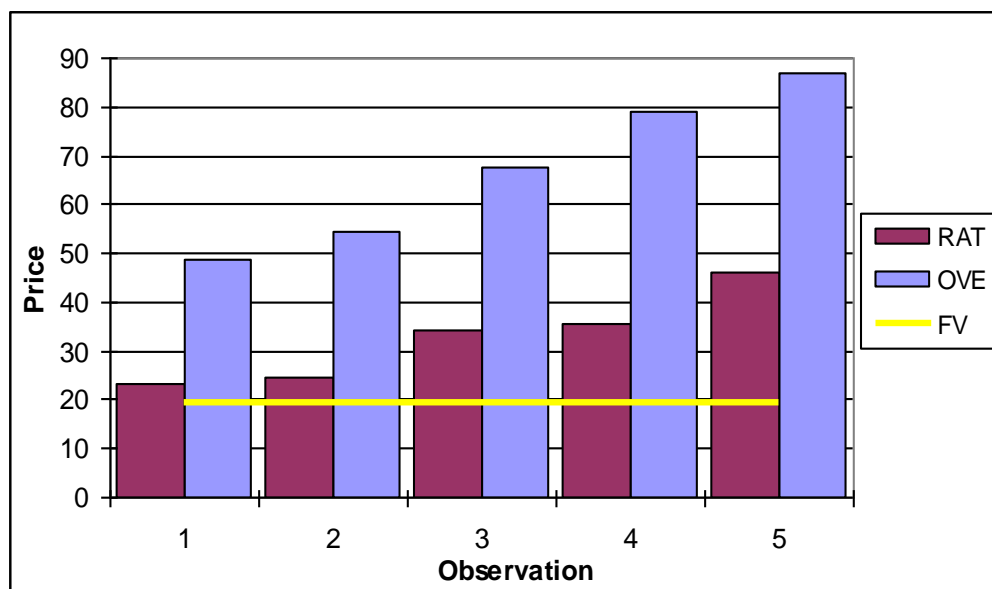


Figure 2.1: Average trade prices in both types of markets

I start by the comparison of the average contract prices in the rational and overconfident markets. Figure 2.1 demonstrates that on average prices in the overconfident market tend to be higher than in the rational market. The average market price for the rational markets was

33 ECUs (SD = 9.41<sup>18</sup>) and for the overconfident market 67 ECUs (SD = 16.02<sup>18</sup>). Statistical test of the difference between the average prices supports the initial conclusion from the visual analysis - average prices in the overconfident market are significantly higher than the rational market prices (Mann-Whitney U = 0.0, p < 0.01, one-sided). Now I turn to the comparison of the average prices obtained in the experiment to the average intrinsic value of the asset (fundamental value) 19.20 ECU. Figure 2.1 indicates that experimental average prices exceed the average fundamental value (from now on FV). Wilcoxon Signed Rank test supports that prices both in the rational and in the overconfident markets are higher than the fundamental value (Wilcoxon T = 1.89, p < 0.05, one-sided), i.e. prices in both types of the experimental market are shifted to the right from the fundamental value.

### *Evolution of the Average Price*

Figure 2.2 presents the development of the joint average transaction prices for all five rational and all five overconfident markets. On the horizontal axis trading periods are indicated; vertical axis measures average price of the transaction for that period. Fundamental asset value, which is found as the sum of the expected dividends for the periods left till the end of the game, is presented alongside.

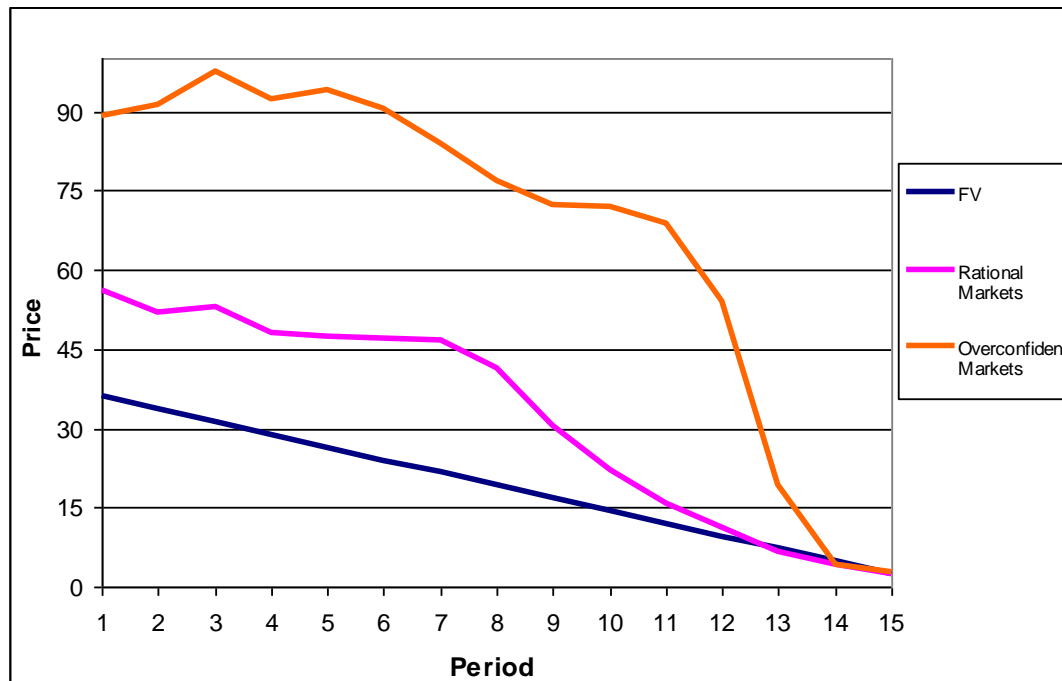


Figure 2.2: Development of the average market price

<sup>18</sup> Here aggregated average price and the standard deviation, which are based on the five average prices for each type of the market, are presented.

Visual data analysis suggests that prices deviate from the fundamental values in both types of the market. However prices in the rational market deviate from the fundamental value to a smaller extent than in the overconfident market. Although prices in both types of the market stay away from the fundamental value for almost the whole duration of the session, prices in the rational market tend to track the fundamental asset value more accurately than prices in the overconfident market. It can also be seen that in the aggregated overconfident market the bubble and burst pattern is more pronounced than in the aggregated rational market, where no sudden drop of the aggregated market price to the fundamental value is observed.

### *Volatility*

Prior to the experiment I hypothesized that prices in the rational market would be less volatile than in the overconfident market. Figure 2.3 presents price variations in both types of the market, measured as standard deviations<sup>19</sup>. Initial analysis suggests that this intuition was right. The conducted Mann Whitney U test confirms that variation in prices in the overconfident markets is significantly higher than in the rational markets (Mann-Whitney U = 4,  $p < 0.05$ , one-sided). For both types of the market, Wilcoxon Signed Rank test enabled rejection of the null hypothesis that the volatility of the prices was not less than the volatility of the fundamental value at the significance level of 0.05 (Wilcoxon T = 1.89,  $p < 0.05$ , one-sided) in favor of the alternative hypothesis that the median volatility of rational/overconfident market exceeded the volatility of the fundamental value (SD = 10.73).

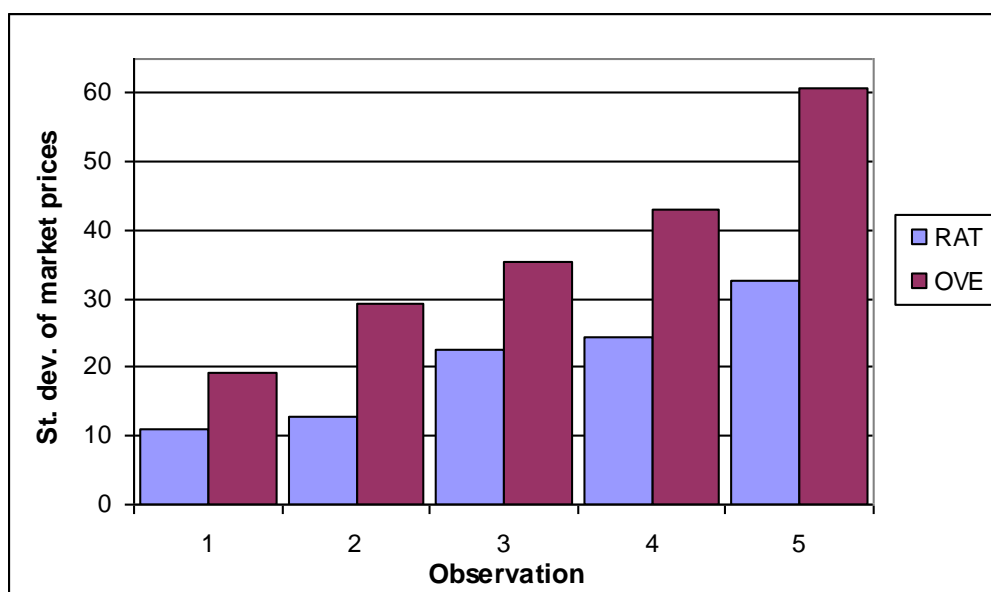


Figure 2.3: Volatility of asset prices in both types of markets

<sup>19</sup> Based on all prices of that market.

### Trading Activity

In this subsection measure of market volume is introduced – average asset turnover rate (market turnover). Market turnover is obtained by dividing the number of the asset units traded in that market by the total number of the asset units in that market (18 units in our case).

I start by testing if the propositions of the No-Trade Theorem by Milgrom and Stokey (1982) applied in the conducted experimental markets. This theorem states that rational agents who differ from each other only in terms of information and who have no reason to trade in the absence of information will not trade (Milgrom and Stokey, 1982). Even though one type of the market was constructed so as to be on average rational and there was no private information Figure 2.4 suggests that trading activity in neither market is zero. Wilcoxon Signed Rank test of the hypothesis that there was no trading activity in the overconfident/rational market is rejected in favor of the alternative hypothesis that the trading activity is significantly higher than zero (Wilcoxon T = 1.896,  $p < 0.05$ , one-sided).

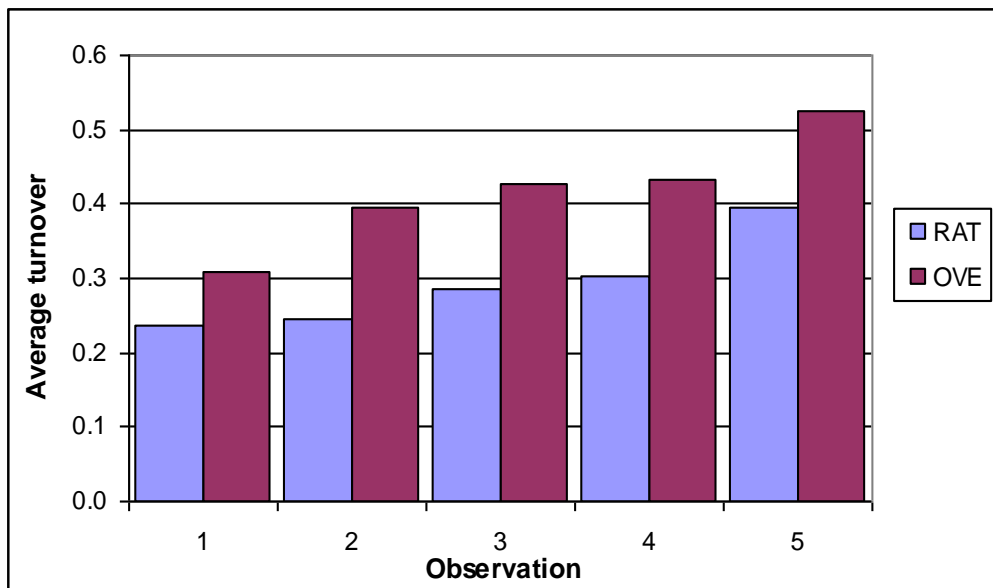


Figure 2.4: Average trading activity (turnover) in both types of markets

Trading activity in the rational market is lower than in the overconfident one: average market turnover in rational market sessions is 28% (5 units of asset) and 44% (8 units of asset) in overconfident. Mann-Whitney U Test was conducted to test if the average asset turnover rate in rational markets was the same as in overconfident markets, or whether alternatively market turnover in overconfident markets was higher. Trading in overconfident markets is found to be significantly higher than in rational markets (Mann-Whitney U = 1.5,  $p < 0.05$ , one-sided).

Evolution of the joint average market turnover for five experimental sessions of rational market and five overconfident markets is shown in Appendix 2.C. It can be observed that the joint average market turnover decreased over the trade periods in both types of markets. Increase in trading activity in the last period can be attributed to the so-called end-game effect<sup>20</sup>.

## 2.6.2 OVERCONFIDENCE MEASURE FROM THE FORECASTING TASK

Bias score (BS) from the price forecasting task was calculated for each session separately, as an average from all participants' forecasts about the next period's average price and their confidence in the answer. The score was calculated based on the "binary" methodology: if the average price was equal to the forecast it got a weight of 1, if not – 0. Overconfidence measure from the pre-experimental test is strongly correlated with the overconfidence measure from the forecasting task (Spearman's rho (8) = 0.65,  $p < 0.05$ , one-sided). According to Cohen, (1988) this correlation coefficient is considered to be large, thus I assume that both constructs measure the same phenomenon. This result also suggests that overconfidence is a robust phenomenon in our sample.

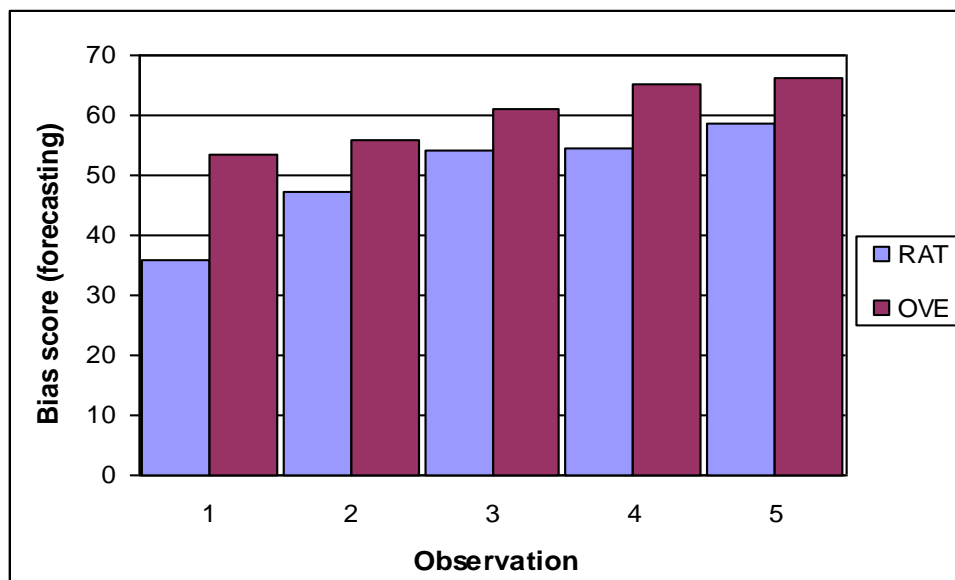


Figure 2.5: Average overconfidence in both types of markets

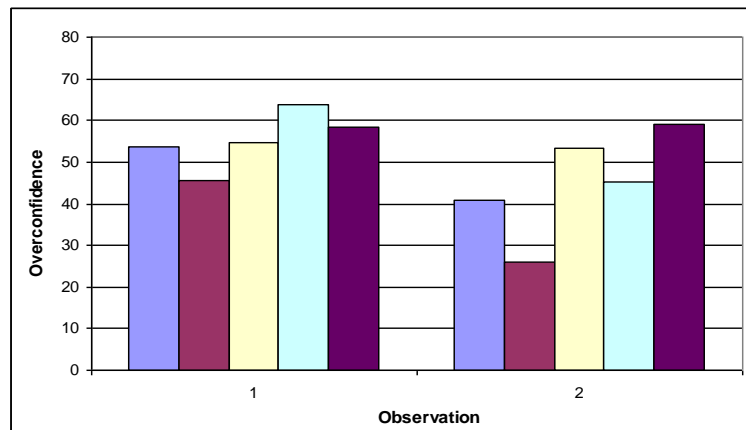
Figure 2.5 indicates that on average the bias score from the price forecasting task was higher in the overconfident markets than in the rational ones. On average overconfidence in price prediction task differed between the two types of market by 10 units (BS in rational markets

<sup>20</sup> The end-game effect occurs in repeated-round experiments, and is defined as the change in subjects' behavior that is attributed to the end of the experiment rather than being a part of subjects' behavior during the course of the experiment.

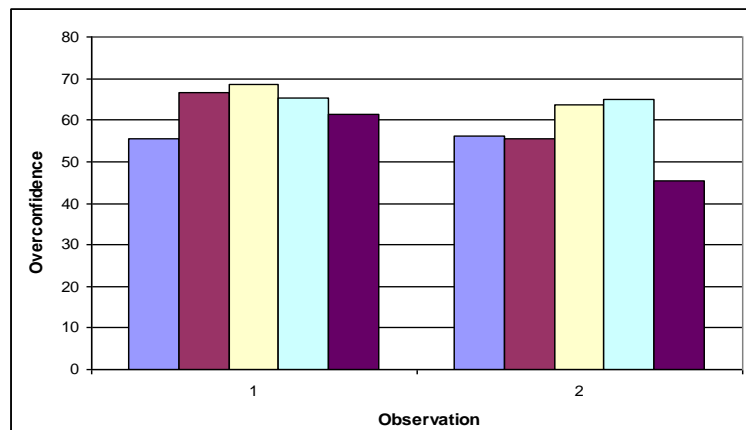
M = 50.08 (SD = 8.96); in overconfident markets M = 60.31 (SD = 5.02). BS value for the overconfident market is significantly higher than the BS for the rational market (Mann-Whitney U = 4.0,  $p < 0.05$ , one-sided).

*Evolution of the Bias Score*

To check if the proposition that overconfidence reduces to the end of the game compared to the beginning of the game holds true in the conducted experimental sessions, data on price prediction task were divided into two time intervals of seven periods in each, and two overconfidence measures for each market were calculated: one score for the first seven periods BS(2-8), and the second for the last seven periods BS(9-15). Figure 2.6 demonstrates that for most of the markets overconfidence measures calculated from the data on the price prediction for the first seven periods are higher than those calculated from the seven last periods. Wilcoxon Signed Ranks test confirms that BS(2-8) are significantly higher than BS(9-15) ( $Z = -2.429$ ,  $p < 0.01$ , one-sided). This finding could serve as an explanation why bubbles burst close to the end (or in some cases middle) of the experiment.



a.



b.

Figure 2.6: Comparisons of BS(2-8) and BS(9-15): a. rational market; b. overconfident market

### *Relationship between Overconfidence and Market Parameters*

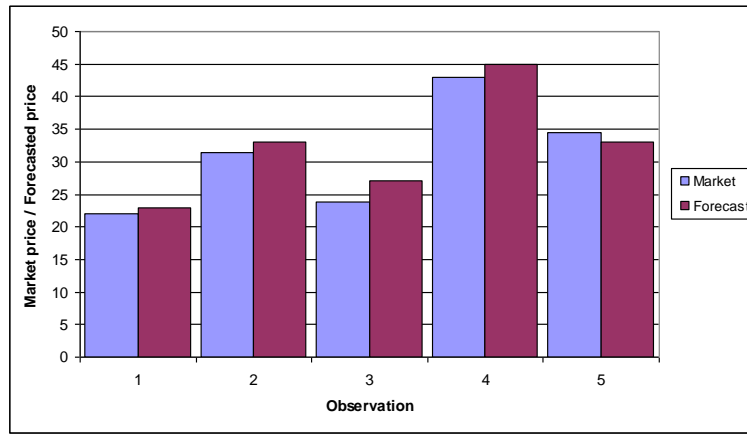
To check if the constructs that are used to measure overconfidence are related to the experimental results, analysis of correlation between the market parameters (specifically average market prices and the measure of trading activity) and the bias scores found from the pre-experimental overconfidence measurement and from the price prediction task is conducted. A positive significant linear relationship between the constructs and the average market price was found (pre-experimental BS: Spearman's rho (8) = 0.75,  $p < 0.01$ , one-tailed; forecasting BS: Spearman's rho (8) = 0.71,  $p < 0.05$ , one-tailed). These correlation coefficients are considered to be large. It can be concluded that an increase in overconfidence is associated with an increase in average prices.

Linear relationship between the bias scores found from the pre-experimental overconfidence measurement and found from price prediction, and average trade volume for the whole sample is found to be large (pre-experimental BS: Spearman's rho (8) = 0.69,  $p < 0.05$ , one-tailed; forecasting BS: Spearman's rho (8) = 0.58,  $p < 0.05$ , one-tailed). This is in line with the previous research, which found that the increase in overconfidence was paired with the increase in the trading activity, and the decrease in overconfidence was paired with the decrease in trading activity.

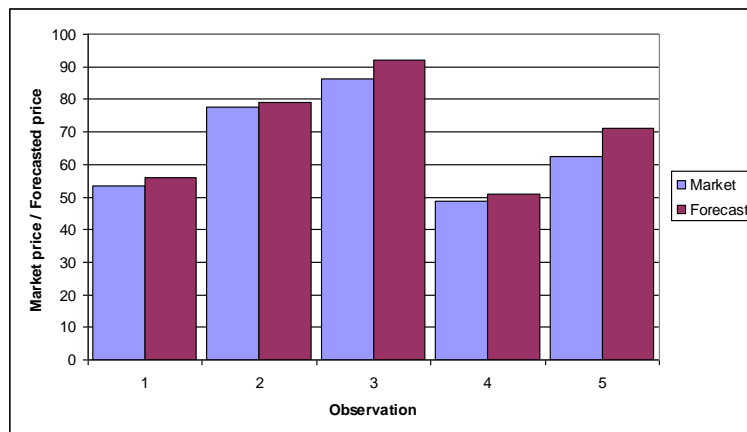
### 2.6.3 FORECASTING PRECISION

In this section precision of the forecasting activity for the two types of markets is analyzed. I start by graphical comparison of the average price predictions to the average market transaction prices to see whether any conclusions can be drawn about which type of the players (rational or overconfident) was better in forecasting. Figure 2.7 indicates that the differences between players' predictions and actual prices are small for both types of the markets, thus no clear conclusion can be drawn.

To conduct a statistical test of which group provided more accurate forecasts, differences between the average forecast and the average transaction prices are taken for the each type of the market. Then the hypothesis is tested, that the difference between the average forecast and the average market price equals to zero, versus the alternative one that the difference is not zero, or more than zero. After conducting the Wilcoxon Signed Rank test the null hypothesis cannot be rejected for the rational market ( $Z = 1.079$ ,  $p = 0.28$ , two-sided). In the case of the overconfident market it is concluded that the forecasts tend to over-predict the real market price significantly ( $Z = 1.89$ ,  $p < 0.05$ , one-sided).



a.



b.

Figure 2.7: Average forecast and market price: a. rational market; b. overconfident market

In a following step, a test is run if there is a linear relationship between the price prediction and the price. An almost one to one correspondence between the average forecast and the average market price is found (Spearman's  $Rho(8) = 0.997$ ,  $p < 0.001$ , one-tailed), which could mean that on average the convergence of beliefs occurred, alongside with the “anchoring” of the subjects on their predicted price values while trading in the market in order to make more money.

#### 2.6.4 BUBBLE MEASURES

From the previous analysis I obtained some evidence, that although prices, volatility and turnover in rational markets are significantly lower than in overconfident markets, they are still much higher than I have initially hypothesized, and that rational market might also be prone to bubbles, but of a smaller magnitude. Thus in this section experimental treatments will be analyzed in terms of their effect on the bubble's size. I use several measures of the magnitude of bubbles in laboratory markets that were developed by previous authors (e.g. King et al., 1993; Van Boening et al., 1993; Porter and Smith, 1995 as in Noussair and



Tucker, 2003; Dufwenberg et al., 2005). These measures are: Haessel-R<sup>2</sup>, Price Amplitude, Normalized Absolute Deviation, Normalized Average Deviation, and Velocity. Table 2.1 reports the values of the measures by session and treatment. Appendix 2.D presents graphs of average market prices and turnover values per period in each of the ten markets.

*The Hassel-R<sup>2</sup>* (Haessel, 1978) measures goodness-of-fit between average market price per period and the intrinsic asset value. It determines how well the variation in actual market prices (around their mean) is accounted for by the variation in the fundamental value (around its mean), or, in other words, the proportion of the variation in market price which can be explained by variation in fundamental value. Hassel-R<sup>2</sup> converges to 1 if trading prices converge to fundamental values<sup>21</sup>. It is estimated by the R<sup>2</sup> associated to the regression of market prices on the fundamentals, where fundamental value is seen as an estimator for the average market price obtained from some linear model. A comparison of average contract prices obtained from the rational market with those obtained from the overconfident one, reveals that variation in the average prices in the rational market fit variation in the intrinsic value better in most of the sessions. Thus goodness of fit measure is significantly higher in rational markets (Mann-Whitney U = 3, p < 0.05, one-sided).

*The Normalized (Average) Price Deviation* is calculated by summing up all deviations of market contract prices from fundamental value and dividing this sum by the total number of stocks in the market (see Equation 3).

$$NPD = \frac{\sum_{i=1}^q [P_{it} - FV_t]}{TSN} \quad (2.3)$$

Here,  $P_{it}$  is the price of the  $i$ th share in period  $t$ ,  $FV_t$  is the fundamental value in period  $t$ ,  $q$  is the number of contracts in period  $t$ , and  $TSN$  is the total number of shares in the market<sup>22</sup>. This measure is calculated for each period. Table 2.1 presents average value for each of the markets.

