Overconfidence and bubbles in experimental asset markets

Michailova, Julija and Schmidt, Ulrich