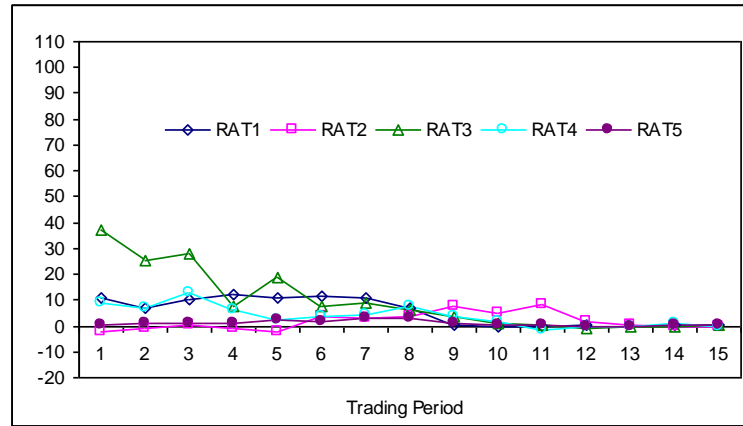
From the analysis of the NPD it can be determined whether stocks in that period were overpriced or underpriced relative to the fundamental value (a value of under- or overvaluation per-share). Average market value of the NDP can be treated as an indicator of the aggregated average under- or overvaluation per-stock in that market. Figure 2.8 depicts, for each of the two types of asset market, normalized price deviations from fundamental value per period. Results from rational markets are presented in the upper part of the panel, and

---

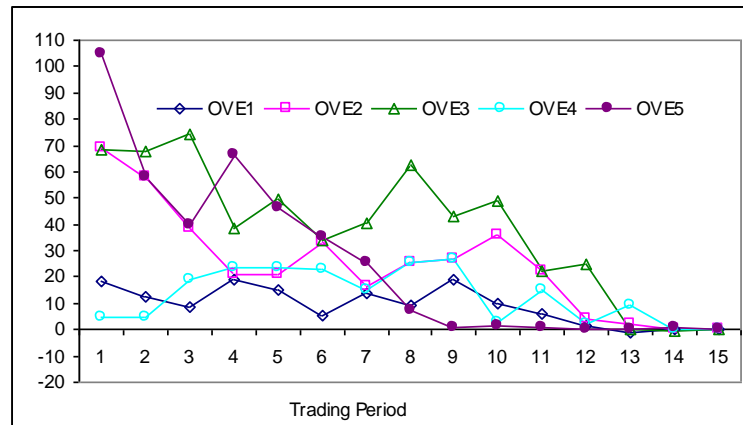
<sup>21</sup> See Dufwenberg et al. (2005) for an explanation why this measure is appropriate to experimental settings with known to the subjects last period.

<sup>22</sup> 18 in each of the ten experimental markets.

from overconfident markets - in the lower part. From Table 2.1 one can see that prices are on average much more overvalued in the overconfident market than in the rational market and this difference is significant (Mann Whitney U = 1, p < 0.01, one-sided).



a.



b.

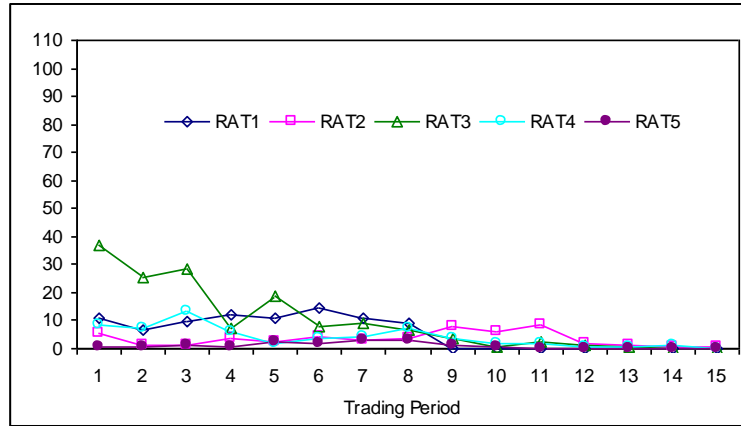
Figure 2.8: Normalized price deviations from FV by trading period and overconfidence level:  
a. rational market, b. overconfident market

The *Normalized Absolute Deviation* is similar to the NPD, and is found as the sum, over all transactions of that period, of the absolute deviations of the market prices from fundamental value, divided by the total number of stocks in the market (see Equation 4). This measure is calculated for each of the periods. Table 2.1 presents the average value of NAD for each of the markets.

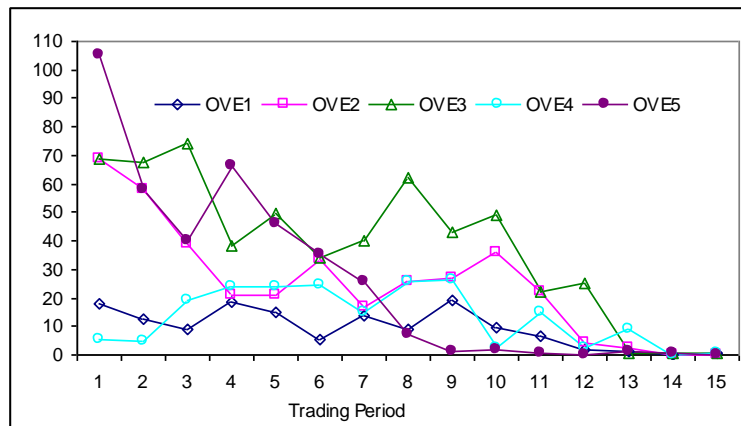
$$NAD = \frac{\sum_{i=1}^q |P_{it} - FV_t|}{TSN} \quad (2.4)$$

NAD measures the dispersion of the contract prices around the fundamental value, and high values of NAD point out that large number of transactions are being conducted at prices above the fundamental value. Figure 2.9 depicts absolute price deviations from fundamental

value per period for each of the two types of asset market. Results from rational markets are presented in the upper part of the panel, and from overconfident - in the lower part. From the Table 2.1 one can say that on average contract prices in the overconfident market differ by more experimental units from the fundamental value (22.36 ECU) than in the rational market (4.92 ECU), and this difference is significant (Mann Whitney U = 1,  $p < 0.01$ , one-sided). Since there are not many cases of per-share undervaluation relative to fundamental value, there are no considerable differences in the values of NAD and NPD.



a.



b.

Figure 2.9: Normalized absolute price deviations from FV by trading period: a. rational market, b. overconfident market

*The Price Amplitude (APL)* is the maximum value of the shift of average contract price from the fundamental value for an experimental session. It is found as the difference between the maximum positive and the maximum negative deviation of the average period price from the fundamental value of that period, normalized by the initial fundamental value (see Equation 5).

$$APL = \max \left\{ \frac{(P_t - FV_t)}{FV_1} \right\} - \min \left\{ \frac{(P_t - FV_t)}{FV_1} \right\} \quad (2.5)$$

Here,  $P_t$  is the average contract price and  $FV_t$  is the fundamental value in period  $t$ . Initial fundamental value  $FV_1 = 36$  ECU.

Higher price amplitudes imply greater bubbles, and larger swings in the market price of the asset relative to fundamental value, evidence that prices have grown away from their fundamental values. From the Table 2.1 one sees that the price amplitudes in the overconfident market are on average more than twice higher than the amplitudes in the rational markets, and this difference is highly statistically significant (Mann-Whitney U = 1.00,  $p < 0.01$ , one-sided).

*Velocity of the Asset* is found by dividing the total number of transactions over the experimental session by the total number of stocks in the market. This is the measure of how many times an asset “turned over” the market. This measure is connected to the volume of trade: the higher is the velocity, the higher is the volume of trade, suggesting, according to Noussair and Tucker (2006), either heterogeneous expectations or biases in decision making prompting trade. From the Table 2.1 one can see that the velocity of stocks in the overconfident market is significantly higher than in the rational market: on average each stock “turned over” 6.27 times in the overconfident market, and only 4.38 times in the rational market. This difference is significant (Mann Whitney U = 1.50,  $p < 0.05$ , one-sided).

Correlation coefficients between bubble measures and measures of overconfidence (pre-experimental and forecasting bias scores) are large and significant (see Appendix 2.E). This provides additional evidence that overconfidence has a significant effect on pricing and trade behavior in experimental asset markets.

Table 2.1: Bubble measures in each session

Session	Treatment	Hassel-R2	NPD	NAD	Amplitude	Velocity
1	OVE	0.581	9.144	9.308	1.69	4.61
2	OVE	0.535	24.908	24.939	2.25	5.94
3	OVE	0.414	38.257	38.380	2.87	7.89
4	OVE	0.288	13.008	13.196	1.32	6.50
5	OVE	0.877	25.874	25.961	3.33	6.39
6	RAT	0.906	5.745	6.133	1.09	4.56
7	RAT	0.571	1.769	3.412	0.67	5.94
8	RAT	0.944	9.593	9.924	1.67	4.28
9	RAT	0.805	3.781	4.099	1.15	3.56
10	RAT	0.942	0.983	1.017	0.30	3.67

### *Comparison to other experiments*

Table 2.2 presents data<sup>23</sup> from several experiments which had similar structure to mine: in which 1) the asset market had duration of 15 periods, and 2) the fundamental value was declining each period. In the experiments of Smith, Suchanek, and Williams (1988), Porter and Smith (1995), Van Boening, Williams, and Le Master (1993) bubble and crash pattern in prices is widely observed. On the contrary, experimental sessions of Noussair and Tucker (2006) yield practically no bubbles. On average values of Normalized Absolute Deviation<sup>24</sup> and the Amplitude from the rational treatment are higher than the values from the “no-bubbles” experiment of Noussair and Tucker (2006) but are much lower than those obtained from the other three experiments, thus there is evidence of the smaller deviations from the fundamental value in the rational market treatment. However the turnover value is more than four times higher than that of Nourssar and Tucker (2006). Measures obtained from the overconfident market treatment are consistent with those observed in previous studies of markets of this type.

Table 2.2: Average values of some of the bubble measures from previous studies

<i>Average values from my experiment</i>	NAD	Velocity	Amplitude
Overconfident market treatment	2.24	6.27	2.29
Rational market treatment	0.49	4.40	0.98
<i>Average values from previous research</i>			
Noussair and Tucker (2006)	0.24	0.99	0.33
Smith, Suchanek, and Williams (1988)	5.68	4.55	1.24
Porter and Smith (1995)	--	5.49	1.53
Van Boening, Williams, and Le Master (1993)	5.12	5.05	4.19

Results presented in this section demonstrate that although bubbles in the rational markets are not completely eliminated, they are less severe in comparison to the bubbles in overconfident

<sup>23</sup> Data on the experiments of Smith, Suchanek, and Williams (1988), Porter and Smith (1995), Van Boening, Williams, and Le Master (1993) are taken from the paper of Noussair and Tucker (2006).

<sup>24</sup> For the comparison of NAD measure from my experiment to those of the other experiments, it has to be divided by 10. The reason is that, previous studies used an expected dividend equal to 0.24 ECU is each period; in my experiment it is 2.40 ECU.

markets: bubble measures calculated for the rational sessions are statistically significantly smaller than the ones obtained from the overconfident sessions. Moreover size of the bubble measures increases with the increase in market overconfidence.

## 2.7 CONCLUSIONS

In this paper results of the experiment, designed to investigate the role of market overconfidence in the occurrence of stock-prices' bubbles, are reported. The design of the experiment follows Smith, Suchanek and Williams (1988) and is extended by a new feature, in which markets are constructed on the basis of subjects' overconfidence, measured in pre-experimental studies. In the experiment two types of the markets are conducted: rational and overconfident. Empirical evidence presented in this paper refines differences between market outcomes in the experimental treatments and suggests the existence of the connection between market overconfidence and market outcomes.

When there are no asymmetries in information and all traders have identical assets' and currency endowments, and all traders are "homogenous ... with statistically rational dividend and price expectations" (Gilette et al., 1999) a theory predicts that either no trading should occur or some marginal trading at the prices around the fundamental value. Findings from my experiment contradict this assumption. I find that trading activity in rational markets is significantly higher than zero; however it is significantly lower than in the overconfident ones. In the ten sessions of this experiment, it is observed that, higher market overconfidence is accompanied by the higher average market prices and larger deviations of the security prices from fundamental value. Although average prices in both types of markets significantly exceed the fundamental value, prices in rational markets tend to track the fundamental asset value more accurately than the prices in the overconfident markets, and are significantly lower than the average overconfident prices. Moreover, bubble and burst pattern was observed in the aggregated overconfident market, whereas in the rational market no sudden drop of the aggregated market price to the fundamental value occurred. Volatility of the prices and trade volume proved to be significantly lower in the rational market, as it was hypothesized.

Results show that both constructs that were used in the experiment to measure overconfidence (pre-experimental and price-forecasting task bias scores) are linearly strongly dependent, thus overconfidence is a robust phenomenon in my sample. Also it is taken as a proof that both constructs measure the same phenomenon. The strong positive correlation between market outcomes (average market price and trade volume) and overconfidence measures draws

conclusion, that an increase in market overconfidence is associated with the increase in average price and trading activity. The reduction of the aggregated average market price and trade volume over the experiment's periods is observed. Thus hypothesis that overconfidence also reduces to the end of the game was tested. For that, based on the data from the first and last seven periods, two bias scores for each market were constructed. Overconfidence measure of the first part of the experiment is, in most markets, lower than that of the second part and this difference is significant. This finding could serve as an explanation why bubbles burst close to the end (or in some cases middle) of the experiment. Menkhoff, Schmidt, and Brozynski (2006) find similar results of decrease in overconfidence with experience; however Kirchler and Maciejovsky (2002) report that overconfidence increases with the experience.

Analysis of the five bubble measures (NPD, NAD, Amplitude, Hassel-R2, and Velocity) revealed that in the markets formed of overconfident subjects bubbles are more likely to occur and that they are significantly larger in magnitude than in rational markets. Large and significant correlation between bubble measures and measures of overconfidence provide additional evidence that overconfidence has significant effect on price and trading behavior in experimental asset markets. Comparison of the selected bubble measures, averaged over five rational and overconfident markets, to the measures obtained in other experiments in which bubble-crash pattern was observed (e.g. Smith, Suchanek, and Williams, 1988) and the experiment of Noussair and Tucker (2006) in which bubbles were practically eliminated, suggests that there is evidence of the smaller deviations from the fundamental value in the rational market treatment than those observed in previous studies of markets of this type. To conclude, the analysis of the bubble measures demonstrates that although bubbles in the rational markets are not completely eliminated, they are less severe in comparison to the bubbles in overconfident markets. Moreover bubble measures increase with the increase in market overconfidence.

Although results presented in this paper shed some light on the effect of overconfidence on the processes in experimental financial markets, further investigation of the topic is desirable. A promising direction in research is examination of what proportion of overconfident subjects in the market is sufficient to influence price departures from fundamental value. For this purpose mixed markets of overconfident and rational subjects should be introduced. It would also be interesting to study if the results obtained were not dependent on the other factors, e.g. risk aversion (if higher overconfidence is not correlated to higher risk aversion). Finally, one could investigate the differences in personal behavior and outcomes on the individual level of the two types of players (rational and overconfident).

## REFERENCES

- Allwood, C.M., Granhag, P. A., Jonsson, A. C., (2006), Child witnesses' metamemory realism. *Scandinavian Journal of Psychology*, Vol. 47(6), p 461-447.
- Alpert, M., Raiffa, H., (1982), A progress report on the training of probability assessors. In Kahneman, D., Slovic, P., Tversky, A. (Eds.), *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, p. 294-305.
- Barber, B., and T. Odean, (2001), Boys Will Be Boys: Gender, Overconfidence and Common Stocks Investments. *Quarterly Journal of Economics*, 116, p. 261-292.
- Bar-Hillel, M., (2001), Subjective Probability Judgments. In Smelser, N.J. and Baltes, D.B. (Eds.) "International Encyclopedia of the Social and Behavioral Sciences", Amsterdam: Elsevier Science Ltd., p. 15247-15251.
- Baron, R. A., (2000), Psychological Perspectives on Entrepreneurship: Cognitive and Social Factors in Entrepreneurs' Success. *Current Directions in Psychological Science*, Vol. 9(1), p. 15-18.
- Benos, A., (1998), Aggressiveness and Survival of Overconfident Traders. *Journal of Financial Markets*, Vol. 1, p. 353-383.
- Biais, B., Hilton, D., Mazurier, K., Pouget, S., (2005), Judgmental overconfidence, self-monitoring and trading performance in an experimental financial market. *Review of economic studies*, Vol. 72(2), p. 287-312.
- Caballé, J. and Sákovics J., (2003), Speculating Against an Overconfident Market. *Journal of Financial Markets* Vol. 6, p. 199-225.
- Cambridge, R. M., Shreckengost, R. C., (1978), Are you sure? The subjective probability assessment test. Langley, VA: Office of Training, Central Intelligence Agency.
- Clarke, F. R., (1960), Confidence Ratings, Second-Choice Responses, and Confusion Matrices in Intelligibility Tests. *Journal of the Acoustical Society of America*, Vol. 32(1), p. 35-46.
- Cohen, J., (1988), *Statistical Power Analysis for the Behavioural Sciences* (2nd ed.). New Jersey: Lawrence Erlbaum, 567 p.
- Daniel, K., Hirshleifer D., Subrahmanyam A., (1998), Investor Psychology and Security Market Under-And Overreactions. *Journal of Finance*, Vol. 53(6), p 1839-1885.



- Deaves, R., Lüders, E., Luo, G. Y., (2009), An Experimental Test of the Impact of Overconfidence and Gender on Trading activity. *Review of finance*, Vol. 13(3), p. 555–575.
- De Bondt, W. F. M., Thaler R., (1984), Does the Stock Market Overreact? *The Journal of Finance*, Vol. 40(3), p. 793-805.
- De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., (1991), The Survival of Noise Traders in Financial Markets. *The Journal of Business*, Vol. 64(1), p. 1-19.
- Dufwenberg, M., Lindqvist, T., Moore, E., (2005), Bubbles and Experience: An Experiment, *The American Economic Review*, Vol. 95(5), p. 1731-1737.
- Fischbacher, U., (2007), z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, Vol. 10(2), p. 171-178.
- Fischhoff, B., Slovic, P., Lichtenstein, S., (1977), Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 3, p. 552-564.
- Fraser, S., Greene, F. J., (2006), The Effects Of Experience On Entrepreneurial Optimism And Uncertainty, *Economica*, Vol. 73(290), p 169-192.
- Gervais, S., Odean, T., (2001), Learning to be Overconfident. *Review of Financial Studies*, Oxford University Press for Society for Financial Studies, Vol. 14(1), p. 1-27.
- Glaser, M., Weber, M., (2007), Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, Vol. 32(1), p. 1-36.
- Griffin. D, Brenner, L., (2005). Perspectives on Probability Judgment Calibration. In Koehler, D. J., Harvey, N. (Eds.) “Blackwell Handbook of Judgment and Decision Making”, p. 177-199.
- Haruvy, E., Lahav, Y., Noussair, Ch., (2007), Traders’ Expectations in Asset Market: Experimental Evidence. Working Paper.
- Hazard, T. H., Peterson, C. R., (1973), Odds versus probabilities for categorical events. Tech. Rep. No. 73-2, McLean, VA: Decisions and Designs, Inc.
- Hirshleifer D., Subrahmanyam, A., Titman S., (1994 ), Security Analysis and Trading Patterns When Some Investors Receive Information Before Others. *Journal of Finance*, Vol. 49(5), p. 1665-1698.
- Hvide, H. K., (2002), Pragmatic Beliefs and Overconfidence. *Journal of Economic Behavior and Organization*, Elsevier, Vol. 48(1), p. 15-28.

- Hynes, M., Vanmarcke, E., (1976), Reliability of embankment performance predictions. Proceedings of the ASCE Engineering Mechanics Division Specialty Conference, Waterloo, Ontario, Canada: University of Waterloo Press.
- Kim, A. K., Nofsinger, J. R., (2003), The Behavior and Performance of Individual Investors In Japan. Working paper.
- Kirchler, E., Maciejovsky, B., (2002), Simultaneous Over- and Underconfidence: Evidence from Experimental Asset Markets. *Journal of Risk and Uncertainty*, Springer, Vol. 25(1), p. 65-85.
- Koriat A, Lichtenstein S, Fischhoff B., (1980), Reasons for Confidence. *Journal of Experimental Psychology: Human Learning and Memory*, Vol. 6, p. 107-18.
- Lakonishok, J., Shleifer, A., Vishny, R. W., (1992), The Impact of Institutional Trading on Stock Prices. *Journal of Financial Economics*, Elsevier, Vol. 32(1), p. 23-43.
- Lei, V., Noussair, Ch. N., Plott, Ch. P, (2001), Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality. *Econometrica*, Vol. 69(4), p. 831-859.
- Lichtenstein, S., and Fischhoff, B, (1980), Training for Calibration. *Organizational Behavior and Human Performance*, Vol. 26, p. 149-171.
- Lichtenstein, S., Fischhoff, B., Phillips, L. D., (1977), Calibration of Probabilities: The State of the Art. In Jungermann, H., deZeeuw, G. (Eds.) “Decision Making and Change in Human Affairs”, Amsterdam: D. Reidel.
- Lichtenstein, S., Fischhoff, B., Phillips, L. D., (1982), Calibration of Probabilities: The State of the Art to 1980. In Kahneman, D., Slovic, P., Tversky, A. (Eds.) “Judgment under Uncertainty: Heuristics and Biases”, Cambridge: Cambridge University Press, p. 306-334.
- Milgrom, R., Stokey N., (1982), Information, Trade and Common Knowledge. *Journal of Economic Theory*, Vol. 26(1), p. 17-27.
- Menkhoff, L., Schmidt, U., Brozynski, T., (2006), The Impact of Experience on Risk Taking, Overconfidence, and Herding of Fund Managers: Complementary Survey Evidence. *European Economic Review*, Vol. 50(7), p. 1753-1766.
- Moore, P. G., (1977), The manager’s struggle with uncertainty. *Journal of the Royal Statistical Society*, Vol. 140, p. 129-165.
- Neale, M. A., Bazerman, M. H., (1990), *Cognition and rationality in negotiation*. New York: The Free Press, p. 240.

- Noussair, Ch., Tucker, S., (2006), Futures Markets and Bubble Formation in Experimental Asset Markets. *Pacific Economic Review*, Vol. 11(2), p. 167–184.
- Odean, T., (1998), Volume, Volatility, Price and Profit When All Traders are Above Average. *Journal of Finance*, Vol. 53(6), p. 1887 -1934.
- Odean, T., (1999), Do Investors Trade Too Much? *American Economic Review*, Vol. 89(5), p. 1278-1298.
- Oskamp, S., (1962), The relationship of clinical experience and training methods to several criteria of clinical prediction. *Psychological Monographs*, Vol. 76(547), p. 1-28.
- Pitz, G.F., (1974), Subjective probability distributions for imperfectly known quantities. In Gregg, L. W. (Ed.), *Knowledge and Cognition*. New York: Wiley, p. 29-41.
- Powel, W. D., Bolich, C., (1993), Changing predictions based on knowledge of past performance in preschool and young grades-school. *Child Study Journal*, Vol. 23(3), p. 209.
- Porter, D., Smith, V., (1995), Futures Contracting and Dividend Uncertainty in Experimental Asset Markets. *Journal of Business*, Vol. 68, p. 509–541.
- Russo, J. E., Schoemaker, P. J. H., (1992), Managing Overconfidence. *Sloan Management Review*, Vol. 33, p.7-17.
- Scheinkman, J. A., Xiong, W., (2003), Overconfidence and Speculative Bubbles. *Journal of Political Economy*, Vol. 111, p. 1183-1219.
- Shiller, R. J., (2000), *Irrational Exuberance*. Princeton, N.J.: Princeton Univ. Press, 344 p.
- Sieber, J. E., (1974), Effects of Decision Importance on Ability to Generate Warranted Subjective Uncertainty. *Journal of Personality and Social Psychology*, Vol. 30, p. 688-694.
- Smith, V. L., Suchanek G. L., Williams A. W., (1988), Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econometrica*, Vol. 56(5), p. 1119-1151.
- Smith, V. L., (1994), Economics in the Laboratory. *Journal of Economic Perspectives*, American Economic Association, Vol. 8(1), p. 113-31.
- Staël von Holstein, C. A. S., (1972), Probabilistic forecasting: An experiment related to the stock market. *Organizational Behavior and Human Performance*, Vol. 8, p. 139-158.
- Statman, M., Thorley, S., Vorkink, K., (2006), Investor Overconfidence and Trading Volume. *Review of Financial Studies*, Vol. 19, p. 1531–1565.

Tversky, A., Kahneman, D., (1982), Judgment under Uncertainty: Heuristics and Biases. In Kahneman, D., Slovic, P., Tversky, A. (Eds.), Judgment under Uncertainty: Heuristics and Biases, Cambridge University Press, p. 3-20.

Van Boening, M., Williams A., LaMaster, S., (1993), Price Bubbles and Crashes in Experimental Call Markets. Economics Letters, Vol. 41, p. 179–185.

Wagenaar, W. A., Keren, G., (1986), Does the expert know? The reliability of predictions and confidence ratings of experts. In E. Hollnagel, G. Maneini, and D. D. Woods (Eds.), Intelligent decision support in process environments, Berlin: Springer, p. 87-107.

Weinstein, N. D., (1980), Unrealistic Optimism about Future Life Events. Journal of Personality and Social Psychology, Vol. 39, p. 806-820.

White, E. N., (2006), Bubbles and Busts: The 1990s in the Mirror of the 1920s. NBER Working paper Nr. W12138.

Zakay, D., Glicksohn, J., (1992), Overconfidence in a multiple-choice test and its relationship to achievement. Psychological Record, Vol. 42(4), p 519-524.

APPENDIX 2.A: BIAS SCORE OF THE VARIOUS (EXPERIMENTAL) SUBGROUPS

<i>Pre-experimental Test</i>					
OBS	Group	Mean	SD	Min	Max
201	All	11.78	10.57	-11.33	43.5
93	Female	9.62	10.68	-11.33	38.89
108	Male	13.37	10.28	-10.28	43.50
	Male vs. female diff.	3.75 (0.57)	--	--	--
<i>Experiments</i>					
OBS	Group	Mean	SD	Min	Max
60	All	11.20	12.08	-5.89	43.50
25	Female	9.96	12.45	-5.89	38.89
35	Male	12.08	11.91	-4.72	43.50
30	Overconfident	21.33	8.26	10.17	43.50
30	Rational	1.06	4.03	-5.89	6.78
	Male vs. female diff.	2.13 (0.81)	--	--	--
	Overconfident vs. rational diff.	20.27 (0.00)	--	--	--
OVE market	Male vs. female diff.	-0.65 (0.64)	--	--	--
RAT market	Male vs. female diff.	0.68 (0.76)	--	--	--

## **INSTRUCTIONS**

In this experiment we are going to create a market in which you will trade units of a fictitious asset (i.e. “shares” of a “stock”) that earn a dividend over a series of trading periods. The instructions are simple, and if you follow them carefully and make appropriate decisions **YOU MAY EARN A CONSIDERABLE AMOUNT OF MONEY** which will be **PAID TO YOU IN CASH** at the end of the experiment.

The currency used in the market is called Gulden. All trading and earnings will be in terms of Guldens. At the end of experiment, the Guldens that you have accumulated will be converted to euros at the exchange rate of 0.27 EUR for each 10 Guldens and you will be paid in euros. Note that the more Guldens you earn, the more euros you get!

### ***Duration of the experiment***

The market will take place over a sequence of 15 trading periods. You may think of each trading period as a “business or trading day”. Each trading period has a maximum length of 180 seconds at which time the market will close for that period. The remaining time left in each period will be shown by a clock on your computer screen.

The market period can be ended before the trading time expires by a **UNANIMOUS** vote of all participants in the market to end trading for that period. This alternative stopping rule allows the group as a whole to bypass the usual 180 second stopping rule. Each participant can vote by pressing the key labeled **VOTE**. Pressing **VOTE** and thus voting to end that market period does not eliminate you from participating further in trading for that period; it simply says that you are ready to end trading in the current period and move on to the next period.

### ***Initial Endowments of Participants***

Each trader at the beginning of the trading game is endowed by **STARTING CAPITAL** equal to 300 Guldens and 3 units of assets. During the experiment you may purchase or sell assets. At the **END** of each trading period you will receive a **DIVIDEND** on **EACH UNIT** asset unit in your inventory.

### ***Dividend Process***

You will not know the exact value of your dividend per unit prior to the end of each trading period. At the end of each trading period you will be told the value of your dividend per unit

and your dividend earnings (dividend earnings = assets × dividend per unit). They will be added to your working capital.

Your dividends are drawn randomly each period. The possible values of your dividend per unit and the associated probability of occurrence are given below:

dividend	0.0 Gulden	0.8 Gulden	2.8 Gulden	6 Gulden
probability	1/4	1/4	1/4	1/4

Thus, the average dividend over many draws is 2.4 Gulden ( $=0.0 \cdot 1/4 + 0.8 \cdot 1/4 + 2.8 \cdot 1/4 + 6 \cdot 1/4$ )

Before each trading period information on potential income from holding your assets till the end of the experiment (15<sup>th</sup> period) is provided to assist you in formulation of your market decisions. The following information is given to you: maximum, average and minimum possible dividends (the same in each period), and maximum, average and minimum earnings per inventory unit over the remaining experiment periods.

### ***Reward scheme***

Your decisions regarding the purchase and sale of asset units and your end-of-period inventory level (dividend earnings = dividend per unit × end-of-period inventory) should rest on the fact that at the end of the experiment your cash earnings are based on your final working capital which equals:

(300 Gulden starting capital) + (dividend earnings) + (asset sales revenue) - (asset purchase cost).

At the end of the game your assets have no value!

### ***The rules of the Experimental Market***

Suppose we open the market for Trading Period 1 and that you wish to enter your bid or offer. To enter bid (price at which you wish to buy an asset): type in the price for which you wish to buy an asset. Then click the box labeled “ENTER BID”. To enter offer (price at which you wish to sell an asset): type in the price at which you wish to sell your asset and then click on the box “ENTER OFFER”.

Notice that bids are going to be ranked in the decreasing order on the right side of the screen, and sale offers in the increasing order on the left-hand side of the screen.

Suppose now, that you wish to accept Seller’s offer and purchase one unit of the asset. To do this first click the appealing price, standing in the column named “SALES OFFERS”, and

then click the button labeled “ACCEPT OFFER”. If you wish to accept Buyer’s bid click on the appealing price, standing in the column “BIDS” and then click the button labeled “ACCEPT BID”. Note that after a contract has been made, all bids and offers are erased and a new auction begins.

Upon buying/selling one unit of the commodity the transaction price (sales or purchase) will be added to (if you have sold), or subtracted from (if you have bought) your working capital immediately, same is valid for the assets’ inventory.

Your inventory at the end of a trading period is carried over to the beginning of the next trading period. At the end of each trading period your working capital will be increased by the amount of your dividend earnings (dividend earnings = number of units in your inventory × dividend per unit).

You can buy asset units as long as your working capital is greater than or equal to the purchase price. If you attempt to enter a bid or accept a seller’s offer that is greater than your working capital, the action will be ignored and you will receive an error message on your display screen. You can sell assets as long as your inventory is greater than zero. If you attempt to enter an offer or accept a buyer’s bid, when you have no assets in your inventory, the action will be ignored and you will receive an error message on your display screen.

### ***Market Information***

At the end of each trading period you will have the opportunity to see the market price summary information from the past trading periods, which will include such information as average market contract price, the highest, and the lowest market price, volume traded and dividend for that period.



### ***Additional Means to Earn***

At the end of each trading period you will be asked to enter a forecast of the average contract price in the next trading period. Information on the current period's mean price will be available for your inspection prior to entering a forecast. Information on your forecasting accuracy, consisting of the actual price, and your price forecast from the past periods will be available to your inspection after entering a forecast.

You will be paid for your predictions, based on their accuracy. The closer the prediction is to the actual average market price, the higher is the reward. Reward scheme for predictions' accuracy:

Level of Accuracy	Earnings
+/- 5% from the actual price	3 Gulden
+/- 12.5% from the of actual price	1 Gulden
+/- 25% from the actual price	0,5 Gulden

Your income from "forecasting part" will be converted to euros at the same rate as mentioned above and paid to you at the conclusion of the experiment.

In the gap marked "Confidence level" you have to write how confident you are that your price forecast is correct! You can use any number between 0% and 100% to express your confidence, that your forecast is correct. Thus 0% means that you completely do not believe that your forecast can be true, and 100% means that you are completely sure that your Forecast will be correct.

This is the end of the instructions!

---

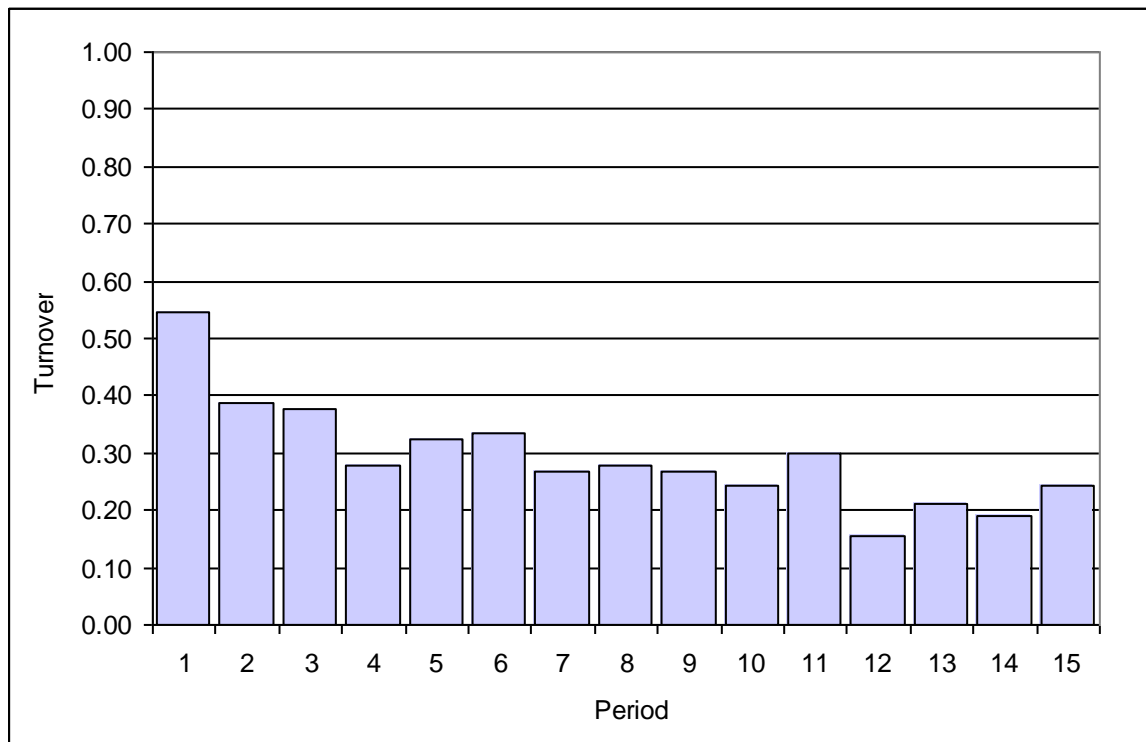
If you have a question that was not fully answered by the instructions please raise your hand and ask the experiment monitor before proceeding.

---

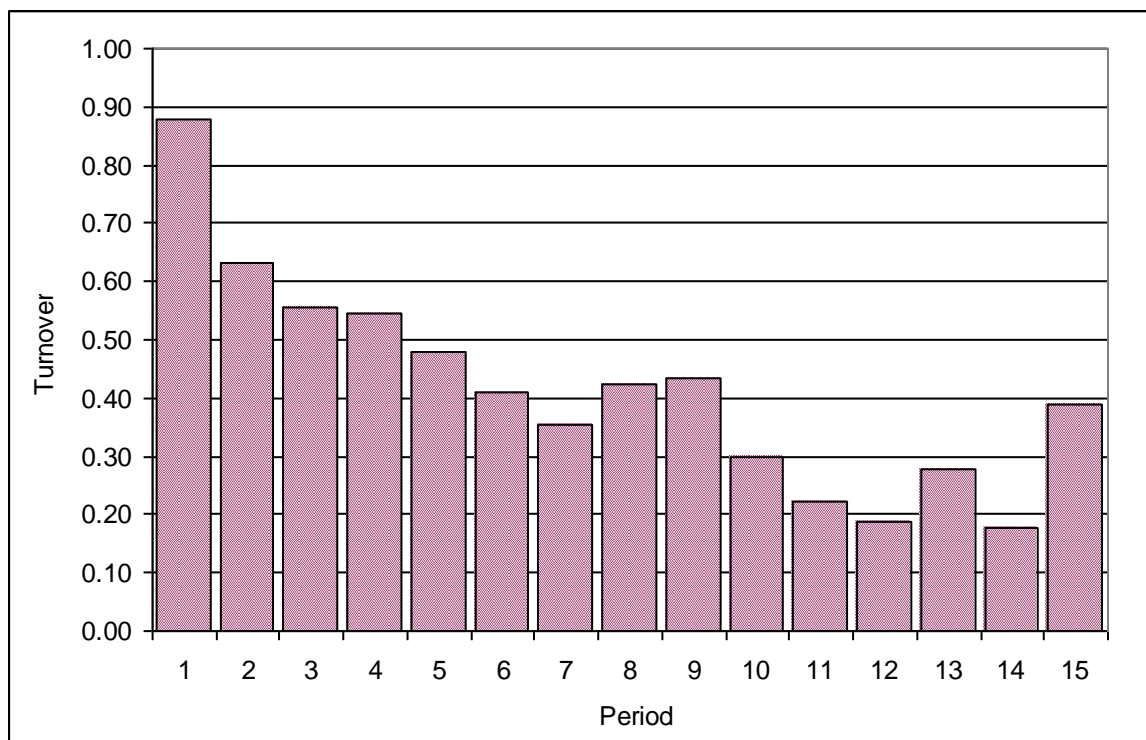
**BEWARE! YOUR EARNINGS MAY SUFFER IF YOU PROCEED INTO THE  
MARKETPLACE WITHOUT UNDERSTANDING THE INSTRUCTIONS!**

---

APPENDIX 2.C: JOINT AVERAGE TURNOVER DEVELOPMENT (a. Rational market, b. Overconfident market)



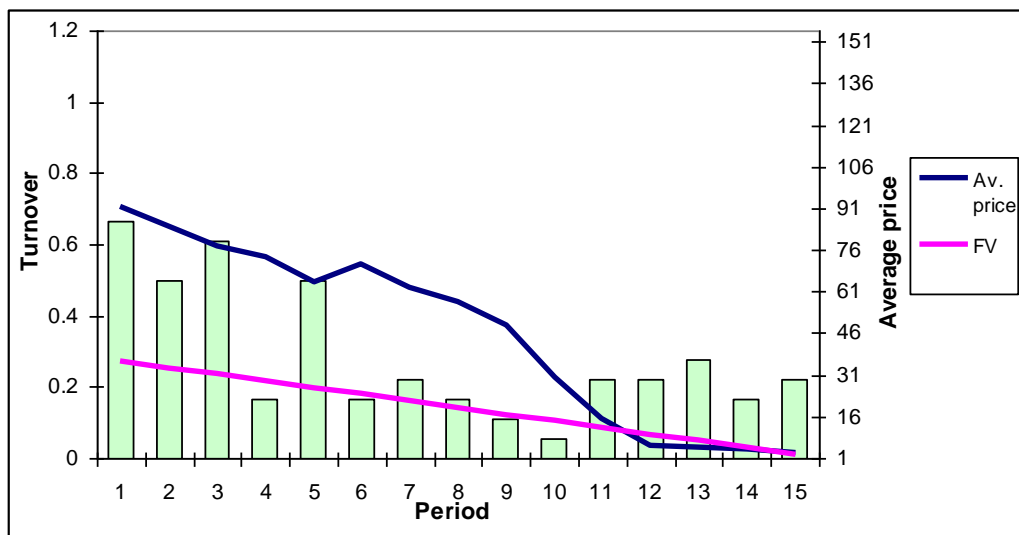
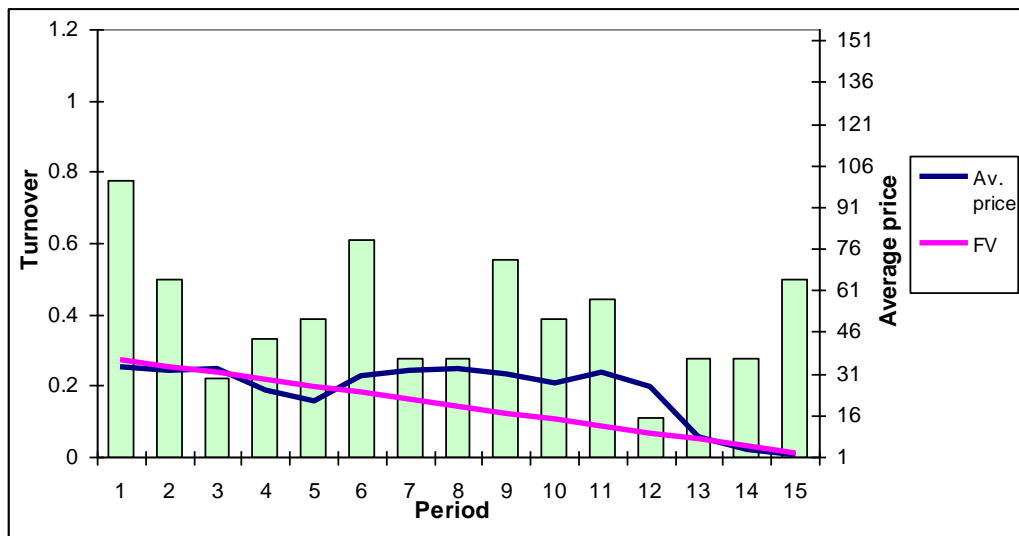
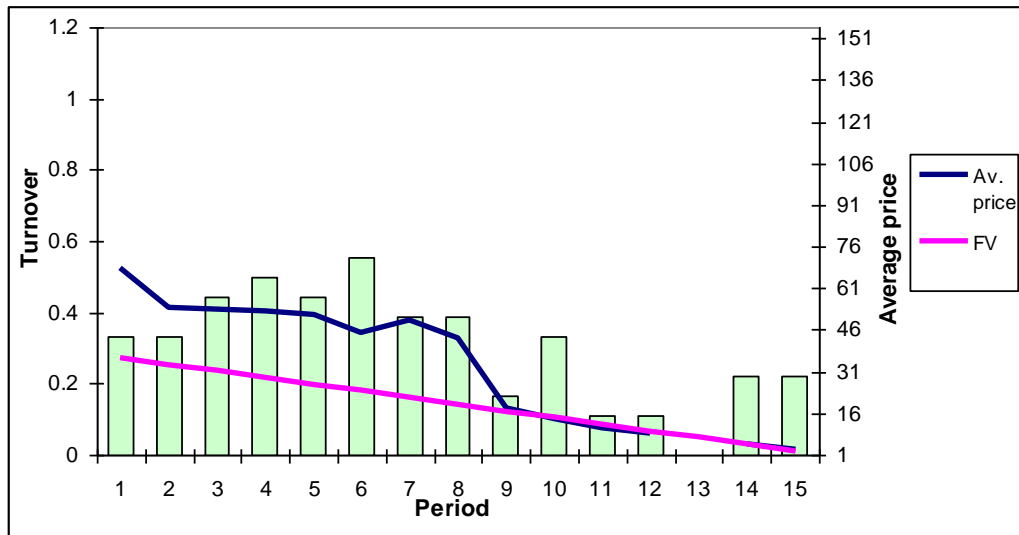
a.

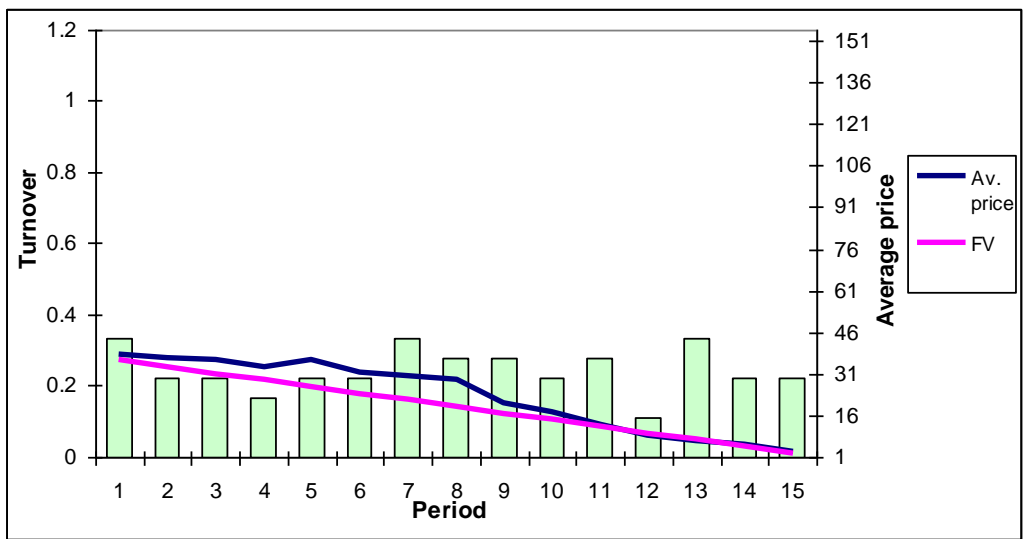
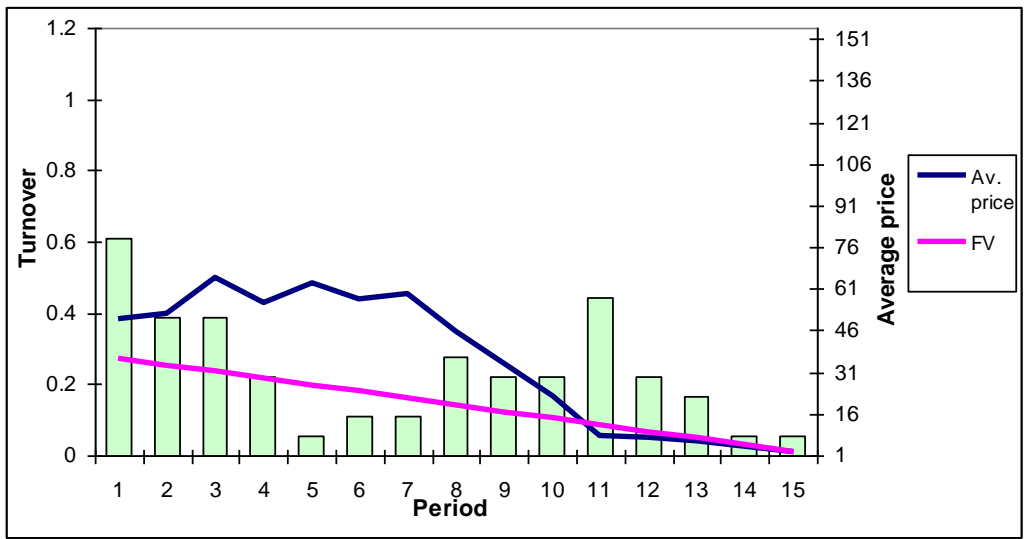


b.

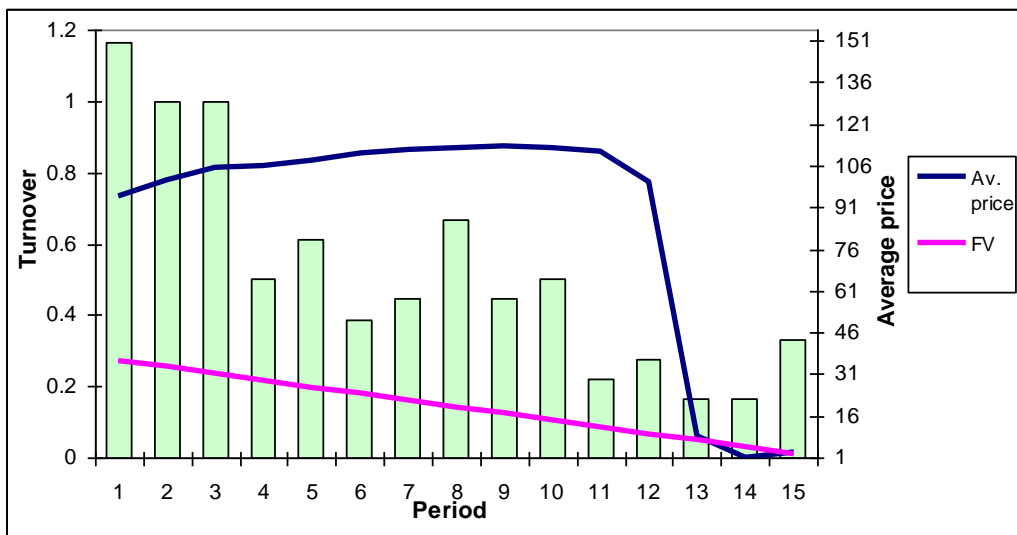
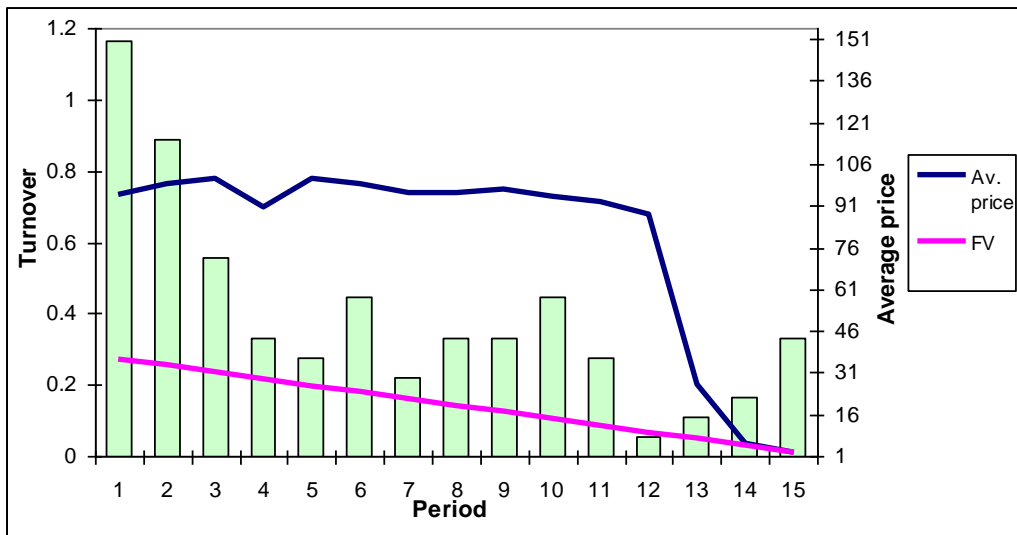
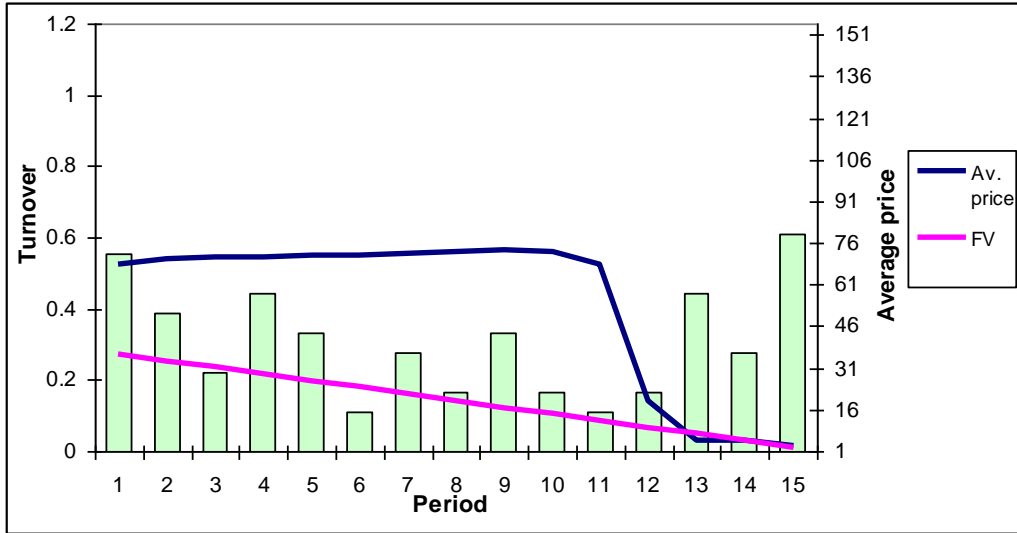
APPENDIX 2.D: DEVELOPMENT OF AVERAGE PRICE AND TRADE VOLUME IN EACH MARKET

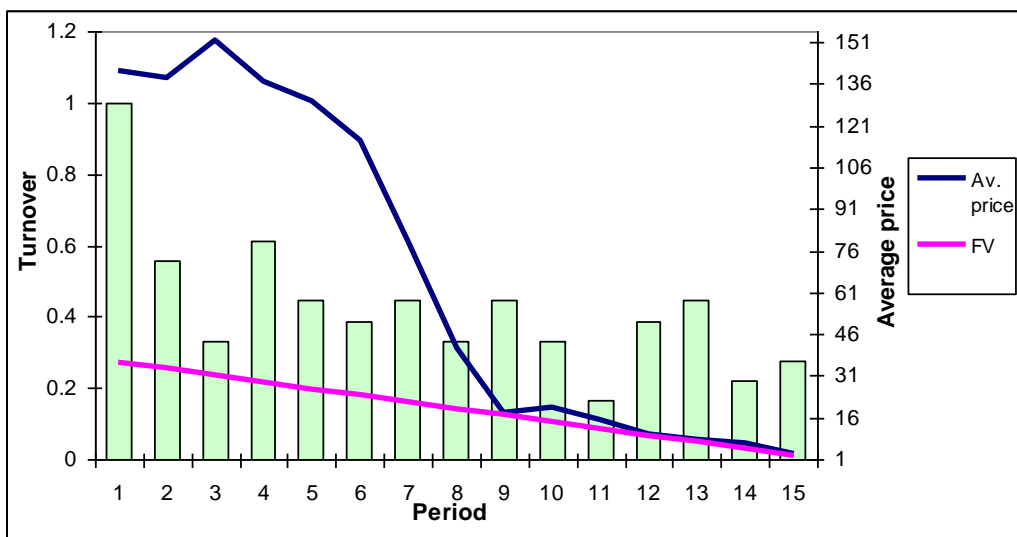
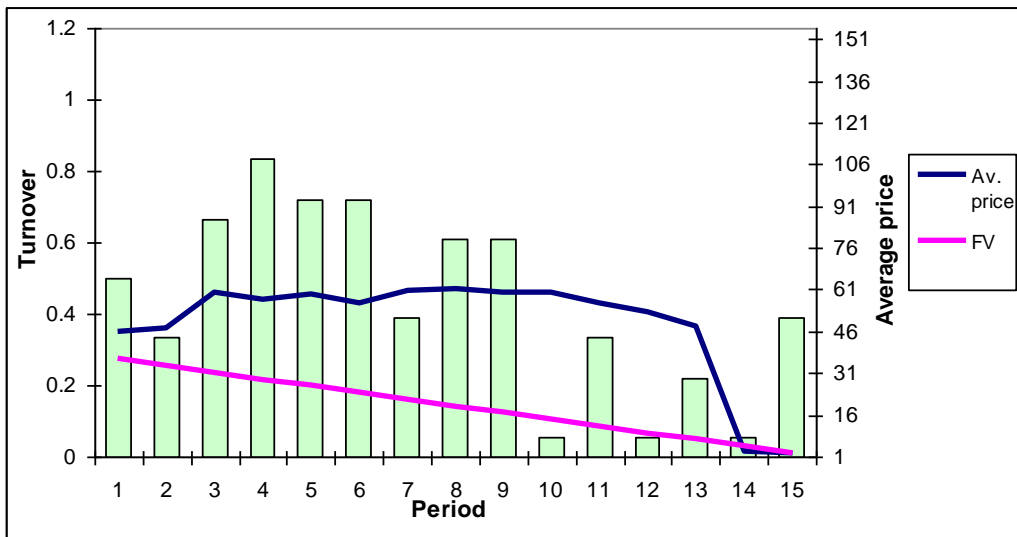
*Rational Markets*





Overconfident Markets





APPENDIX 2.E: SPEARMAN'S RHO CORRELATION COEFFICIENT BETWEEN BIAS SCORES AND BUBBLE MEASURES

	BS (pre-experimental)	BS (forecasting)
Hassel R2	-0.770 (p<0.05, one-sided)	-0.673 (p<0.05, one-sided)
NPD	0.745 (p<0.01, one-sided)	0.636 (p<0.05, one-sided)
NAD	0.745 (p<0.01, one-sided)	0.636 (p<0.05, one-sided)
Velocity	0.717 (p<0.01, one-sided)	0.550* (p < 0.1, one-sided)
Amplitude	0.661 (p<0.05, one-sided)	0.515 (p<0.1, one-sided)

\* - based on 9 observations

## CHAPTER 3

# OVERCONFIDENCE, RISK AVERSION AND (ECONOMIC) BEHAVIOR OF INDIVIDUAL TRADERS IN EXPERIMENTAL ASSET MARKETS



### 3.1 INTRODUCTION

By allowing psychological bias and emotion to affect their investment decisions, investors can do serious harm to their wealth (Baker and Nofsinger, 2002). One of such biases, inducing deviation from rational economic behavior, is overconfidence. Kahneman and Riepe (1998) emphasized the importance of overconfidence in financial decision-making and suggested that, combined with optimism, it would make individuals overestimate their knowledge, underestimate risks, and exaggerate their ability to control events. Overconfidence in investors can result in aggressive trade (e.g. Deaves et al., 2009), portfolio undiversification (e.g. Odean, 1999), pursuit of the active portfolio management strategy (e.g. De Bondt and Thaler, 1994) and suboptimal performance (e.g. Fenton-O’Creevy et al., 2003). Another personality trait, which determines individual investment choices and strategy, is the degree of risk aversion. Some authors suggest existence of a direct link between overconfidence and risk aversion (see Russo and Shoemaker, 1992; De Long et al., 1991), namely that greater overconfidence leads to risk underestimation and excessive risk-taking by investors e.g. by engagement in aggressive trade or choice of riskier portfolios.

Interest in the topic of economic consequences of investors’ overconfidence (irrationality) generated a large body of literature. Scheinkman and Xiong (2004) point out, that an issue occupying attention of researchers starting with Friedman (1953) “is whether traders with irrational beliefs will lose money trading with rational traders and eventually disappear from the market?” Consequently most of the foregoing research was focused on the analysis of the “mixed” asset market setting, where both overconfident and rational traders interacted. According to a widespread opinion, in such markets rational investors will take advantage of the overconfident ones (noise traders) and the former will eventually incur losses and die out (e.g. Hirshleifer and Luo, 2001). This proposition is supported by findings from experimental markets of this type, which present evidence that overconfident subjects, compared to rational ones, engage in more trading and face reductions to their welfare. Several articles, examining the effect of overconfidence on variation in subjects’ trading activity and performance in the “mixed” market setting, which are the closest in spirit to mine are by Biais, Hilton, Mazurier, and Pouget (2005), Glaser and Weber (2007), Deaves, Lüders and Luo (2009), and Kirchler and Maciejovsky (2002). Biais et al. (2005), presented empirical evidence of the negative association between overconfidence and traders’ performance; yet it had a more significant effect on males’ than on females’ performance. Overconfidence did not lead to increase in trading activity in their sample. In the study of Glaser and Weber (2007) overconfidence was assessed both as miscalibration score and the better than average effect. The former was found

to be unrelated to trading volume, however investors who thought that they were above average traded more. Linear relationship between portfolio performance, overconfidence, and trade frequency was not detected. Results of Deaves et al. (2009) indicated that greater overconfidence has led to increased trading activity and had a negative effect on individual trading performance. Kirchler and Maciejovsky (2002) discovered no significant difference in terms of risk attitude between their experimental markets, and concluded that any distinctions between experimental outcomes were not risk-attitude dependent.

My experiment was created with the following assumptions in mind. First, previous studies have created at least two sources of exogenous heterogeneity in their experiments by: 1) construction of “mixed” markets, consisting of both overconfident and rational traders and by 2) asymmetries in the information and/or its quality, which different types of players received from the experimenter. Conclusions of these papers were mainly based on the antagonizing principle, that one group would take advantage of the other. Second, following the examination of the instruments, which were used by previous authors to assess subjects’ overconfidence, there were good reasons to suspect that overconfidence was measured inadequately and that these instruments did not offer comprehensive measure of individual miscalibration. And last but not least, to my best knowledge, there were no other experiments that tried to explore the link between individual risk aversion, overconfidence and market outcomes. The only experiment that measured both subjects’ overconfidence and risk aversion was by Kirchler and Maciejovsky (2002), but they rather focused on the market-level distinctions. My paper will differ from the foregoing research in some important fashion. In contrast to the abovementioned papers in present experiment subjects, based on their pre-experimental overconfidence scores, were assigned to the two types of markets. Individuals with the lowest score formed “rational” markets and individuals with the highest score “overconfident” markets. In the course of experiment participants have interacted only with subjects of their own “type”. Within each of the conducted markets, the degrees of individual overconfidence (slightly) varied. This was the only source of subjects’ heterogeneity, as in my experiment participants have differed neither in terms of their initial endowments, nor in terms of information quality and access to it. Subjects’ overconfidence was assessed through a specially created test, weighted for the inclusion of easy, hard and medium difficulty questions, which has also accounted for the possible gender and country biases. I have also chosen a different test format which, due to its simplicity, was clearer to subjects. None of the previous experiments made use of such test.

This paper reports the results of experiments designed to investigate the influence of behavioral factors, namely the degree of overconfidence and risk aversion, on financial

decision making of economic subjects. For this purpose two kinds of experiments were conducted. The first one, whose design followed Smith, Suchanek and Williams (1988) and Deaves et al. (2009), was the asset market experiment. The second experiment was aimed at individual risk aversion measurement and followed Holt and Laury (2002) and Baker, Laury and Williams (2008). Hypotheses, tested in the context of the suggested experimental design, were formulated upon the analysis of findings from the previous research in the “mixed” market setting. Data, collected in the first part of the experiment from five overconfident and five rational markets, enabled investigation of the effect that the degree of overconfidence had on individual trading activity and performance. In the second part of experiment, post hoc assessment of risk aversion was implemented in a sample of former participants of the asset market experiment. These data were used for the determination of the importance of individual differences in risk attitude for explanation of trading behavior and outcomes.

Main findings from my experiment can be summarized as follows. The data analysis confirmed the hypothesized positive impact that overconfidence had on individual engagement in trading activity; yet, with increase in overconfidence, females have completed more stock market transactions than males. Contrary to the formulated hypothesis, data suggest that overconfidence had positive effect on gains from trade. As it was expected, individual gains were negatively affected by active involvement in trade, and low turnover players have significantly outperformed high turnover players. However, with increase in the number of market transactions males incurred smaller losses compared to females. It can be concluded that in the suggested experimental setting, where the two “types” of subjects were separated from each other, performance and trading activity were overconfidence dependent and even small variations in miscalibration among players belonging to the same “type” were sufficient to cause this effect. At the completion of subsequent risk aversion measurements thirty two former participants of the asset market experiment were found to be on average risk averse. Inconsistent with the initial proposition that overconfident subjects were more risk loving, tests detected no statistical difference between the two types of participants in terms of the average number of safe choices. Also no linear relationship between the bias score and risk aversion was detected. The presented evidence implied that, risk aversion was not among the factors which had significant influence on engagement in trading activity or performance. It can be concluded, that in the reduced sample, experimental market outcomes were overconfidence and not risk aversion driven.

Paper proceeds as follows. Section 3.2 gives an overview of the findings of financial overconfidence and risk aversion literature; along analysis of the similar experimental papers

and discussion of the paper's contributions is presented. Section 3.3 lists the research hypotheses. Section 3.4 provides the description of experimental procedures of both experiments. Sections 3.5 and 3.6 present data analysis of, correspondingly, experimental asset market and risk aversion measurement experiment. Finally Section 3.7 concludes.

## 3.2 PSYCHOLOGICAL DETERMINANTS OF TRADING ACTIVITY AND PERFORMANCE

### 3.2.1 OVERCONFIDENCE

Interest in the topic of economic consequences of investors' overconfidence generated a large body of literature. According to Fischhoff, Slovic, and Lichtenstein (1977) most people hold unrealistic positive beliefs about their personal skills and their knowledge. Giardini et al. (2008) suggest that constant overestimation of personal talents and abilities, and chances of positive outcomes "can have important consequences, and sometimes results in suboptimal decisions". Kahneman and Riepe (1998) point out at the importance of overconfidence for financial decision taking, in that, overconfidence combined with optimism, produces overestimation of individual knowledge, exaggeration of the ability to control events, and risk underestimation. Barber and Odean (2000) indicate that overconfident investors "hold unrealistic beliefs about how high their returns will be". Thus cognitive bias of overconfidence creates various distortions in the way traders perceive objective market reality, and this can result in trade aggressiveness, portfolio undiversification, risk underestimation, pursuit of the active portfolio management strategy, and suboptimal performance (decrease in wealth, sales of the wrong assets, etc.).

Most of the research findings support the proposition that greater overconfidence leads to higher trading volume (see e.g. Deaves et al., 2009). Glaser et al (2003) showed that overconfident investors traded more aggressively, and the higher the degree of overconfidence of a trader was, the higher was her trading volume. Odean (1998) calls this outcome "the most robust effect of overconfidence". Odean (1999) analyzed the trades of 10,000 individuals with discount brokerage accounts. His results indicate that these investors reduced their returns by trading, and thus he concluded that their trading volume was excessive. Kourtidis et al., (2010) suggest that overconfidence is the fact that makes investors believe that they can predict the winners and thus makes them to engage in excessive trading activity and take too much risk. Barber and Odean (2000) analyzed a data set of trading activity of the 66,465 households for a period from 1991-1996, and found that the frequency

of trading and its cost explained the poor investment performance of the households, and that overconfidence was the explanation for the high trading volume of individual investors (and the resulting poor results). In a further study Barber and Odean (2001) showed that men, who according to their analysis of psychological research findings were more overconfident, traded 45% more than women, and that trading reduced their net returns by more percentage points per year than for women (2.65 vs. 1.72).

What concerns the welfare effect of overconfidence, the evidence from research is contradictory. Some authors indicate that, in comparison to rational investors, the expected utility (De Long et al., 1991; Odean, 1998) and welfare (Fenton-O’Creevy et al., 2003; Barber and Odean, 2002; Nöth and Weber, 2003) of overconfident traders is reduced. The results of Kirchler and Maciejovsky (2002) indicate that higher degree of overconfidence is negatively correlated with earnings of participants in their experiments. Deaves et al. (2009) showed that greater overconfidence lead to higher trading volume and earnings’ reduction. Likewise in the experiment of Biais et al (2004) miscalibration decreased subjects’ trading performance. Overconfidence was also found to have a negative impact on trading performance in the paper by Fenton-O’Creevy et al. (2003). Barber and Odean (2000) presented evidence, that high turnover households underperformed low turnover households, in terms of investment returns, by about 7.1% annually. On the contrary, Benos (1998) reported higher earnings of overconfident investors. Kyle and Wang (1997) indicated that the expected returns of overconfident traders might outperform those of rational traders, and that this could be obtained through the aggressive trading. Also, Hirshleifer and Luo (2001) presented evidence of higher profits of overconfident traders. The third strand of literature reports no connection between overconfidence and individual gains. E.g. Glaser and Weber (2007) detected no linear relationship between gross returns (portfolio performance), trade frequency and overconfidence. Thus they concluded that underperformance of investors who trade more is transaction costs driven.

De Bondt and Thaler (1994) suggest that overconfidence explains excessive trade of portfolio managers, and active portfolio management pursuit by financial economists and pension funds. According to Lakonishok et al. (1992) overconfident subjects choose active portfolio management. Also, Mullainathan and Thaler (2000) named overconfidence as the factor, helping to explain why, even the “most professionally managed portfolios are turned-over once a year or more”. They have also noticed that individual investors traded too much, and that assets sold by them tended to outperform the new, subsequently acquired, assets. De Bondt and Thaler (1994) presented evidence, that irrespective the fact that portfolio managers

usually underperformed index funds, most stock portfolios were still managed actively. As well, overconfidence makes investors choosing undiversified, thus riskier, portfolios (Odean, 1998, 1999; Lakonishok, Shleifer and Vishny, 1992). Scheinkman and Xiong (2003) argue that investors who perceive prospects for the future dividends from the assets as more optimistic will “bid up the price of the asset and eventually hold the total supply” of inventories.

Some authors suggest that there is direct link between overconfidence and risk aversion. Russo and Shoemaker (1992) note that overconfidence causes risk underestimation and encourage traders to take excessive risk. Odean (1998) has shown that investors with a higher degree of overconfidence chose in general more risky portfolios than those with a lower degree of overconfidence. Evidence presented by Barber and Odean (2001) implied that overconfident investors (men) held riskier portfolios. De Long et al. (1991), who examined traders that were overconfident in the sense that they underestimated risk, found that as a result of risk underestimation, these traders held more of the risky asset. Chuang and Lee (2005) presented empirical evidence in the support of the hypothesis that overconfident investors traded more in riskier securities, since they underestimated risk. Not only risk underestimation contributes to the more aggressive trade by overconfident investors, in comparison to rational traders, it is also boosted by the fact that they overestimate the expected profit from their trading strategies (Hirshleifer and Luo, 2001).

### 3.2.2 RISK AVERSION

Risk aversion is connected to the desire of individuals to avoid uncertainty. Yet, almost every economic decision involves some sort of risk and uncertainty. Thus analysis of individual differences in risk attitudes is important in a sense, that understanding these distinctions could help in predicting real economic behavior. A commonly used approximation for modeling individuals in economic theory is that of the risk neutral economic human. However, empirical and experimental studies suggest that most people exhibit risk aversion. In the experiment of Holt and Laury (2002) about two-thirds of subjects exhibited risk aversion with low payoffs. As payoffs grew, risk aversion tended to increase as well. Binswanger (1980), who conducted his experiment with low-income farmers in Bangladesh, showed that most farmers were significantly risk averse, and that their risk aversion increased with the increase in payoffs. Demographic factors such as age, gender, education, intelligence, etc., serve as determinants of individual differences in risk attitudes.

A large body of literature addresses the topic of gender differences in risk tolerance. Holt and Laury (2002) showed that women were more risk averse in comparison to men in low-payoff decisions. Brachinger et al. (1999) and Schmidt and Traub (2002) have presented evidence that females often had higher degree of risk and loss aversion than males. Menkhoff et al. (2006) found that female fund-managers showed higher degree of risk aversion. According to Dohmen et al., (2005) at all ages women are less willing to take risks. There is as well an interaction between subject's gender and marital status. In general, married individuals are more risk-averse (Roszkowski, 1998), where married females comprise the least risk tolerant group (Yao and Hanna, 2005), and single males - the most risk tolerant; single men are followed correspondingly by married males, and single females. Households headed by females are less likely to be risk tolerant, compared to households headed by men or married couples (Sung and Hanna, 1996). As well, gender has influence on subjects' investment choices, e.g. women would invest more often in risk-free assets (Hariharan et al., 2000). Compared to single men, single women tend to hold a smaller part of their wealth in the form of risky assets (Jianakoplos and Bernasek, 1998). However, an increase in knowledge in a financial decision making context can "create a near role reversal between men and women in attitudes toward uncertainty" (Gysler et al., 2002).

The level of obtained education is importantly related to subject's risk attitude. Grable (2000) pointed out that increased risk tolerance is associated with greater levels of attained education. The results of Sung and Hanna (1996) indicated that risk tolerance in their sample increased with the schooling degree, where the lowest predicted risk tolerance (43%) was in a group of subjects who did not graduate from high school, and the highest predicted risk tolerance (71%) was obtained for those subjects who had a college degree. Better education of parents also has an important influence on individual's risk attitude: individuals whose parents are highly-educated are more willing to take risk (Dohmen et al., 2005). Educational attainments are to a large extent determined by subjects' cognitive ability. Results, concerning the relationship between cognitive ability and risk taking, are rather mixed. Frederick (2006) presented evidence, that cognitive ability was positively correlated with willingness to take risk in lotteries when outcomes included gains, and negatively – when outcomes included losses. Benjamin et al. (2005) suggested that students with lower math scores were less likely to be risk neutral. However Dohmen et al., 2007 argued that people with higher cognitive ability were significantly more risk loving.

No definite conclusion regarding the relationship between experience, age and risk attitude can be drawn. Some studies suggest that less experienced subjects are willing to take more

risk (e.g. Menkhoff et al., 2006) whereas others argue that willingness to take risk increases with experience (e.g. Hong et al., 2000; Chevalier and Ellison, 1999). In the article by Hanna and Lee (1995) the predicted risk tolerance was approximately the same for all ages under 55. However, starting with 55, it decreased with an increase in age. The results of Dohmen et al., (2005) indicated that increase in subjects' age was negatively associated with their willingness to take risk. Authors note that "the impact of age implies increased financial conservatism in ageing societies" (Dohmen et al., 2005).

As in the case with overconfidence, there is some empirical evidence that risk attitude affects trading behavior. Higher propensity for risk taking manifests itself through an increase in trade frequency (Durand et al., 2006). Fellner and Maciejovsky (2007) in their experiment explored the relation between market activity and risk attitude. They concluded that higher degree of risk aversion was accompanied by lower market activity. As well, attitude to financial risk is a significant positive predictor of willingness to invest in stocks (Keller and Siergist, 2006). Camacho-Cuena, Requate and Waichman (2009) pointed out that risk attitude affected subjects' performance in the laboratory, where the probability of engagement in speculative activity increased with increment in risk tolerance. In general, females engage in market activity less actively than males, and also are more risk averse. However, according to Fellner and Maciejovsky (2007), one should be cautious in connecting gender differences (in terms of market behavior) only to risk aversion diversity. They argue that "second order characteristics", one of which is overconfidence, might influence risk attitude. Also Croson and Gneezy (2008) suggested, that differences in risk attitudes may be explained by differences in risk perception, where overconfidence can manifest itself via "reduced estimate of the riskiness of a given investment".

### 3.2.3 SIMILAR RESEARCH AND VALUE ADDED

This paper builds on several previous survey and experimental articles, which investigate the effect of the degree of overconfidence on variation in subjects' trading activity and performance (earnings), but with some new aspects in experimental design.

The approach of Biais et al. (2005) relies on the asymmetric information trading game, and focuses on determining the link between subjects' psychological characteristics and their earnings. For this purpose, their sample was broken in four quartiles in terms of overconfidence degree, and analysis of the average trading activity and profit patterns was conducted. Biais et al. (2005) showed that the association between subjects' miscalibration and their performance (trading profits) was negative, and this relation was robust across the



samples. Yet the strength of the impact differed, namely overconfidence had more significant effect on males' performance, than on females'. The reported results also indicated that overconfidence did not lead to an increase in trading activity neither in the whole sample, nor in the male and female sub-samples.

Glaser and Weber (2007) empirically tested the hypothesis of theoretical models that overconfident investors trade more than the rational ones. They correlated overconfidence scores of 215 investors, who responded to their on-line questionnaire, with several measures of trading volume. In their study overconfidence was assessed both as miscalibration score and the better than average effect. Glaser and Weber (2007) discovered that miscalibration scores were unrelated to trading volume. Yet investors who thought that they were above average traded more. Although only secondary to their main interest, authors have also tested for the existence of the linear relationship between gross returns (portfolio performance), overconfidence, and trade frequency. No such connection was discovered. Also, there was no link between gender and trading volume.

Experiment of Deaves et al. (2009) was aimed at finding support for the premise that overconfidence leads to enhanced trading activity. This hypothesis was tested both at individual and market levels by regressing, previously obtained, overconfidence scores on the measure of trading activity (assessed as executed trades). In their experiment Deaves et al. (2009) conducted a limited amount of sessions where subjects were assigned, based on their overconfidence degree, to the two "high overconfidence" and the two "low overconfidence" markets. In the remaining four sessions subjects were separated by gender, while roughly maintaining average overconfidence within the markets. The reported results indicate that greater overconfidence lead to increased trading activity and that overconfidence had a negative effect on trading performance (reduced earnings). No connection between overconfidence, trading activity and gender was discovered in their sample.

Kirchler and Maciejovsky (2002) run a multi-period experimental asset market and investigate the development of overconfidence in the course of experiment. Prior to opening of the market sessions, subjects' risk aversion measurement was implemented by the methods of certainty equivalents' assessment and lottery pair's choice. Overconfidence was measured before each trading period, through the two average market trading price forecasting tasks: point prediction and interval prediction. Participants of the experiment were found to be well-calibrated in certain periods, and under- or overconfident in other periods. Higher degree of subjects' overconfidence was negatively correlated with their earnings. Between the experimental groups there was no significant difference in terms of risk aversion, thus authors

concluded that any distinctions that were observed between the experimental conditions were not risk-attitude dependent.

Based on the analysis of previous research, my experiment was designed with the following assumptions in mind:

First of all, foregoing researchers have created at least two sources of exogenous heterogeneity in their experiments by: 1) construction of “mixed” markets, consisting of both overconfident and rational traders and by 2) asymmetries in the information and/or its quality, which different types of players received from the experimenter. According to Glaser et al. (2003), such heterogeneity between investors could have created a potential for trade.

Overconfident or rational subjects were not the only type of players in the market in most of the above-mentioned papers, and their conclusions were mainly based on the antagonizing principle, that one group would take advantage of the other. Thus, none of the prior experimenters have employed such market construction principle in which miscalibrated and well-calibrated subjects were separated from each other. E.g. participants of the study by Glaser and Weber (2007) operate in real markets where one cannot control for the differences in traders’ endowments, information, or for the trade between overconfident and rational traders. Although Deaves et al. (2009) run four sessions to which subjects were assigned, based on their degree of overconfidence, for the analysis purposes these data were combined with the data obtained from another treatment (four gender-based markets). Not to mention that in their experiment different overconfidence measurement methodology and market structure were employed. In my experiment subjects were assigned to the overconfident and rational markets, based on their miscalibration scores, before the start of the experimental sessions, and in the course of experiment they interacted only with subjects of their own type. Subjects with the highest score comprised “overconfident” markets, and subjects with the lowest score - “rational” markets.

A second source of exogenous heterogeneity in the previous experiments was created by supplying participants with asymmetric pieces of information, which also differed in their quality. E.g. in the experiment of Kirchler and Maciejovsky (2002) half of the participants had no information about the dividend distribution, and the other half had complete information. Subjects of Biais et al. (2005) received three different types of private signals: bullish, bearish, and neutral. Deaves et al. (2009) also supplied their participants with signals of different quality. Moreover, they tried to manipulate subjects’ beliefs and make them think that their signals were more accurate than these of the other subjects. In my experiment all traders were given the same information, and no artificial belief of possessing a better

information piece was created. Likewise, my participants did not differ in terms of the initial endowments<sup>25</sup>.

Second, following the examination of the instruments, which were used in prior research to assess subjects' overconfidence, there were good reasons to suspect that overconfidence was measured inadequately and that these instruments did not offer comprehensive measure of individual miscalibration<sup>26</sup>. Most of the afore-named researchers followed the famous work by Russo and Schoemaker (1992) and used confidence interval elicitation tasks to measure overconfidence. However, Klayman et al. (1999) argue, that interval estimation tasks are prone to produce extreme miscalibration. E.g. in the experiment of Deaves et al. (2009) none of the subjects got close to the perfect calibration measure, and even the best calibrated participants exhibited quite high levels of overconfidence. Findings from psychological research also indicate, that overconfidence is the most pronounced for hard questions (a few persons know the right answer), and the least pronounced for the easy questions (the right answer is known to many persons). However, questions' difficulty was not assessed in the foregoing research, and the constructed scales were not pre-tested prior to their application for experimental measurements. In my experiment, a specially created test, weighted for the inclusion of easy, hard and medium difficulty questions, was used to estimate subjects' overconfidence. This instrument has also accounted for the possible gender and country biases, and, compared to some authors, included more items. Unlike previous authors, I chose another test format, which was not inherently prone to production of extreme overconfidence levels and, due to its simplicity, was clearer to subjects – multiple choice discrete propositions' task format. Prior to experimental use, the developed scale was pre-tested with students enrolled in different disciplines of social sciences. Overconfidence estimation phase was administered and financially rewarded, which increased reliability of the measurements.

And last but not least, to my best knowledge, there were no other experiments before that tried to explore the link between subjects' risk aversion, overconfidence and experimental outcomes. The only experiment that measured both subjects' overconfidence and risk aversion was by Kirchler and Maciejovsky (2002). However analysis of the individual differences among subjects in terms of risk aversion was not the aim of their research; in contrast, they focused on the existing group (market-level) differences. Dependence between

---

<sup>25</sup> However I cannot completely avoid heterogeneity of subjects as within each of the constructed rational and overconfident markets subjects still differ from each other (even slightly) in terms of the bias scores.

<sup>26</sup> More information about the overconfidence measurement instrument is in the paper "Development of the Overconfidence Measurement Instrument for the Economic Experiment".

experimental outcomes, overconfidence and risk aversion was also not in their focus of interest.

### 3.3 HYPOTHESES

Hypotheses, to be tested in this article, are built on the analysis of findings from overconfidence research in financial markets, and some of them were previously tested in the “mixed” market setting, where both types of subjects (overconfident and rational) interacted.

To begin with, there is plenty of evidence that greater overconfidence leads individuals to engage in more trading activity (e.g. Odean, 1998, Deaves et al., 2009) or pursue active portfolio management (Lakonishok, Shleifer and Vishny, 1992). The mentioned findings suggest testing of the hypothesis that in my experimental setting the higher degree of subject’s overconfidence is also accompanied by more active engagement in trading activity.

*Hypothesis 1:* Individual trading activity increases with the increase in overconfidence, measured as the bias score.

Experimental findings from the “mixed” market setting suggest that overconfident traders, who engage in trade more actively, incur losses. In other words, they are outperformed by low turnover traders (e.g. Barber and Odean, 2000). These results raise a question of whether in my experiment, where overconfident and rational traders are separated from each other, excessive trading activity still has negative impact on traders’ performance.

*Hypothesis 2:* High turnover traders underperform low turnover traders, i.e. there is negative relationship between individual gains from trade and trading activity.

Upon analysis of the foregoing papers, it can be suspected that the final portfolio size is positively connected to the degree of overconfidence. E.g. model of Scheinkman and Xiong (2003), suggests that investors who perceive prospects for the future dividends from the assets as more optimistic will “bid up the price of the asset and eventually hold the total supply” of inventories. Likewise, some authors noted that overconfident investors chose undiversified portfolios (Odean, 1998, 1999; Lakonishok, Shleifer and Vishny, 1992). In this regard, I posit the hypothesis that the cognitive bias of overconfidence has positive influence on the asset portfolio size at the end of the experiment.

*Hypothesis 3:* Increase in the number of assets in trader’s final inventory is accompanied by growth of her bias score.

In the experimental asset markets, where overconfident traders trade against rational traders, higher degree of traders’ overconfidence reduces their welfare (e.g. Fenton-O’Creevy et al.,

2003; Biais et al., 2005; Nöth and Weber, 2003; Kirchler and Maciejovsky 2002). In line with these results, it is important to explore whether in the context of suggested experimental design, subjects' market performance deteriorates with the increment in their miscalibration.

*Hypothesis 4:* Individual gains from trade decrease with the greater degree of overconfidence. Foregoing research presented results, implying existence of the link between overconfidence and risk aversion, in the form of risk underestimation (e.g. Ruso and Schoemaker, 1992; Kahneman and Riepe, 1998) and increased risk taking by overconfident investors, e.g. by riskier portfolios choices (Odean, 1997, 1998; Lakonishok, Shleifer and Vishny, 1992). Risk loving individuals also engage in trading activity more actively (Durand et al., 2006) and are more willing to invest in stocks (Keller and Siergist, 2006). If, on average, participants of the experiment had the same degree of risk aversion then, their final holdings of assets would be approximately the same (e.g. Lei et al., 2001). However, as dividend value changes in a probabilistic manner from period to period, each stock is perceived as some sort of the lottery by players. More risk-averse participants, who like uncertainty less, would try to sell their assets at the early stages of the experiment. On the contrary, more risk-loving subjects would try to acquire as many asset items as possible. Thus, I expect all stocks to be in the inventory of the more risk-loving participants at the end of experiment.

The above discussion gives justification to test three hypotheses:

*Hypothesis 5:* Overconfident subjects are (more) risk loving, and therefore they make more risky choices in the lottery-choice experiment.

*Hypothesis 6:* Trading activity is negatively dependent on the degree of risk aversion, thus more risk loving subjects engage more in trading activity.

*Hypothesis 7:* The number of assets in the subject's final inventory negatively depends on her degree of risk aversion: the more risk-averse a subject is, the fewer assets she has in her final portfolio.

### 3.4 EXPERIMENTAL PROCEDURE AND THE RULES OF THE GAME

This section first briefly describes the experimental procedure of the asset market experiment<sup>27</sup>, and then it presents the risk aversion experiment.

---

<sup>27</sup> For a detailed description of the experimental procedure see the working paper "Overconfidence and Bubbles in Experimental Asset Markets".

### 3.4.1 ASSET MARKET EXPERIMENT

Data analyzed in this article were obtained from 60 students, enrolled in social sciences at the Christian-Albrechts University of Kiel, who participated in the asset market experimental sessions that were conducted in the academic year 2008-2009. Totally ten experimental sessions were carried out, and six subjects took part in each of them. Thirty five males and 25 females, aged 19 to 28 ( $M = 22.73$ ,  $SD = 2.06$ ) participated in the experimental asset market. On average, one session lasted 1 hour and 40 minutes, and subjects have earned approximately 390.36 units of experimental currency (ECU)<sup>28</sup> ( $SD = 197.89$ ) in the asset market (excluding reward for the forecasting activity). Males, on average, gained significantly more ECUs than women: 447 ECU vs. 335 ECU (Mann-Whitney  $Z = -2.646$ ,  $p < 0.01$ , one-sided). Descriptive statistics of the group, including age, semester, the bias score, and profit are presented in the Appendix 3.A.

Prior to conducting the asset market experiment, subjects' overconfidence was measured in the pre-experimental psychological test sessions. For that, students had to fill-in an 18-questions' general knowledge quiz and state how confident they were that their answers were correct. Based on the difference between subject's average accuracy and her confidence, individual overconfidence was assessed. The obtained measure is called a bias score, where the negative bias score indicated underconfidence, and the positive bias score indicated overconfidence; an equal to zero bias score denoted perfect calibration. For the participation in the experimental asset markets two types of subjects were invited: those who had the lowest bias score (rational subjects) and those who had the highest bias score (overconfident subjects). Of them two types of asset markets were constructed: rational and overconfident. Thus, in the course of the experiment, subjects interacted only with traders of their own type. Yet, within each of the constructed rational and overconfident markets subjects still differ from each other (even slightly) in terms of the individual degree of overconfidence. Appendix 3.B presents characteristics of subjects in the different experimental (sub-) groups in terms of their age, duration of studies and the bias score: all participants of the experimental sessions, and separately overconfident and rational subjects, and male and female participants.

Experiment was programmed and conducted in the computer lab with the software z-Tree (Fischbacher, 2007). Experimental design followed the pioneering work of Smith, Suchanek, and Williams (1988). Every experimental market consisted of the sequence of 15 trading periods, lasting at maximum 180 seconds, during which each trader could post her bid and ask

---

<sup>28</sup> 10.54 EUR.

price of the asset unit. Prior to the start of the experiment each subject was endowed with an equal amount of experimental assets and cash: 300 ECU and 3 experimental asset units. At the end of the trading period, each asset in participants' inventory paid a dividend with possible values of 0.0, 0.8, 2.8, or 6.0 ECU. Probability of each dividend value was 0.25. An average dividend, which subjects could expect through many draws, was 2.4 ECU. Fundamental value of the stock can be obtained according to the formula  $n \times 2.4$  ECU, where  $n$  stands for the number of periods remaining to the end of the session. In the first period of the experiment fundamental asset value was 36 ECU.

At the end of each trading period subjects were asked to predict the average market price in the next period, as well as to state how confident they were that this forecast was correct. Any value between 0% and 100% could be used to express subject's confidence, where 0% meant complete disbelief that the forecast was true, and 100% meant complete certainty that the forecast was correct. Participants were awarded for their predictions based on their accuracy: the closer the prediction was to the actual average market price, the higher was the reward. Gains from the forecasting task were not added to subjects' working capital, but paid to them at the end of the experiment.

At the termination of the experiment, each participant was paid in cash the amount of money, which was based on her final working capital and total gains from the forecasting activity, converted at the predefined exchange rate to Euros.

### 3.4.2 RISK AVERSION EXPERIMENT

A set of experimental sessions aimed at assessment of the individual degree of risk aversion was conducted at the Christian-Albrechts University of Kiel in January, 2010. For each session former participants of the asset market experiment, described in the previous section, were recruited. Thirty two subjects took part in the experimental sessions; 16 of them were overconfident and 16 rational subjects. Nineteen males and 13 females, whose average age was 24.34 years ( $SD = 1.94$ ) and who on average studied for 5.69 semesters ( $SD = 2.05$ ), took part in the experiment. The duration of one experimental session was approximately 20 minutes. The part of the sessions was implemented in the computer lab; another part was paper-pen based. Computer-based experimental sessions were programmed and conducted with the software z-Tree (Fischbacher, 2007). Participants, on average, have earned 5.73 EUR ( $SD = 1.83$ ), including the show-up fee of 2 EUR. Descriptive statistics of subjects in the various experimental (sub-)groups in terms of the number of taken safe choices, age, semester

and the bias score are presented in Appendix 3.C: for the whole group, and separately for rational and overconfident, and male and female subjects.

Experimental design followed the work of Holt and Laury (2002) and Baker, Laury and Williams (2008). Instructions were sent to participants in advance, and subjects were asked to read them prior to arriving to the experimental session. At the beginning of the session students received a copy of instructions, which they were sent via E-mail. The experimenter repeated the most important information concerning the experiment and participants were given time to ask questions. Then, either on their computer screen, or on the paper form, students were presented with a table of ten paired lotteries - *Option A* and *Option B*. Option A represented a “safe” choice in comparison to Option B. Safety of the options was assessed on the basis of the difference between the low payoff and the high payoff outcomes, where the small difference between the two indicated the safe option. The ten decisions had equal payoffs, however, moving down the table, the probability of the high-payoff outcome has gradually increased in steps of 10%, until it has reached 100% for the tenth decision; correspondingly, probability of the low-payoff outcome has gradually decreased in steps of 10%. Lottery payoffs in the experiment were as in the low-payoff treatment of Holt and Laury (2002) multiplied by 1.5. See Table 3.1 for the menu of lottery decisions that were used in the experiment, and the difference in their expected payoffs.

Table 3.1: The ten paired lottery-choice decisions

Option A	Option B	Difference in expected payoffs
1/10 of 3.00 EUR, 9/10 of 2.40 EUR	1/10 of 5.78 EUR, 9/10 of 0.15 EUR	1.75 EUR
2/10 of 3.00 EUR, 8/10 of 2.40 EUR	2/10 of 5.78 EUR, 8/10 of 0.15 EUR	1.24 EUR
3/10 of 3.00 EUR, 7/10 of 2.40 EUR	3/10 of 5.78 EUR, 7/10 of 0.15 EUR	0.74 EUR
4/10 of 3.00 EUR, 6/10 of 2.40 EUR	4/10 of 5.78 EUR, 6/10 of 0.15 EUR	0.24 EUR
5/10 of 3.00 EUR, 5/10 of 2.40 EUR	5/10 of 5.78 EUR, 5/10 of 0.15 EUR	-0.27 EUR
6/10 of 3.00 EUR, 4/10 of 2.40 EUR	6/10 of 5.78 EUR, 4/10 of 0.15 EUR	-0.77 EUR
7/10 of 3.00 EUR, 3/10 of 2.40 EUR	7/10 of 5.78 EUR, 3/10 of 0.15 EUR	-1.27 EUR
8/10 of 3.00 EUR, 2/10 of 2.40 EUR	8/10 of 5.78 EUR, 2/10 of 0.15 EUR	-1.77 EUR
9/10 of 3.00 EUR, 1/10 of 2.40 EUR	9/10 of 5.78 EUR, 1/10 of 0.15 EUR	-2.28 EUR
10/10 of 3.00 EUR, 0/10 of 2.40 EUR	10/10 of 5.78 EUR, 0/10 of 0.15 EUR	-2.78 EUR



Students were asked to make ten decisions and pick one *Option* for each of the lottery pairs. Only one decision was chosen at the end of the experiment to determine participant's earnings. A random number from one to ten was generated for each player individually, either through the random number generator in z-Tree program, or by rolling a ten-sided dice. The obtained number specified the decision that was used to determine subject's payment. Prior to start of the experiment, students were informed that every number had an equal probability to occur, and that they had to think carefully about each of their choices. The outcome of the selected lottery was determined by a second random number from one to ten also generated, either by the computer program, or by rolling a ten-sided dice. At the conclusion of the experiment students were paid their earnings in cash. Additionally, each of them received a show up fee of 2 EUR. English translation of instructions is included in Appendix 3.D.

As in Holt and Laury (2002) and Baker, Laury, and Williams (2008) a total number of safe choices taken by subjects was used to assess their risk aversion. Individuals base their choice of lotteries on the difference in the expected values of the competing lotteries, and on the degree of their own risk aversion. Consequently a risk neutral person, before switching to the risky lottery, would make four safe choices; a risk-averse person would make more than four choices and a risk loving person would make less than four choices. Table 3.2, based on Holt and Laury (2002), classifies degrees of risk aversion, based on the number of taken safe choices in the lottery experiment.

Table 3.2: Risk aversion classification based on the number of safe lottery choices

Number of safe choices	Risk preference classification
0-1	Highly risk loving
2	Very risk loving
3	Risk loving
4	Risk neutral
5	Slightly risk averse
6	Risk averse
7	Very risk averse
8	Highly risk averse
9-10	Stay in bed

## 3.5 EXPERIMENTAL RESULTS

In this section the above-named hypotheses about the consequences of the cognitive bias of overconfidence on individual trading behavior and performance are tested. First, univariate and bivariate analyses are carried out. Upon it, in order to investigate these questions more thoroughly, measures of trading activity and gains from trade are regressed on several explanatory variables.

### 3.5.1 UNIVARIATE AND BIVARIATE ANALYSIS

#### *Trading activity*

Compared to other experiments where subjects had different portfolio endowments and actively used market-place to balance their portfolios, e.g. SSW (1988), in my experiment subjects had equal endowments, thus there was no need for them to use market for portfolios' balancing. Yet the empirical data suggest that average trading activity (MTA), defined as the mean of transactions (purchases and sales) conducted by an individual summed over the session divided by the number of shares outstanding in that market<sup>29</sup>, was quite high (see Appendix 3.A). On average, per session, traders have transacted 0.89 times the outstanding stock of shares; trading activity of some subjects equaled several times the stock. To test the hypothesis, that overconfident investors trade more, I calculate the correlation coefficient between individual average trading activity and overconfidence, expressed as the bias score. Results imply that individuals with the higher bias score engage in trading activity more actively (Pearson correlation(58) = 0.350,  $p < 0.01$ , one-sided; medium correlation). Traders in overconfident markets engage in trading activity significantly more often than traders in rational markets (Mann-Whitney  $Z = -2.610$ ,  $p < 0.01$ , one-sided). There is no significant difference in transactions' frequency between female and male participants (Mann-Whitney  $Z = -0.105$ ,  $p = 0.916$ , two-sided). To further analyze gender dimension in the relationship between trading activity and bias score, correlation coefficients were recalculated separately for female and male participants. No significant linear relationship between overconfidence and trading activity was found for males (Pearson correlation(33) = 0.118,  $p = 0.249$ , one-sided). Yet, for females the correlation coefficient is high and significant (Pearson correlation(23) = 0.635,  $p = 0.000$ , one-sided), implying that female-participants engage in trading activity more lively with increase in their bias score.

---

<sup>29</sup> 18 shares.

Now I turn to testing the proposition that high turnover traders are outperformed by low turnover traders. Normalized profits of the participants, calculated as individual gains from trade scaled by the initial portfolio value ( $36 \text{ ECU} \times 3 = 106 \text{ ECU}$ ), and corresponding to them average trading activity are presented on Figure 3.1. Data in the graph are arranged in the increasing order of trading activity values and are plotted with the linear trend line. The trend line implies that increase in trading activity is accompanied by decrease in individual gains. Average normalized profits equaled 3.614 times the value of the initial portfolio ( $SD = 1.83$ ). In the total mass of individual earnings two values, namely 7.915 and 8.391 (in absolute value consequently 854.8 ECU and 906.2 ECU) look rather like exceptions. This observation will be taken in account, while conducting statistical tests.

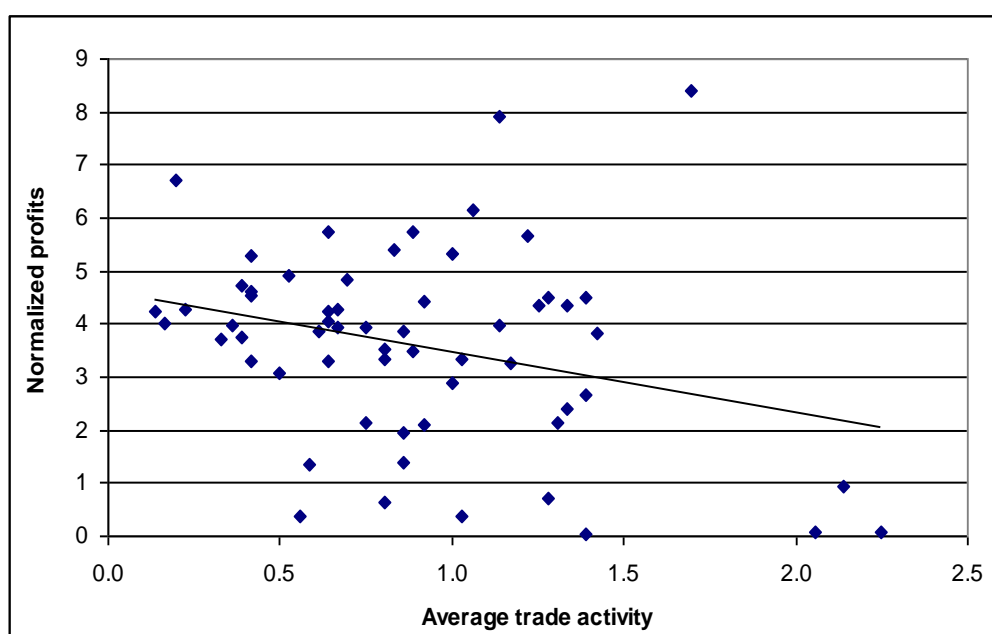


Figure 3.1: Normalized profits per participant

To verify the relationship between trading activity and gains from trade the sample is broken in five equal sub-samples ranked in terms of trading activity (quintiles). In Figure 3.2 distribution of the average total profits from trade across trading activity quintiles is presented. One can notice the pattern of reduction in earnings with growth of trading activity (with an exception of the fourth quintile). Can the hypothesis that gains are equal in turnover quintiles be rejected? The conducted Kruskal-Wallis test cannot reject this hypothesis ( $Kruskal-Wallis(4) = 7.562, p = 0.109$ , two-sided). However, as mentioned above, there are two possible outliers in the sample of individual performance, thus the analysis is repeated without these observations. For the reduced sample, proposition of gains' equality across quintiles is rejected ( $Kruskal-Wallis(4) = 9.351, p = 0.053$ , two-sided). This result serves as weak evidence that gains from trade depend on trading activity. To get more clarity on this

issue, correlation coefficient between trading activity and individual earnings is estimated. The obtained coefficient is small but significant, implying that increase in trading activity is paired with poorer performance (Pearson Correlation(58) = -0.292,  $p < 0.05$ , one-sided). If analysis is repeated without the two outlier values, strength of the linear relationship increases to medium (Pearson Correlation(56) = -0.456,  $p = 0.00$ , one-sided), which suggests that increase in trading activity is moderately associated with reduction in gains from trade.

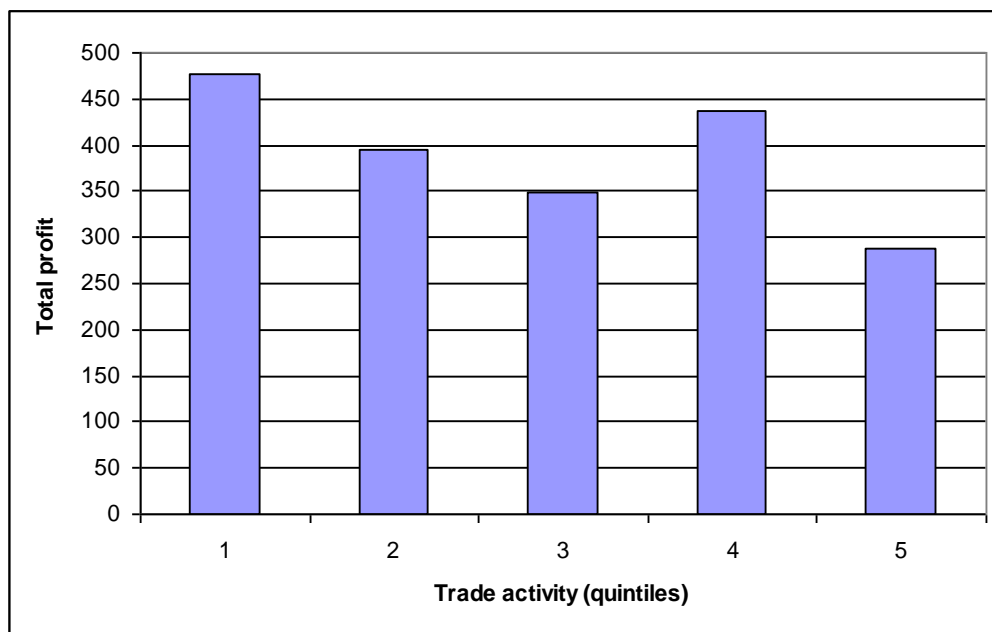


Figure 3.2: Distribution of individual profits across trading activity quintiles

As a next step, comparison between earnings in the subgroup, whose trading activity was lower than the median value (0.847), and subgroup, whose trading activity was higher than the median, was conducted. Although, on average the latter made 15.17% less in terms of individual earnings, a statistically significant conclusion that this group profited less from trade cannot be drawn (Mann-Whitney  $Z = -1.271$ ,  $p = 0.102$ , one-sided). Yet, when analysis is repeated with the reduced sample (without the possible outliers)<sup>30</sup>, results suggest that the *low trade* group outperformed on average the *high trade* group by 23.12%, and this difference is significant (Mann-Whitney  $Z = -1.438$ ,  $p < 0.10$ , one-sided). To explore this issue further, individual gains in the lowest and the highest trading activity quartiles were compared (for descriptive statistics see Table 3.3(a)). The conducted Mann-Whitney test revealed that traders in the lowest quartile significantly outperformed traders in the highest quartile ( $Z = -1.555$ ,  $p < 0.10$ , one-sided) and gained on average 38% more ECUs at the end of the experiment. Without the outliers this difference is even higher – 55.7% (Mann-Whitney  $Z = -2.095$ ,  $p <$

<sup>30</sup> Median equals 0.819.

0.05, one-sided; see Table 3.3(b). This is in line with the results of Barber and Odean (2000) who showed that high turnover households were outperformed by low turnover households.

Table 3.3: Profit and trading activity in the lowest and highest trading activity quartiles

a: The whole sample

	1 <sup>st</sup> quartile		4 <sup>th</sup> quartile	
	Profit	Trade	Profit	Trade
Mean	442.69	0.363	321.13	1.515
Median	456.6	0.389	286.6	1.389
SD	145.73	0.129	260.5	0.347
Min.	42.4	0.14	4.2	1.22
Max.	726	0.56	906.2	2.25
N	15	15	15	15

b: Sample without outliers

	1 <sup>st</sup> quartile		4 <sup>th</sup> quartile	
	Profit	Trade	Profit	Trade
Mean	442.69	0.363	284.33	1.48
Median	456.6	0.389	286.6	1.333
SD	145.73	0.129	205.02	0.354
Min.	42.4	0.14	4.2	1.17
Max.	726	0.56	612.6	2.25
N	15	15	15	15

*Gains from Trade*

This subsection is devoted to the analysis of traders' performance in the experimental asset market. Subject's performance is assessed as her relative profit, which is calculated based on Hirota and Sunder (2007), as individual gains from trade (in ECUs) divided by the fundamental value of the initial portfolio of 3 stocks ( $36 \text{ ECU} \times 3 = 106 \text{ ECU}$ ) minus the cross-sectional average of this ratio. Figure 3.3 shows the cross-sectional distribution of

players' relative profits in the two types of experimental market. The value of each marker represents one trader's relative profit. Standard deviation of individual earnings from the mean value (0) in the overconfident market sessions is 2.32 and in the rational market sessions – 1.19. Empirical evidence suggests that overconfident sessions were characterized by the larger dispersion of gains from trade, compared to the rational sessions. The conducted Siegel-Tukey test confirmed that difference in profits' variation within the two groups was statistically significant (Siegel-Tukey = 2.329,  $p < 0.05$ , two-sided).

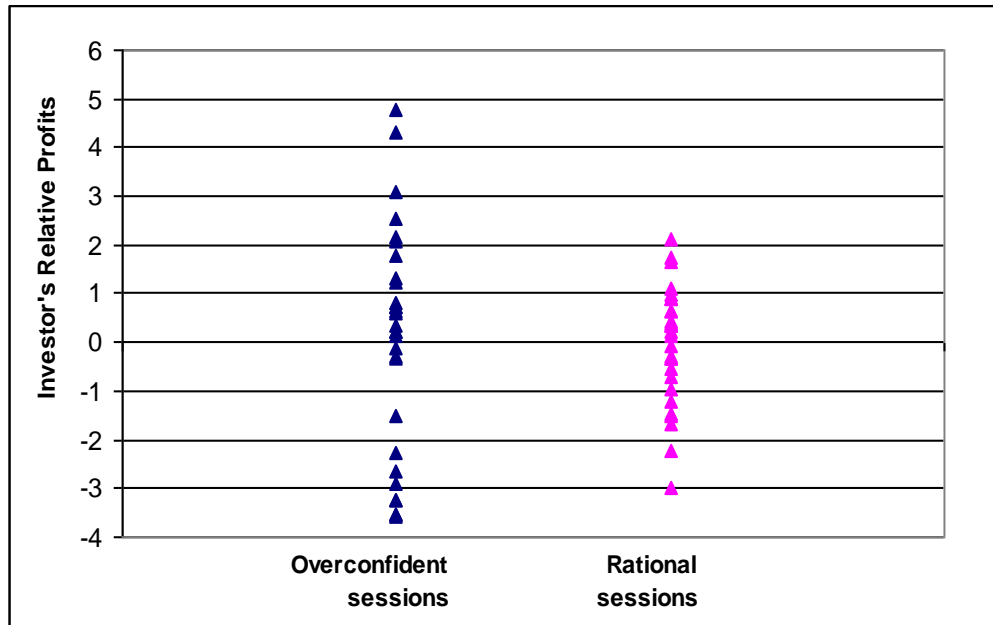


Figure 3.3: The cross-sectional distribution of relative profits by treatment

In the experiment by SSW (1988) better forecasters have also enjoyed higher gains from trade in the experimental market. To determine the relationship between accuracy of average price prediction and individual earnings and its' direction in my experiment several statistical tests are performed. Forecasting precision is expressed as the Total Absolute Error (TAE) of prediction and/ or the Average Absolute Error (MAE) (see Appendix 3.A for sample values of MAE and TAE):

$$Total\ Absolute\ Error\ (TAE)_i = |Sum\ (P_t - F_{it})| = |Sum\ (P_t) - Sum\ (F_{it})| \quad (3.1)$$

$$Average\ Absolute\ Error\ (MAE)_i = Sum\ |(P_t - F_{it})| / 15 \quad (3.2)$$

Here,  $F_{it}$  is the forecast of subject  $i$  for the period  $t$ , and  $P_t$  is the observed average price in period  $t$ .

The calculated correlation coefficient between forecasting precision and relative profits detected moderate and significant negative correlation (MAE: Pearson Correlation (58)= -

0.360,  $p < 0.01$ , one-sided; TAE: Pearson Correlation(58) = -0.365,  $p < 0.01$ , one-sided). This is in line with previous research which suggests that increase in price prediction precision is paired with higher gains from trade in the experimental asset markets. I continue by comparing traders' accuracy in both of the groups. Subjects in the rational group committed significantly less errors, measured both as MAE and TAE (MAE: Mann-Whitney  $Z = -3.512$ ,  $p < 0.00$ , one-sided; TAE: Mann-Whitney  $Z = -2.610$ ,  $p < 0.01$ , one-sided). Linear relationship between overconfidence and average absolute error (MAE) is moderate and significant (Pearson Correlation(58) = 0.350,  $p < 0.01$ , one-sided), and between overconfidence and total absolute error (TAE) is small but significant (Pearson Correlation(58) = 0.225,  $p < 0.05$ , one-sided). Thus, it can be concluded that increase in the degree of overconfidence is accompanied by (moderate) reduction in accuracy of prediction.

Another factor that had negative impact on earnings was the number of assets in participants' inventory at the end of experiment. Namely there was a weak tendency for subjects with more assets in their final portfolios to have lower gains. Correlation between the final asset-holdings and relative profit is small but significant (Pearson Correlation(58) = -0.225,  $p < 0.05$ , one-sided). An interesting fact is that, although female subjects had fewer assets in their inventories at the end of experiment (see below), their gains from trade were significantly lower than gains of male subjects (Mann-Whitney  $Z = -2.646$ ,  $p < 0.01$ , one-sided).

To test the proposition of Scheinkman and Xiong (2003) that more overconfident investors would hold most of the assets at the end of experiment, I first estimate the correlation between individual bias scores and the number of assets in players' final inventory. The correlation coefficient is almost equal to zero and insignificant (Pearson Correlation (28) = -0.031,  $p = 0.407$ , two-sided), thus providing no evidence of linear relationship between overconfidence and final portfolio size. The same is true for the rational and overconfident markets separately (overconfident: Pearson Correlation (28) = -0.097,  $p = 0.306$ , one-sided; rational: Pearson Correlation (28) = 0.028,  $p = 0.442$ , one-sided). To test for the differences in dispersion (scale) between end inventories of the two types of traders, a Siegel-Tukey variance test was conducted, which revealed that rational and overconfident traders did not differ in terms of final portfolios' variance (Siegel-Tukey = 0.531,  $p = 0.595$ , two-sided). See Appendix 3.E for distribution of the number of assets in the final inventories of all traders (a.), overconfident traders (b.), and rational traders (c.). There was significant difference in the final portfolio size of male and female players, where the latter had significantly less assets than the former at the end of experiment (Mann-Whitney  $Z = -3.121$ ,  $p < 0.01$ , one-sided). No linear relationship between asset-holdings and the bias scores in none of the groups was detected (women:

Pearson Correlation(23) = 0.105, p = 0.309, one-sided; men: Pearson Correlation(33) = -0.126, p = 0.235, one-sided).

### 3.5.2 MULTIVARIATE ANALYSIS

#### *Trading activity*

This subsection presents results of cross-sectional regressions estimating the relationship between average trading activity of an individual (MTA) per experiment and several explanatory variables that might affect efficiency of financial decision making: the normalized bias score<sup>31</sup> (NBS), gender dummy (this variable takes value 1 if subject is male), an interaction term between the bias score and gender (NBS\*Gender), subject's experience expressed as age (Age) or duration of studies in semesters (Semester), and price forecasting precision measured as average absolute error (MAE) or total absolute error (TAE). In parenthesis error terms are shown. Equations 8 and 9, instead of average trading activity use the measure of total trading activity. All specifications of equations (see Table 3.4) confirm the proposition that overconfidence has explanatory power for trading activity<sup>32</sup>.

To detect which variables significantly affect the number of average stock market transactions per person in my experiment, I start with the simplest model specification in which average trading activity (MTA) is regressed on the individual degree of overconfidence, measured as the normalized bias score (NBS). Subsequently a range of alternative specifications are estimated by adding other regressors to the model. From Equation 1 it can be seen that, as predicted, NBS has significant positive influence on trading activity: a one standard deviation increase in the bias score raises trading activity by 0.35 standard deviations<sup>33</sup> (p < 0.05, two-sided).

---

<sup>31</sup> A sample of bias scores of the participants is normalized on an interval [0,1], where one stands for the most overconfident person.

<sup>32</sup> Equations 10 and 11 are discussed later in the section on "Risk Aversion Analysis: Experimental Results".

<sup>33</sup> In all specifications, one of the key variables of our interest, namely the normalized bias score, is measured on a scale that is not easy to interpret. Thus to assess the effect that independent variables had on a dependent variable standardized  $\beta_{std}$  coefficients were calculated, which allowed expressing this effect not in terms of the original units of the variables, but in standard deviation units, i.e. to see how the dependent variable changed, if the independent variable grew by one standard deviation. Standardized beta coefficients were obtained by multiplying the original beta coefficient by the ratio of the sample standard deviation of the corresponding independent variable to the sample standard deviation of the dependent variable:

$$\hat{\beta}_{std} = (\hat{\sigma}_j / \hat{\sigma}_y) \hat{\beta}_j \quad \text{for } j = 1, \dots, k.$$



Table 3.4: Trading activity (all errors are heteroskedasticity corrected)

	1	2	3	4	5	6	6a	7	8	9	10	11
C	0.660**** (0.085)	0.693**** (0.086)	0.499**** (0.102)	1.340** (0.623)	0.734**** (0.117)	1.220* (0.618)	1.301 (0.618)	1.302 (0.637)	1.318**** (0.171)	2.431* (1.239)	1.815** (0.869)	2.111** (0.836)
NBS	0.669** (0.266)	0.680** (0.268)	1.286*** (0.370)	1.237*** (0.377)	1.149*** (0.371)	0.994** (0.384)	1.052*** (0.385)	1.167*** (0.364)	1.339** (0.530)	1.985** (0.768)	0.859* (0.774)	0.854* (0.757)
Gender		-0.063 (0.119)	0.301* (0.158)	0.253*, a (0.160)	0.253*, a (0.157)	0.317**, a (0.163)	0.268**, a (0.154)	0.279* (0.160)		0.635**, a (0.326)	0.470 (0.269)	0.446 (0.243)
NBS*Gender			-1.074** (0.460)	-0.996** (0.481)	-0.998** (0.475)	-0.939**, a (0.476)	-0.956** (0.475)	-1.008** (0.480)		-1.871**, a (0.950)	-1.307*, a (0.878)	-1.329*, a (0.857)
Age				-0.036*, a (0.026)		-0.040*, a (0.025)	-0.041* (0.026)	-0.038*, a (0.027)		-0.081*, a (0.050)	-0.060**, a (0.033)	-0.067** (0.029)
Semester					-0.046** (0.019)							
MAE						0.036** (0.015)	0.027**, a (0.012)			0.072** (0.031)	0.040** (0.017)	
TAE								0.002*, a (0.001)				0.004** (0.002)
Safe choices											-0.023 (0.026)	-0.033 (0.026)
N	60	60	60	60	54	60	59	60	60	60	32	32
adj. R2	0.11	0.10	0.16	0.17	0.17	0.23	0.25	0.19	0.11	0.23	0.15	0.23
SERegr	0.44	0.44	0.43	0.42	0.43	0.41	0.38	0.42	0.88	0.82	0.40	0.38

\*\*\*\* 0.001; \*\*\* 0.01; \*\* 0.05; \* 0.1. <sup>a</sup> one-sided test.

In Equation 2 it is also controlled for the role of traders' gender on variation in trading activity, and it is detected that gender has no significant influence on trading activity. Impact of overconfidence remains significant and positive: a one standard deviation increase in NBS raises trading activity by 0.356 standard deviations ( $p < 0.05$ , two-sided).

In Equation 3, an interaction term between overconfidence and gender is added; this increases the explained variation in trading activity. The regression results imply that, the strength of the impact of overconfidence on trade intensity varies across genders, i.e. with increase in overconfidence the rate of increment in trading activity is lower for males, than for females. For every unit increase in overconfidence when gender dummy is equal to one (i.e. a trader is a man) trading activity grows by only 0.212 units, whereas for a female subject this increase equals to 1.286 units (interaction term is significant at 5%, two-sided). This contradicts the result of Barber and Odean (2001) that men traded more than women. Yet, Deaves et al. (2009) found no interaction between gender and overconfidence in their sample. They noted that, after controlling for overconfidence, there was no difference in trading activity between genders. Also in the paper of Glaser and Weber (2007) gender was not significantly related to the trading volume measures.

In addition to previous regressors, in Equation 4 it is tested for the explanatory power of subjects' experience, namely age, for trading activity. Results indicate that a one standard deviation increment in age decreases trading activity by -0.162 standard deviations ( $p < 0.1$ , one-sided). Overconfidence and gender interaction term is significant at 5% level (two-sided) and for every unit increase in NBS for male subjects there is growth in trading activity of 0.241 units; in contrast, for female subjects increase of 1.237 units in trading activity is observed.

In Equation 5, I control for another subjects' experience proxy, such as duration of studies in semesters. Negative impact of subjects' experience on the number of average market transactions persists: a one standard deviation increase in semester variable decreases trading activity by -0.207 standard deviations ( $p < 0.05$ , two-sided). Interaction term between gender and overconfidence is negative and significant at 5% level (two-sided), and for every unit increase in NBS for male subjects there is growth in trading activity of 0.150 units; in contrast, for female subjects increase of 1.340 units in trading activity is observed.

To test if forecasting errors boost trading activity, in Equation 6 it is controlled for individual average absolute error (MAE) in forecasting. Analysis reveals that the more an individual is mistaken about future asset prices, the more actively she engages in trading activity: a one standard deviation increase in MAE raises trading activity by 0.291 standard deviations ( $p < 0.05$ , two-sided). Other significant control variables are interaction term between participants'

overconfidence and gender, and age: a one standard deviation increment in age reduces trading activity by -0.184 standard deviations ( $p < 0.1$ , one-sided); interaction term is negative and significant at 5% level (one-sided), and for every unit increase in NBS for male subjects there is growth in trading activity of 0.055 units; in contrast, for female subjects increase of 0.994 units in trading activity is observed.

After analysis of the residuals, one possible outlier was detected and, after exclusion of it, the regression was re-run. Equation 6a indicates that increase in forecasting error by one standard deviation increases trading activity by 0.231 standard deviations ( $p < 0.1$ , one-sided). Increase in age by one standard deviation reduces trading activity by -0.201 standard deviations ( $p < 0.1$ , two-sided). Interaction term between gender and overconfidence is negative and significant at 5% level (two-sided), and for every unit increase in NBS for male subjects there is growth in trading activity of 0.095 units; in contrast, for female subjects increase of 1.052 units in trading activity is observed. Exclusion of the outlier resulted in increase in the explained variation in trading activity and reduction of the regression error, implying that model without the outlier better represents the data.

In Equation 7 forecasting imprecision is assessed as total absolute error (TAE). The goodness of fit of the regression model deteriorates and regression error increases; the direction of the relationship between the variables remains as in Equation 6: a one standard deviation increment in age reduces trading activity by -0.173 standard deviations ( $p < 0.1$ , one-sided); a one standard deviation increase in forecasting error, measured as TAE, raises trading activity by 0.168 standard deviations ( $p < 0.1$ , one-sided); interaction term is negative and significant at 5% (two-sided), and for every unit increase in NBS for male subjects there is growth in trading activity of 0.159 units; in contrast, for female subjects increase of 1.167 units in trading activity is observed.

In Equations 8 and 9 instead of average trading activity another dependent variable namely total trading activity (TTA) is used. Total trading activity is defined as the sum of all trade contracts (purchases and sales) per fifteen periods divided by the number of shares outstanding in that market (i.e. 18 shares in each of the markets). Replacement of the dependent variable does not bring about any substantial changes in relationships between regressors and the dependent variable. The goodness of fit of the models does not improve, yet the error term doubles.

Results, presented in Table 3.4, lead to the following conclusions 1) subject's experience has a significant negative effect on her engagement in trading activity, 2) impact of overconfidence on trade is positive, however, holding all other factors constant, with increase

in overconfidence men engage in stock market transactions less than women, 3) forecasting errors, that induce false future price expectations, force subjects to involve in trading activity more actively. Modest success in explaining variation in trading activity in the sample by means of selected models suggests that other unobserved factors that were not included in the regression, also have impact on the average number of market transactions by an individual participant. I will come back to this issue in the section on risk aversion analysis, where the regression model is re-estimated for a sample of participants whose risk aversion measures was obtained.

### *Gains from Trade*

Many theoretical and empirical works predict reduction in welfare, which is faced by overconfident traders. In this subsection I describe the results of the cross-sectional regressions, estimating the relationship between subjects' performance, assessed as relative profit, and several explanatory variables: the normalized bias score (NBS), gender dummy (this variable takes value 1 if subject is male), average trading activity (Trading activity), an interaction term between gender and trading activity (Gender\*Trading activity), price forecasting precision measured as average absolute forecasting error (MAE) or total absolute forecasting error (TAE), subject's experience expressed as age (Age) or duration of studies in semesters (Semester), and the number of assets in the final inventory (End assets). In parenthesis the error terms are shown. For the specifications of the estimated models see Table 3.5.

I start again with the simplest model specification and, by adding other regressors to the model, test which variables significantly affect relative profit from trade of participants in my experiment. In Equation 1 it is tested for the influence of subject's degree of overconfidence and gender on gains from trade. The normalized bias score has no significant effect on performance, while gender has significant impact on individual profit: male students enjoy higher earnings in comparison to their female counterparts ( $p < 0.01$ , two-sided).

In addition to previous regressors in Equation 2 it is controlled for the impact of active trade engagement on variations in relative profit. From the results, one can see that this influence is negative and significant: a one standard deviation increase in trading activity reduces performance by -0.325 standard deviations ( $p < 0.05$ , two-sided). The effect of overconfidence on gains from trade is positive, yet insignificant. Males have earned significantly more ECUs compared to females ( $p < 0.01$ , two-sided).

To analyze, whether strength of the effect that active trade engagement has on performance varies across female and male participants, an interaction term between gender and trading

activity is introduced in Equation 3. Adding this regressor increases the explained variation in relative profits. Results indicate that with increase in the average number of market transactions, the rate of decrease in gains varies across genders, and is lower for male subjects. For every unit increase in trading activity for male traders there is reduction in earnings by -0.206 units, whereas for female traders earnings decrease by -2.655 units (interaction term is significant at 5%, two-sided). Overconfidence is a factor that has significant positive effect on earnings: for a one standard deviation increase in NBS there is growth in relative profit by 0.210 standard deviations ( $p < 0.1$ , one-sided).

In Equation 4, impact of forecasting errors on relative profit is analyzed. As in the work by SSW (1988), forecasting errors have negative and significant consequences for gains from trade: a one standard deviation increase in average absolute error (MAE) reduces relative profit by -0.221 standard deviations ( $p < 0.1$ , one-sided). Subjects' relative profits significantly increase with the increment in the degree of overconfidence ( $p < 0.1$ , two-sided). For every unit increase in trading activity for male traders there is reduction in earnings by -0.074 units, whereas for female subjects earnings decrease by -2.314 units (interaction term is significant at 5%, two-sided).

In Equation 5 forecasting imprecision is assessed as total absolute error (TAE). TAE has negative effect on relative profit, which is comparable to the effect obtained in Equation 4; however there is increase in the goodness of fit of the model and decrease in the error term. A one standard deviation increment in TAE reduces individual earnings by -0.277 standard deviations ( $p < 0.05$ , one-sided). Growth in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains by 0.253 standard deviations ( $p < 0.1$ , two-sided). For every unit increase in trading activity for male traders there is reduction in earnings by -0.130 units, whereas for female subjects earnings decrease by -2.360 units (interaction term is significant at 5%, two-sided).

In Equation 6 one more regressor is added, and it is tested for the explanatory power of subjects' experience, proxied by age, for relative profit. From the estimation outcome one can see that age has no significant effect on individual earnings. A one standard deviation increment in forecasting error, measured as TAE, reduces relative profit by -0.293 standard deviations ( $p < 0.05$ , one-sided). Growth in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains by 0.241 standard deviations ( $p < 0.1$ , two-sided). For every unit increase in trading activity for male traders there is reduction in earnings by -0.112 units, whereas for female subjects earnings decrease by -2.124 units (interaction term is significant at 5%, two-sided).

Table 3.5: Gains from trade (all errors are heteroskedasticity corrected)

	1	2	3	4	5	6	7	8	9	10	11
C	-0.776* (0.423)	0.108 (0.539)	1.159*** (0.377)	1.626*** (0.498)	1.525*** (0.442)	-1.030 (2.881)	-0.029 (2.702)	-0.035 (2.700)	1.790** (0.788)	1.465**** (0.417)	1.624**** (0.399)
NBS	0.065 (1.085)	0.932 (1.037)	1.577*, a (1.045)	1.995* (1.069)	1.896* (0.987)	1.805* (0.957)	1.440*, a (0.919)	1.436*, a (0.919)	1.549*, a (1.013)	1.474*, a (0.936)	1.628* (0.899)
Gender	1.293*** (0.442)	1.212*** (0.395)	-1.012 (0.812)	-1.064* (0.813)	-0.950 (0.770)	-0.692 (0.760)	-0.431 (0.740)	-0.429 (0.740)	-0.695 (0.820)	-0.564 (0.738)	-0.577 (0.746)
Trading activity (average)		-1.276** (0.520)	-2.655**** (0.488)	-2.314**** (0.568)	-2.360**** (0.564)	-2.124*** (0.661)	-2.122**** (0.559)		-2.366**** (0.567)	-2.255**** (0.469)	-2.402**** (0.446)
Trading activity*Gender			2.449** (0.996)	2.240** (0.986)	2.229** (0.977)	2.013** (0.977)	2.237** (0.978)	1.120** (0.489)	2.362** (1.023)	2.370** (0.969)	2.555** (0.947)
MAE				-0.108*, a (0.069)							
TAE					-0.012**, a (0.006)	-0.013**, a (0.007)	-0.008*, a (0.006)	-0.008*, a (0.006)	-0.007 (0.006)	-0.008*, a (0.005)	-0.011*** (0.004)
Age						0.104 (0.118)	0.061 (0.109)	0.061 (0.109)			
End assets							-0.136**, a (0.073)	-0.136**, a (0.073)	-0.118*, a (0.085)	-0.143** (0.071)	-0.153** (0.069)
Semester									-0.067 (0.105)		
Trading activity (total)								-1.061**** (0.280)			
N	60	60	60	60	60	60	60	60	54	60	59
adj. R2	0.09	0.17	0.26	0.28	0.32	0.32	0.38	0.38	0.32	0.39	0.47
SERegr	1.75	1.67	1.58	1.55	1.51	1.51	1.45	1.45	1.49	1.44	1.34

\*\*\*\* 0.001; \*\*\* 0.01; \*\* 0.05; \* 0.1. <sup>a</sup> one-sided test.

Equation 7 analyzes the effect of the final portfolio size on gains from trade. The number of assets in the final inventory is a significant determinant of reduction in relative profit: a one standard deviation increase in final portfolio size decreased earnings by -0.301 standard deviations ( $p < 0.1$ , one-sided). A one standard deviation increment in forecasting error, measured as TAE, reduces relative profit by -0.189 standard deviations ( $p < 0.05$ , one-sided). Growth in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains from trade by 0.192 standard deviations ( $p < 0.1$ , one-sided). For every unit increase in trading activity for male traders there is growth in earnings by 0.115 units, whereas for female subjects earnings decrease by -2.366 units (interaction term is significant at 5%, two-sided).

In Equation 8, instead of average trading activity another dependent variable, namely total trading activity (TTA), is used. This specification, compared to Equation 7, results in no changes to the goodness of fit of the model and standard error of regression. A one standard deviation increment in forecasting error, measured as TAE, reduces earnings by -0.189 standard deviations ( $p < 0.1$ , one-sided). Subjects' relative profit deteriorates with growth in the size of their final portfolio: a one standard deviation increase in the end inventory reduces individual earnings by -0.301 standard deviations ( $p < 0.05$ , one-sided). Increment in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains from trade by 0.192 standard deviations ( $p < 0.1$ , one-sided). For every unit increase in trading activity for male traders there is growth in earnings by 0.059 units, whereas for female subjects earnings decrease by -2.122 units (interaction term is significant at 5%, two-sided).

In Equation 9, I control for another subjects' experience proxy, such as duration of studies in semesters. As in the case with age variable, this relationship is also insignificant. Forecasting error, measured TAE, has no significant effect on gains from trade. Subjects' relative profit deteriorates with growth in the size of their final portfolio: a one standard deviation increase in the end inventory reduces individual earnings by -0.260 standard deviations ( $p < 0.1$ , one-sided). Increment in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains from trade by 0.207 standard deviations ( $p < 0.1$ , one-sided). For every unit increase in trading activity for male traders there is reduction in earnings by -0.004 units, whereas for female subjects earnings decrease by -2.122 units (interaction term is significant at 5%, two-sided).

The insignificant experience variables (age and semester) are removed from Equation 10. This specification, compared to Equations 7 and 9, is characterized by better fit of the regression

model and reduction of the regression error. A one standard deviation increment in forecasting error, measured as TAE, decreases earnings by -0.175 standard deviations ( $p < 0.1$ , one-sided). Subjects' gains deteriorate with growth in the size of their final portfolio: a one standard deviation increase in the end inventory reduces individual earnings by -0.315 standard deviations ( $p < 0.05$ , one-sided). Increment in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains from trade by 0.197 standard deviations ( $p < 0.1$ , one-sided). For every unit increase in trading activity for male traders there is growth in earnings by 0.115 units, whereas for female subjects earnings decrease by -2.255 units (interaction term is significant at 5%, two-sided).

After analysis of the residuals, one possible outlier was detected and, after exclusion of it, the regression was re-run. Equation 11 indicates that a one standard deviation increment in forecasting error, measured as TAE, decreases earnings by -0.249 standard deviations ( $p < 0.01$ , two-sided). Subjects' relative profit deteriorates with growth in the size of their final portfolio: a one standard deviation increase in the end inventory reduces individual earnings by -0.335 standard deviations ( $p < 0.05$ , two-sided). Increment in the degree of overconfidence has positive and significant effect on performance: a one standard deviation increase in NBS raises gains from trade by 0.218 standard deviations ( $p < 0.1$ , two-sided). For every unit increase in trading activity for male traders there is growth in earnings by 0.153 units, whereas for female subjects earnings decrease by -2.402 units (interaction term is significant at 1%, two-sided). Exclusion of the outlier resulted in increase in the explained variation in relative profit and reduction of the regression error, implying that model without the outlier better represents the data.

Results presented in Table 3.5 suggest that 1) overconfidence degree has a significant positive effect on individual earnings, 2) impact of active trade engagement on relative profit is negative, however, holding all other factors constant, with increase in the number of market transactions males incur smaller losses, or even some yield, in comparison to females, 3) forecasting errors, that induce false future price expectations and "cause mistakes in financial decision making" (Biais et al., 2005), produce losses, 4) the number of assets in the final inventory of the subject is a significant determinant of reduction in gains from trade. In general, the described specifications succeeded quite well in explaining variation in relative profits in the sample. Yet, the amount of unexplained variation suggests that other unobserved factors that were not included in the regression also were at play.



### 3.6 RISK AVERSION ANALYSIS: EXPERIMENTAL RESULTS

Experimental results present evidence that on average subjects were risk averse with 5.66 taken safe choices (SD = 1.82). 28.13% of the group has made six safe choices, 25% has made five safe choices, and 18.75% has made seven safe choices. As Table 3.2 suggests, risk preferences of the participants can be classified correspondingly as risk averse, slightly risk averse and very risk averse. **In general**, 71.88% of choices have fallen in the interval of [5, 7] safe options (see Figure 3.4). Rational subjects have taken on average 5.81 safe choices (SD = 1.42), and overconfident subjects 5.50 safe choices (SD = 2.19). Subjects' demographics (age, semester, gender) had no significant impact on their risk aversion. No linear relationship between participants' age, their duration of studies (semester), and the number of safe choices was found (age: Spearman's  $Rho(30) = 0.091$ ,  $p = 0.310$ , one-sided; semester: Spearman's  $Rho(26) = -0.159$ ,  $p = 0.205$ , one-sided). Men have taken on average 5.68 safe choices (SD = 2.00), and women 5.62 safe choices (SD = 1.61). Difference in the average number of safe choices taken by both genders is insignificant (Mann-Whitney  $Z = -0.452$ ,  $p = 0.652$ , two-sided). In Appendix 3.C descriptive statistics of the experiment are presented: the number of safe choices taken by participants, their age, semester and the bias score.

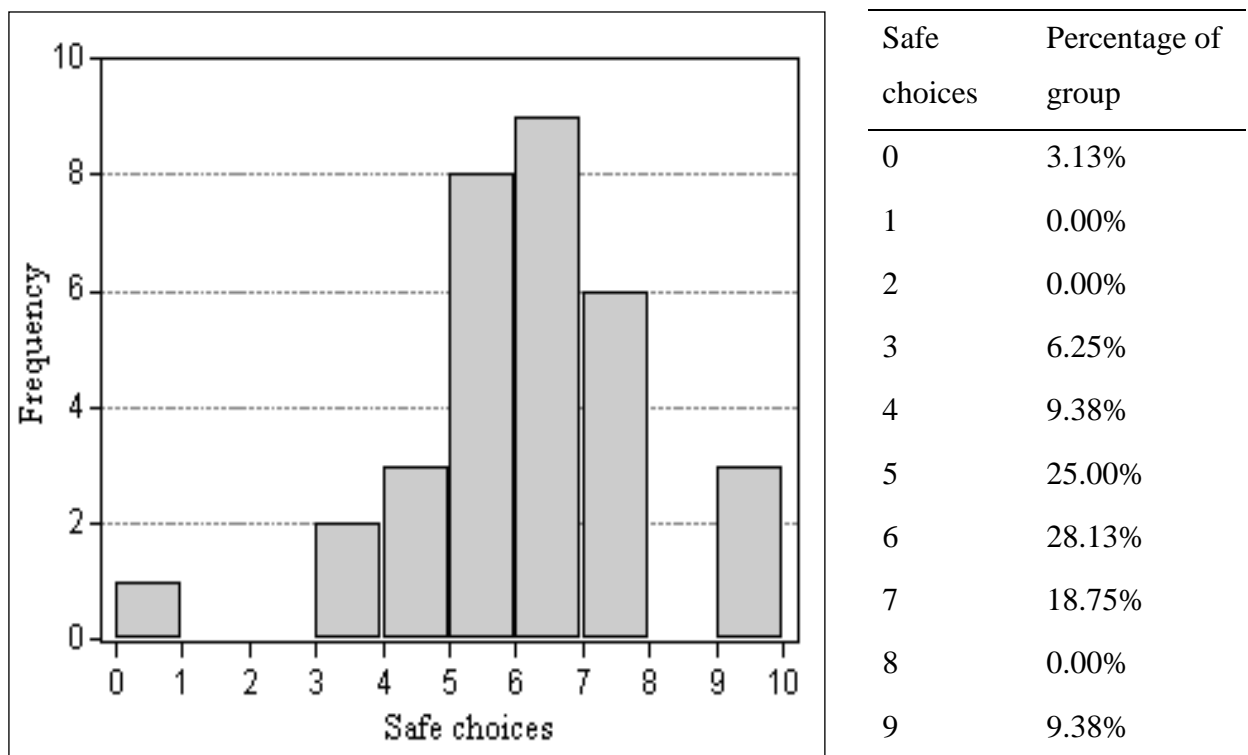


Figure 3.4: Distribution of safe choices in the group

Now, I turn to the examination of the relationship between overconfidence and risk aversion. It was hypothesized that overconfident subjects would be more risk loving, i.e. they would make less safe choices. Statistical tests detected no significant difference between two groups of players, neither in terms of the average number of safe choices (Mann-Whitney  $Z = 0.320$ ,  $p = 0.749$ , two-sided), nor in their variation (Siegel Tukey test = 0.465,  $p = 0.642$ , two-sided). Correlation coefficient between risk aversion, measured as the number of safe choices, and individual bias score implies no linear relationship between them (Spearman's  $Rho(30) = -0.095$ ,  $p = 0.303$ , one-sided). Additionally two regressions were run, where subjects' degree of overconfidence was regressed on the number of safe choices they made, their gender, and either their age (Equation 1) or duration of their studies in semesters (Equation 2) (see Appendix 3.F). In Equation 1, no significant regressors were identified for predicting individual overconfidence; the goodness of fit measure indicates very poor fit. From inspection of Equation 2 it can be seen that a one standard deviation increase in duration of studies reduces overconfidence by -0.0064 standard deviations ( $p < 0.1$ , two-sided), however this impact is almost negligible. Other regressors are insignificant. The presented evidence implies that, in this sample, the number of safe choices has no explanatory power for subject's overconfidence.

The next step is to analyze the relationship between risk aversion and some experimental outcomes: the number of assets in the final inventory of the subject, intensity of the engagement in trading activity and gains from trade. It has been predicted that greater degree of subject's risk aversion will have stronger negative effect on her trading activity and the number of assets in her final inventory. This prediction is based on the assumption that risk-averse subjects would perceive each stock as a lottery and try to sell them in the initial periods of the game. No linear relationship was detected, neither between the final portfolio size and risk aversion (Spearman's  $Rho(30) = -0.001$ ,  $p = 0.498$ , one-sided), nor between gains from trade and risk aversion (Spearman's  $Rho(30) = 0.031$ ,  $p = 0.433$ , one-sided). The correlation coefficient between the number of safe choices and active trade engagement was negative, yet insignificant (Spearman's  $Rho(30) = -0.100$ ,  $p = 0.294$ , one-sided). To clarify more precisely the impact of risk aversion on trading activity, I run two regression models – Equation 10 and Equation 11 (see Table 3.4) – whose specifications are the same as of Equations 6 and 7, with the variable measuring the number of safe choices added.

From inspection of Equation 10 it can be seen that the degree of risk aversion, measured as the number of safe choices, has no significant effect on the intensity of individual engagement in trading activity. Other relations between the variables remain similar to the observed in

Equation 6: 1) a one standard deviation increase in MAE raises trading activity by 0.348 standard deviations ( $p < 0.05$ , two-sided); 2) a one standard deviation increment in age reduces trading activity by -0.266 standard deviations ( $p < 0.05$ , one-sided); 3) interaction term between gender and overconfidence indicates that for every unit increase in NBS for male subjects there is reduction in trading activity by -0.448 units; in contrast, for female subjects increase of 0.859 units in trading activity is observed ( $p < 0.1$ , one-sided). In the model tested in Equation 11 also no significant influence of subjects' risk aversion on their engagement in trading activity was detected. Other relations between the variables remain similar to the observed in Equation 7: 1) a one standard deviation increase in TAE raises trading activity by 0.433 standard deviations ( $p < 0.05$ , two-sided); 2) a one standard deviation increment in age reduces trading activity by -0.300 standard deviations ( $p < 0.05$ , two-sided); 3) interaction term between gender and overconfidence indicates that for every unit increase in NBS for male subjects there is reduction in trading activity by -0.475 units; in contrast, for female subjects increase of 0.854 units in trading activity is observed ( $p < 0.1$ , one-sided).

It can be concluded that, in this sample, differences in experimental market outcomes between the traders were overconfidence and not risk aversion driven.

## CONCLUSIONS

The aim of this article was to investigate the influence of behavioral factors, namely the degree of overconfidence and risk aversion, on financial decision making of economic subjects. For this purpose two kinds of experiments were conducted. The first one, whose design followed Smith, Suchanek and Williams (1988), was the asset market experiment, whereas the second was aimed at individual risk aversion measurement. Hypotheses, which were tested in the context of the suggested experimental design, were built on exploration of findings from the foregoing overconfidence research in financial markets. The presented evidence builds on individual traders' decisions.

The usual finding from experimental literature testing for the impact of overconfidence on variation in subjects' trading activity and performance (earnings) in market settings where both overconfident and rational traders are present, is that the former engage in more trading activity (Odean, 1998; Deaves et al., 2009) and face welfare reductions (e.g. Biais et al., 2005; Kirchler and Maciejovsky, 2002) compared to the latter. In contrast to these works, in the present experiment subjects, based on their pre-experimental overconfidence scores, were assigned to the two types of markets – rational and overconfident – and in the course of

experiment they could interact only with participants of their own “type” (rational or overconfident). Within each of the conducted markets, the degree of overconfidence varied (even slightly) from participant to participant, and this part of experiment was built upon the assumption, that this was the only source of individual heterogeneity.

The design of experimental sessions, where all participants had identical initial endowments and information access, did not create preconditions for the active use of market-place for portfolio balancing purposes; correspondingly (almost) no trade should have occurred. However, the empirical data suggest that average trading activity was rather high and some traders demonstrated trading activity that equaled several times the outstanding market inventory. The results of data analysis supported the hypothesis that individuals with higher degree of overconfidence engaged in trading activity more actively. Yet, holding all other factors constant, with increase in overconfidence men have completed fewer stock market transactions than women. This contradicts the result of Barber and Odean (2001) that men trade more than women. On the other hand Deaves et al. (2009) found that after controlling for overconfidence there was no difference in trading activity between males and females. Also in Glaser and Weber (2007) gender was not significantly related to the trading volume measures. Other control variables, which had significant effect on the number of stock market transactions per person, were: errors in predicting average asset market prices and subjects’ experience, measured as their age or duration of study. Namely, forecasting errors, inducing false future price expectations, forced subjects to engage in trading activity more actively, whereas experience had significant negative impact on their involvement in trading activity.

Statistical data suggest that, contrary to the formulated hypothesis, overconfidence had positive effect on gains from trade. Other factors that significantly affected variation of relative profits were: trading activity, gender, forecasting errors and the final portfolio size. As it was expected, active engagement in trade had negative consequences on individual gains. High turnover players were significantly outperformed by low turnover players, namely participants in the lowest trading activity quartile gained on average 38% more ECUs than participants in the highest quartile. However, holding all other factors constant, with increase in the number of market transactions males incurred smaller losses, or even some yield, compared to females. In line with the previous research, forecasting errors, bringing about mistakes in financial decision making, were associated with losses. The number of assets in trader’s final inventory, which was found to be not overconfidence dependent, proved to be a significant determinant of reduction in gains from trade. A curious fact is that, although females held fewer assets in their portfolios than males, their gains were significantly lower.

Based on the above-mentioned findings, it can be concluded that also in the setting, where two “types” of subjects were separated from each other, performance and trading activity were overconfidence dependent and even small variations in miscalibration among players belonging to the same “type” were sufficient to evoke this effect.

At the completion of subsequent risk aversion measurements in the reduced sample, consisting of sixteen rational and sixteen overconfident former asset market experiment participants, the collected data revealed that subjects on average were risk averse. Inconsistent with the proposition that overconfident subjects would be more risk loving, statistical tests detected no significant difference between the two types of traders in terms of the average number of safe choices. The presented evidence implied that, neither the number of safe choices, nor demographics (age, gender, and semester) had explanatory power for the individual degree of overconfidence. Prior to the experiment it was expected that risk aversion would have negative effect on participant’s trading activity and her final portfolio size. However, no significant relationship between these variables was detected. Other factors, which were found to influence trading activity in the complete sample, remained significant. Hence, it can be concluded that in the reduced sample, differences in experimental market outcomes between the traders were overconfidence and not risk aversion driven.

For future research in the area of psychological motivation of asset market experiment participants’ behavior it would be beneficial to establish the origin of differences between rational and overconfident traders’ economic behavior. Investment decisions in this experiment are based on beliefs concerning the likelihood of the two kinds of independent uncertain events: 1) the size of the dividend at the end of the period, and 2) the probability to resell the asset at a higher price (finding a “greater fool”). Research is called for to examine, whether overconfident traders generate subjective probabilities of occurrence of these events, which significantly deviate from the objective ones, namely by optimistically overestimating probabilities of favorable to them outcomes (e.g. maximum dividend value) and almost neglecting unfavorable ones. Another possible extension could be the assessment and control of personality characteristics other than overconfidence and risk aversion. E.g. students’ intelligence assessed as their IQ score, their attitude towards deception and manipulation aimed at personal gain attainment (Machiavellianism), psychological traits such as extraversion, openness, neuroticism and etc.

## REFERENCES

- Baker, K. H., Nofsinger, J. R., (2002), Psychological biases of investors. *Financial Services Review*, Vol. 11(2), p. 97-116.
- Baker, R. J., Laury, S. K., Williams, A. W., (2008), Comparing small-group and individual behavior in lottery-choice experiment. *Southern Economic Journal*, Vol. 75, p. 367-382.
- Barber, B. M., Odean, T., (2000), Trading is hazardous to your wealth: the common stock investment performance of individual investors. *Journal of Finance*, Vol. 55(2), p. 773-806.
- Barber, B. M., Odean, T., (2001), Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, Vol. 116(1), p. 261-292.
- Barber, B. M., Odean, T., (2002), Online investors: do the slow die first? *Review of Financial Studies*, Vol. 15(2), p. 455-487.
- Benjamin, D., Brown, S., Shapiro, J., (2005), Who is behavioral? Harvard University Working Paper.
- Benos, A., (1998), Aggressiveness and survival of overconfident traders. *Journal of Financial Markets*, Vol. 1, p. 353-383.
- Biais, B., Hilton, D., Mazurier, K., Pouget, S., (2005), Judgmental overconfidence, self-monitoring and trading performance in an experimental financial market. *Review of economic studies*, Vol. 72(2), p. 287-312.
- Binswanger, H., (1980), Attitudes towards risk: experimental measurement in rural India. *American Journal of Agricultural Economics*, Vol. 62, p. 395-407.
- Brachinger, H.W., Schubert, R., Brown, M., Gysler, M., (1999), Financial decision making: are women really more risk averse? *American Economic Review*, Vol. 89(2), p. 381-385.
- Camacho-Cuena, E., Requate, T., Waichman, I., (2009), Investment Incentives under Emission Trading: An Experimental Study. Working paper.
- Chevalier, J., Ellison, G., (1999), Career concerns of mutual fund managers. *Quarterly Journal of Economics*, Vol. 114, p. 389-432.
- Chuang, W. I., Lee, B. S., (2006), An empirical evaluation of the overconfidence hypothesis. *Journal of Banking and Finance*, Vol. 30(9), p. 2489-2515.
- Crosan, R., Gneezy, U., (2008), Gender differences in preferences. *Journal of Economic Literature*, Vol. 47(2), p. 1-27.

- Deaves, R., Lüders, E., Luo, G. Y., (2009), An Experimental Test of the Impact of Overconfidence and Gender on Trading activity. *Review of finance*, Vol. 13(3), p. 555–575.
- De Bondt, W. F. M., Thaler, R. H., (1994), Financial decision-making in markets and firms: a behavioral perspective. NBER Working Paper Nr. 4777.
- De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., (1991), The survival of noise traders in financial markets. *The Journal of Business*, Vol. 64(1), p. 1-19.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G. G., (2005), Individual risk attitudes: new evidence from a large, representative, experimentally-validated survey. Discussion Paper Nr. 1730.
- Dohmen, T., Falk A., Huffman, D., Sunde, U., (2007), Are risk aversion and impatience related to cognitive ability? IZA Discussion Paper Nr. 2735.
- Durand, R.B., Newby, R., Sanghani, J., (2006), An intimate portrait of the individual investor. University of Western Australia Working Paper.
- Fellner, G., Maciejovsky, B., (2007), Risk attitude and market behavior: evidence from experimental asset markets. *Journal of Economic Psychology*, Vol. 28(3), p. 338-350.
- Fenton-O’Creevy, M., Nicholson, N., Soane, E., Willman, P., (2003), Trading on illusions: unrealistic perceptions of control and trading performance. *Journal of Occupational and Organizational Psychology*, Vol. 76, p. 53–68.
- Fischbacher, U., (2007), z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, Vol. 10(2), p. 171-178.
- Fischhoff, B., Slovic, P., Lichtenstein, S., (1977), Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 3, p. 552-564.
- Gysler M., Kruse, J., Schubert, R., (2002), Ambiguity and gender differences in financial decision making: an experimental examination of competence and confidence effects. CER-ETH Economics working paper series 02/23.
- Giardini, F., Coricelli, G., Joffily, M., Sirigu A., (2008), Overconfidence in predictions as an effect of desirability bias. In Abdellaoui, M., and Hey, J.D. (Eds.), “Advances in Decision Making Under Risk and Uncertainty”, p. 163-180.
- Gillette, A. B., Stevens, D. E., Watts, S. G., Williams, A. W., (1999), Price and volume reactions to public information releases: an experimental approach incorporating traders’ subjective beliefs. *Contemporary Accounting Research*, Vol. 16(3), p. 437-479.

- Glaser, M., Nöth, M., Weber, M., (2003), Behavioral finance. In Koehler, D. J., and Harvey N. (Eds.), “Blackwell Handbook of Judgment and Decision Making”, p. 527-546.
- Glaser, M., Weber, M., (2007), Overconfidence and trading volume. *The Geneva Risk and Insurance Review*, Vol. 32(1), p. 1-36.
- Grable, J. E., (2000), Financial risk tolerance and additional factors that affect risk taking in everyday money matters. *Journal of Business and Psychology*, Vol. 14(4), p. 625-630
- Fraser, S., Greene, F. J., (2006), The Effects of experience on entrepreneurial optimism and uncertainty. *Economica*, Vol. 73(290), p 169-192.
- Frederick, S., (2006), Cognitive reflection and decision making. *Journal of Economic Perspectives*, Vol. 19(4), p. 25-42.
- Friedman, M., (1953), The case for flexible exchange rates. In “Essays in Positive Economics”, Chicago: University of Chicago Press, p. 157-203.
- Grable, J.E., (2000), Financial risk tolerance and additional factors that affect risk taking in every day money matters. *Journal of Business and Psychology*, Vol. 14(4), p.625–630.
- Gysler, M., Kruse, J., Schubert, R., (2002), Ambiguity and gender differences in financial decision making: an experimental examination of competence and confidence effects. CER-ETH Economics working paper series, Working paper Nr. 02/23.
- Hanna, S., Lee, H., (1995), Empirical patterns of risk tolerance. *Proceedings: Academy of Financial Services*.
- Hariharan, G., Chapman, K.S., Domian, D.L., (2000), Risk tolerance and asset allocation for investors nearing retirement. *Financial Services Review*, Vol. 9 (2), p.159–170.
- Hirshleifer, D., Luo, G. Y., (2001), On the survival of overconfident traders in a competitive securities market. *Journal of Financial Markets*. Vol. 4, p. 73-84.
- Hirota, S., Sunder, S., (2007), Price bubbles sans dividend anchors: evidence from laboratory stock markets. *Journal of Economic Dynamics & Control*, Vol. 31, p. 1875-1909.
- Holt, Ch. A., Laury, S. K., (2002), Risk aversion and incentive effects. *American Economic Review*, Vol. 92(5), p.1644-1655.
- Hong, H., Kubik, J.D., Solomon, A., (2000), Security analysts’ career concerns and herding of earnings forecasts. *RAND Journal of Economics*, Vol. 31(1), p.121–144.
- Jianakoplos, N. A., Bernasek, A., (1998), Are women more risk averse? *Economic Inquiry*, Vol. 36, p. 620-630.



- Keller, C., Siergist, M., (2006), Investing in stocks: the influence of financial risk attitude and values-related money and stock market attitudes. *Journal of Economic Psychology*, Vol. 27(2), p. 285–303.
- Kyle, A., Wang, F. A., (1997), Speculation duopoly with agreement to disagree: can overconfidence survive the market test? *Journal of Finance*, Vol. 52, p. 2073-2090.
- Lei, V., Noussair, Ch. N., Plott, Ch. R., (2001), Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs. actual irrationality. *Econometrica*, Vol. 69(4), p. 831-859.
- Kahneman D., Riepe, M. W., (1998), Aspects of investor psychology. *Journal of Portfolio Management*, Vol. 24(4), p. 52-65.
- Kirchler, E., Maciejovsky, B., (2002), Simultaneous over- and underconfidence: evidence from experimental asset markets. *Journal of Risk and Uncertainty*, Springer, Vol. 25(1), p. 65-85.
- Kourtidis, D., Šević, Ž., Chatzoglou, P., (2010), Investors' trading activity: a behavioral perspective. *International Journal of Trade and Global Markets*, Vol. 3(1), p. 52–67.
- Lakonishok, J., Shleifer, A., Vishny, R. W., (1992), The impact of institutional trading on stock prices. *Journal of Financial Economics*, Elsevier, Vol. 32(1), p. 23-43.
- Menkhoff, L., Schmidt, U., Brozynski, T., (2006), The impact of experience on risk taking, overconfidence, and herding of fund managers: complementary survey evidence. *European Economic Review*, Vol. 50(7), p. 1753-1766
- Mullainathan, S., Thaler, R., (2000), Behavioral economics. Massachusetts Institute of Technology Department of Economics Working Paper Series, Working Paper 00-27
- Nöth, M., Weber, M., (2003), Information aggregation with random ordering: cascades and overconfidence. *The Economic Journal*, Vol. 113, p. 166-189.
- Odean, T., (1998), Volume, volatility, price and profit when all traders are above average. *Journal of Finance*, Vol. 53(6), p. 1887 -1934.
- Odean, T., (1999), Do investors trade too much? *American Economic Review*, Vol. 89(5), p. 1278-1298.
- Russo, J. E., Schoemaker, P. J. H., (1992), Managing Overconfidence. *Sloan Management Review*, Vol. 33, p. 7-17.

- Roszkowski, M.J., (1998), Risk tolerance in financial decisions. In Cordell, D.M. (Ed.), "Readings in Financial Planning", Bryn Mawr, PA: The American College, p.281–328.
- Scheinkman, J. A., Xiong, W., (2003), Overconfidence and speculative bubbles. *Journal of Political Economy*, Vol. 111, p. 1183-1219.
- Scheinkman, J. A., Xiong, W., (2004), Heterogeneous beliefs, speculation and trading in financial markets. In Carmona, R. A., Cinlar, E., Ekeland, I., Jouini, E., Scheinkman, J. A., and Touzi, N. (Eds.), "Paris-Princeton Lectures on Mathematical Finance", p. 217-250.
- Schmidt, U., Traub, S., (2002), An experimental test of loss aversion. *Journal of Risk and Uncertainty*, Vol. 25, p. 233-249.
- Smith, V. L., Suchanek G. L., Williams A. W., (1988), Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econometrica*, Vol. 56(5), p. 1119-1151.
- Sundén, A.E., Surette, B.J., (1998), Gender differences in the allocation of assets in retirement saving plans. *The American Economic Review*, Vol. 88(2), p.207–211.
- Sung, J., Hanna, S., (1996), Factors related to risk tolerance. *Financial Counseling and Planning*, Vol. 7(1), p.11–20.
- Yao, R., Hanna, S., (2005), The effect of gender and marital status on financial risk tolerance. *Journal of Personal Finance*, Vol. 4(1), p.66–85.

## APPENDIX 3.A: DESCRIPTIVE STATISTICS (ASSET MARKET EXPERIMENT).

	N	M	SD	Min.	Max.
Profit	60	390.36	197.89	4.20	906.20
Relative profit	60	0.00	1.83	-3.58	4.78
Profit / initial portfolio value	60	3.61	1.83	0.04	8.39
Bias score	60	11.20	12.08	-5.89	43.50
Age	60	22.77	2.13	19.00	28.00
Semester	54	3.39	2.10	1.00	12.00
End assets	60	3.00	4.04	0.00	18.00
MAE	60	7.24	3.76	1.79	20.64
TAE	60	55.33	42.40	5.00	202.00
Trading activity (average)	60	0.89	0.47	0.14	2.25

APPENDIX 3.B: BIAS SCORE, AGE, SEMESTER FOR DIFFERENT EXPERIMENTAL SAMPLES

<i>Whole Sample</i>					
	N	M	SD	Min	Max
Bias score	60	11.20	12.08	-5.89	43.50
Age	60	22.77	2.13	19.00	28.00
Semester	60	3.39	2.10	1.00	12.00
<i>Overconfident Traders</i>					
	N	M	SD	Min	Max
Bias score	30	21.33	8.26	10.17	43.50
Age	30	22.83	1.88	19.00	27.00
Semester	30	2.96	1.37	1.00	5.00
<i>Rational Traders</i>					
	N	M	SD	Min	Max
Bias score	30	1.06	4.02	-5.89	6.78
Age	30	22.70	2.38	19.00	28.00
Semester	30	3.81	2.60	1.00	12.00
<i>Male Participants</i>					
	N	M	SD	Min	Max
Bias score	35	12.08	11.91	-4.72	43.50
Age	35	22.51	1.85	19.00	27.00
Semester	35	3.19	1.71	1.00	7.00
<i>Female Participants</i>					
	N	M	SD	Min	Max
Bias score	25	9.96	12.45	-5.89	38.89
Age	25	23.12	2.45	19.00	28.00
Semester	25	3.68	2.59	1.00	12.00

APPENDIX 3.C: RISK AVERSION EXPERIMENT SAMPLE

<i>Whole Sample</i>					
	N	M	SD	Min	Max
Safe choices	32	5.66	1.82	0.00	9.00
Age	32	24.34	1.94	20.00	29.00
Semester	29	5.69	2.05	3.00	9.00
Bias score	32	11.09	10.49	-3.72	35.00
<i>Overconfident Traders</i>					
	N	M	SD	Min	Max
Safe choices	16	5.50	2.19	0.00	9.00
Age	16	24.00	1.97	20.00	27.00
Semester	16	4.93	1.33	3.00	7.00
Bias score	16	20.17	6.48	10.17	35.00
<i>Rational Traders</i>					
	N	M	SD	Min	Max
Safe choices	16	5.81	1.42	3.00	9.00
Age	16	24.69	1.92	22.00	29.00
Semester	13	6.08	1.89	3.00	9.00
Bias score	16	2.01	3.10	-3.72	6.06
<i>Male Participants</i>					
	N	M	SD	Min	Max
Safe choices	19	5.68	2.00	0.00	9.00
Age	19	23.95	1.84	20.00	27.00
Semester	17	5.47	1.62	3.00	8.00
Bias score	19	11.95	10.02	-2.39	35.00
<i>Female Participants</i>					
	N	M	SD	Min	Max
Safe choices	13	5.62	1.61	3.00	9.00
Age	13	24.92	2.02	21.00	29.00
Semester	11	5.45	1.86	3.00	9.00
Bias score	13	9.84	11.45	-3.72	30.50

### INSTRUCTIONS

In this experiment you are to make ten decisions. Each decision is a choice between two paired lotteries - “Option A” and “Option B” such as those presented below:

Decision	Option A	Option B	Your Choice
1	3.00 EUR with 10% chance, 2.40 EUR with 90% chance	5.78 EUR with 10% chance, 0.15 EUR with 90% chance	A <input type="radio"/> <input type="radio"/> B
⋮	⋮	⋮	⋮
10	3.00 EUR with 100% chance	5.78 EUR with 100% chance	A <input type="radio"/> <input type="radio"/> B

You will make ten choices but only one of them will be used in the end of the experiment to determine your earnings. Each decision has an equal chance of being used in the end.

***How one of the decisions is going to be chosen?*** We will use a random number generator in the computer which will generate a number between 1 and 10. The generated number is the number of decision that will be used to determine your payment. Each number has an equal probability to occur, so you have to think carefully about EACH of the Decisions!

***How payment for a lottery of your choice is determined?*** After one of the decisions has been randomly selected, the computer will generate another random number. This random number determines your earnings for the Option (A or B) that you previously selected for the decision being used. E.g. if you choose Option A in the first decision row shown above, you will have a 10% chance of earning 3.00 EUR and a 90% chance of earning 2.40 EUR. Similarly Option, B offers a 10% chance of earning 5.78 EUR and a 90% chance of earning 0.15 EUR.

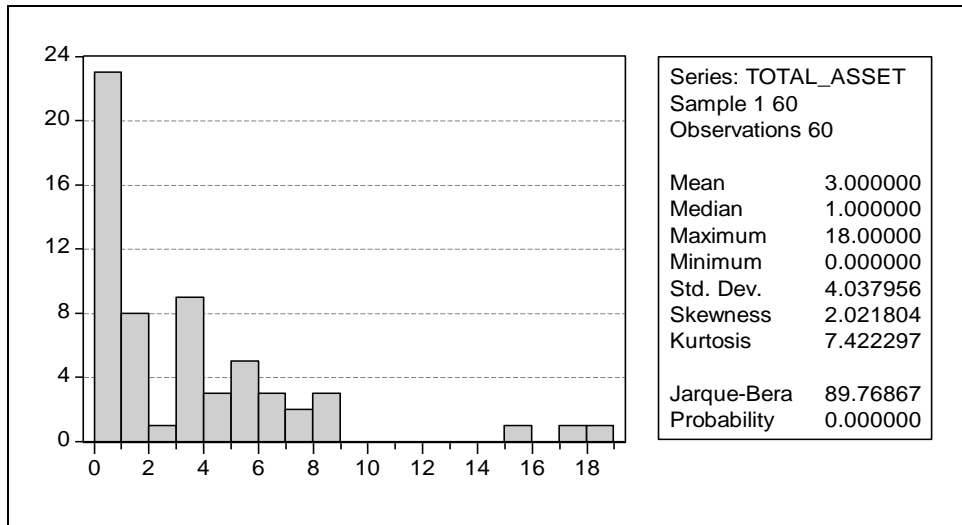
***What is the sequence of your actions?*** For each decision row you will have to choose between Option A and Option B by clicking on one of the “circles” on the right side of the screen. You will make ten choices. You may choose A for some decision rows and B for other rows and you may change your decisions and make them in any order. You have 15 minutes

make your choices. When you are finished, press the “Submit” button. After you have pressed this button you can no longer change your choices. Then a random number generator will choose a number of the decision that will be used for payment. A second random number will determine the outcome of the lottery you chose.

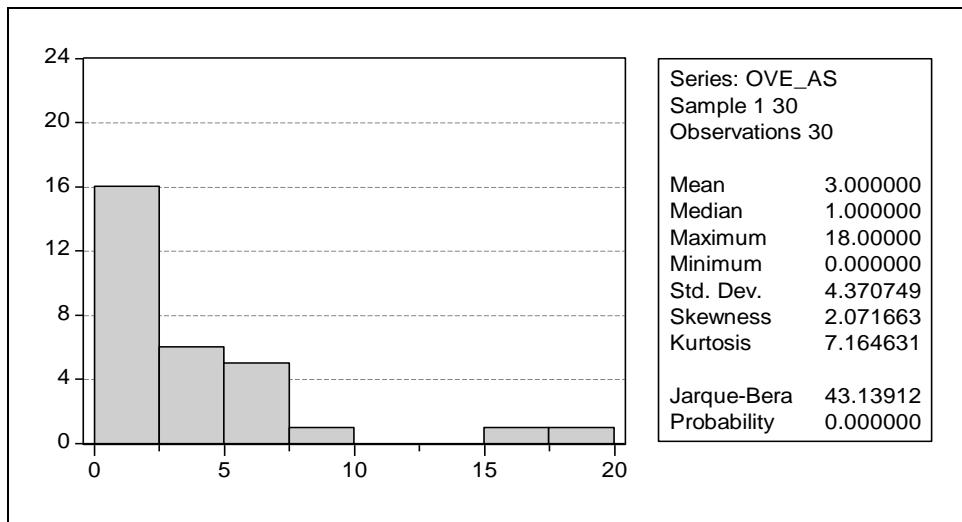
**Important:** please do not talk with anyone! Raise your hand if you have a question.

APPENDIX 3.E: FINAL INVENTORIES DISTRIBUTION IN

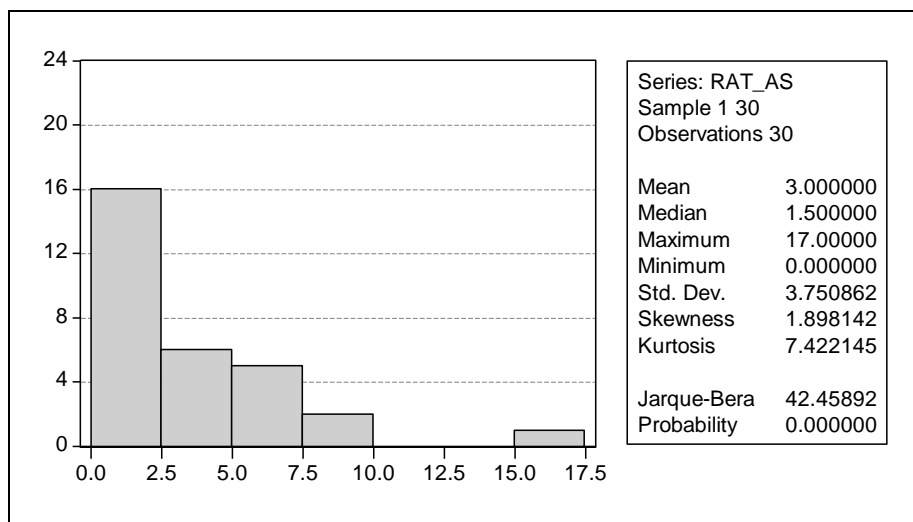
a. The whole sample



b. Overconfident markets sample



c. Rational markets sample





APPENDIX 3.F: REGRESSION OF OVERCONFIDENCE ON RISK AVERSION AND OTHER FACTORS

	1	2
C	0.956 <sup>*</sup> (0.522)	0.690 <sup>***</sup> (0.215)
Safe choice	-0.011 (0.021)	-0.017 (0.019)
Gender	0.021 (0.080)	0.067 (0.084)
AGE	-0.023 (0.021)	
Semester		-0.051 <sup>*</sup> (0.026)
N	32	28
R2	0.065	0.182
SEreg	0.22	0.21

\*\*\* - 0.01,

\* - 0.1

**Affirmation**

I hereby affirm that I have completed my doctoral thesis entitled, “Experimental Studies of Overconfidence in Financial Markets” entirely on my own and unassisted, and that I have specially marked all of the quotes I have used from other authors as well as those passages in my work that are extremely close to the thoughts presented by other authors, and listed the sources in accordance with the regulations I have been given.

11.10.2010

Julija Michailova

**EUROPEAN  
CURRICULUM VITAE  
FORMAT**



**PERSONAL INFORMATION**

Name	<b>JULIJA MICHAILOVA</b>
Gender	Female
Date and place of birth	29.10.1980. Vilnius, Lithuania
Nationality	Lithuanian
Address	Olshausenstraße 66, 24118 Kiel, Germany
Telephone	+491708444578
E-mail	julija_michailova@yahoo.com

**EDUCATION AND TRAINING**

- Dates  
October 2006 – December 2010  
Christian Albrechts Universität, Kiel, Germany
- Name and type of organization providing education and training  
Doctoral studies in Quantitative Economics
- Principal subjects/occupational skills covered  
Thesis title: Experimental Studies of Overconfidence in Financial Markets  
Thesis defended on 21.12.2010
- Dates  
February 2007 – May 2009  
Kiel Institute for the World Economy
- Name and type of organization providing education and training  
Advanced Studies Program in International Economic Policy Research
- Principal subjects/occupational skills covered
- Dates  
October 2005 – October 2006  
Christian Albrechts Universität, Kiel, Germany
- Name and type of organization providing education and training  
Visiting student at the International Doctoral Programme “Quantitative Economics”
- Principal subjects/occupational skills covered
- Dates  
September 2004 to September 2006  
Vilnius Gediminas Technical University, Vilnius, Lithuania
- Name and type of organization providing education and training  
Doctoral studies in Economics  
(terminated since September, 2006)
- Principal subjects/occupational skills covered
- Dates  
September 2002 to June 2004  
Vilnius Gediminas Technical University, Vilnius, Lithuania
- Name and type of organization providing education and training  
Advanced knowledge of business and its economic, legal and social context.  
Focus on business finance.
- Principal subjects/occupational skills covered  
Thesis title: “Construction and Management of Optimal Stock Portfolio”  
Masters degree in Management and Business Administration
- Title of qualification awarded
- Dates  
September 1998 to June 2002  
Vilnius Gediminas Technical University, Vilnius, Lithuania
- Name and type of organization providing education and training  
Basic knowledge of business and its economic, legal and social context.  
Focus on business finance.
- Principal subjects/occupational skills covered  
Thesis title: “Securities Portfolio Construction and Analysis”  
Bachelors degree in Management and Business Administration
- Title of qualification awarded

- Dates
- Name and type of organization providing education and training
- Principal subjects/occupational skills covered

September 2001 to January 2001  
Seinajoki Business School (Seinajoki Polytechnic, Finland)

International Business Program (Erasmus Exchange Semester)

## WORK EXPERIENCE

- Dates
- Name and address of employer
  - Type of business or sector
  - Occupation or position held
- Main activities and responsibilities
  - Note

01.06.2010 – 31.05.2010  
Christian-Albrechts University of Kiel, Olshausenstraße 40, Kiel, Germany  
Institution of Higher Education  
Research assistant at the chair of Public Economics, Social Policy and Health Economics. Faculty of Economics and Social Sciences  
Literature review and analysis, collection of data.  
Part-time

- Dates
- Name and address of employer
  - Type of business or sector
  - Occupation or position held
- Main activities and responsibilities
  - Note

01.08.2006 – 31.07.2007  
Kiel Institute for the World Economy, Düsternbrooker Weg 120, Kiel, Germany  
Center for research in global economic affairs, economic policy consulting, economic education and documentation  
Research Assistant at project "Beyond Europe", work package "Analysis of the Economic and Institutional Consequences of WTO Accession and of Future EU-CIS Free trade Agreements"  
Literature review and analysis, collection of data, communication with regional partners, other.  
Part-time

- Dates
- Name and address of employer
  - Type of business or sector
  - Occupation or position held
- Main activities and responsibilities
  - Note

05.01.2005 – 28.06.2005  
Vilnius Gediminas Technical University, Saulėtekio al. 11, Vilnius, Lithuania  
Institution of Higher Education  
Assistant at the Chair of Financial Engineering, Business Management Faculty  
Data collection, analysis, scientific work, tutorials  
Full-time

- Dates
- Name and address of employer
  - Type of business or sector
  - Occupation or position held
- Main activities and responsibilities
  - Note

26.07.2004 – 20.12.2004  
The Lithuanian Chamber of Auditors, J. Galvydžio g. 5 - 104, Vilnius, Lithuania  
Association, which unifies certified auditors, coordinates their actions, represents and defends their interests or meets other public interests  
Expert on Organization of Training  
Organization of training, seminars, refresher courses, promotion examinations and etc.  
Full-time

- Dates
- Name and address of employer
  - Type of business or sector
  - Occupation or position held
- Main activities and responsibilities
  - Note

22.09.2003 – 07.07.2004  
Ligajos technika Ltd., Žvalgų g. 8, Vilnius, Lithuania  
Trade, installation and maintenance of air conditioning and ventilation systems.  
Office Administrator  
Preparation and control of documents, file-keeping, doing enterprise's correspondence, relations with clients, personnel management, reception of company's guests, marketing, statistics' analysis, debt monitoring, creation and implementation of the quality management system, etc.  
Full-time

## PERSONAL SKILLS AND COMPETENCES

### LANGUAGES

- Reading skills
- Writing skills
- Verbal skills

Russian	Lithuanian	English	Polish	German
Native	Native	Fluent	Good	Good
Native	Native	Fluent	Fair	Good
Native	Native	Fluent	Good	Good

### SOCIAL SKILLS AND COMPETENCES

- Good communication skills, gained both through my work experience as an office administrator, training organizer and as an assistant at the university, and through my personal life.
- Ability to adapt to multicultural environment, gained in Germany, Finland and Lithuania.

### ORGANIZATIONAL SKILLS AND COMPETENCES

- Doctoral thesis: I developed, managed and conducted multi-stage, interdisciplinary experimental project.
- During my work as an Expert on Training Organization I organized numerous refresher courses, trainings, several seminars for auditors, accountants, managerial staff in Vilnius, Kaunas and Palanga (Lithuania) as well as promotion examinations for those, wishing to become auditors.
- Whilst working for "Ligajos technika" Ltd. I was in charge for our company's participation in annual exhibition of building/constructing sector "Resta"; during this exhibition I also organized a seminar on our production presentation for our professional clients. As a direction's representative for quality I organized several lectures for my colleagues on quality management.

### COMPUTER SKILLS AND COMPETENCES

Good command of these Microsoft Office tools: Word, Excel and Power Point; basic knowledge of Access, Publisher.  
Knowledge of EViews, SPSS, zTree (Zurich Toolbox for Readymade Economic Experiments). Basic knowledge of Stata, AutoCAD.

## ADDITIONAL INFORMATION

### SCHOLARSHIPS

10.2005 – 07.2006	Stadt Kiel Stipendium
10.2006 – 07.2007	Stadt Kiel Stipendium
10.2007 – 03.2010	DAAD scholarship

### SEMINARS/ SUMMER SCHOOLS/ COURSES:

- 01.12.2008 – 05.12. 2008 The 7th GARNET PhD seminar "Global Governance, Regionalism & the Role of the EU: The Gender Dimension". Organized by GARNET Network of Excellence and university of Kassel (Germany).
- 18.06.2007 - 22.06.2007 Barcelona LeeX Experimental Economics Summer School in Macroeconomics, organized by Universitat Pompeu Fabra (Spain).
- 28.08.2006 – 07.09.2006 Course "International Trade" of the "Advanced Studies in International Economic Policy Research" program at the Kiel Institute for World Economy.
- 04.07.2005 – 29.07.2005 Potsdam Summer Academy 2005 "Banking, Insurance and the Public Sector: Empirical Experience and Policy Advice", organized by Potsdam University (Germany).
- 30.06.2003 – 13.07.2003 Summer university "Brücken bauen – über Brücken gehen", organized by DAAD, FernUniversität in Hagen and Riga's Technical University in Riga (Latvia).

#### PERSONAL INTERESTS:

Reading books and drawing. Interested in sports activities, especially lawn tennis. Fond of traveling and meeting people of different cultures.

#### PUBLICATIONS:

J. Michailova., I. Melnykovska. Gender, Corruption and Sustainable Growth in Transition Countries // Journal of Applied Economic Sciences. 2009, Vol. IV, Nr. 9.

M. Tvaronavičienė, J. Michailova. Factors Affecting Securities Prices: Theoretical Versus Practical Approach // Journal Of Business Economics And Management. 2006, Vol. 7, Nr. 4.

M. Tvaronavičienė, J. Michailova. Optimal Stock Portfolio Construction by Means of H. Markowitz "Portfolio Theory" // Business: Theory and Practice (in Lithuanian). 2004, Vol. IV, Nr. 3.

#### CONFERENCES:

J. MICHAILOVA. Overconfidence and Bubbles in Experimental Asset Markets // 6th PhD Presentation Meeting of the Royal Economic Society. Conference organized by: City University London (UK), 15-16 January, 2011.

J. MICHAILOVA. Overconfidence and Bubbles in Experimental Asset Markets // 5th Nordic Behavioral and Experimental Economics Conference. Conference organized by: Aalto University of Helsinki (Finland), 12-13 November, 2010.

J. MICHAILOVA. Overconfidence and Bubbles in Experimental Asset Markets // 5th Alhambra Experimental Workshop. Conference organized by: University of Granada and University of Pompeu Fabra (Spain), 21-22 October, 2010.

J. MICHAILOVA. Overconfidence, Risk Aversion and (Economic) Behavior of Individual Traders in Experimental Asset Markets // Symposium Experimental Finance 2010. Conference organized by: University of Gothenburg (Sweden), 7-9 October, 2010.

J. Michailova. Development of an Overconfidence Test for the Economic Experiment // 6th International Student Conference "Empirical Studies in Social Sciences" organized by Izmir Economic University (Turkey), 14-15 April, 2010.

J. Michailova. The Role of Overconfidence in Creation of Stock-Prices' Bubbles: Human Based Experiment // Warsaw International Economic Meeting. Conference organized by: Warsaw University (Poland), 3-5 July, 2009.

J. Michailova. The Role of Overconfidence in Creation of Stock-Prices' Bubbles: Human Based Experiment // Conference on Economic Design. Conference organized by: Society for Economic Design & Maastricht University (Netherlands), 22-24 June, 2009.

J. Michailova., I. Melnykovska. Gender, Corruption and Sustainable Growth in Transition Countries. // The 7th GARNET PhD seminar "Global Governance, Regionalism & the Role of the EU: The Gender Dimension", Kassel, 1- 5 December, 2008.

J. Michailova. Indexes as the Source of Adequate Information on Securities Market // 8th conference of young scientists of Lithuania "Business in XXI century". Conference organized by Vilnius Gediminas Technical University (Lithuania), 10 February, 2005.

A. V. Rutkauskas, J. Michailova. Decision Making Management in Capital Markets of Transition Economies // III International Conference "Management And Engineering": conference organized by Sofia Technical University (Bulgaria), 20-24 June, 2005.

M. Tvaronavičienė, J. Michailova. Application of Markowitz Model for Optimal Stock Portfolio Selection in Capital Markets of Transition Economies: Case of Lithuania // International Conference on Business, Management and Economics. Conference organized by Yasar University (Turkey), 16-19 June, 2005.

M. Tvaronavičienė, J. Michailova. Application of the "Portfolio Theory" for Lithuanian Stock Portfolio Construction in the Context of Accession to the EU // International conference "European prospects for the national economy". Conference organized by University of Economics Varna (Bulgaria), 9-10 June, 2005.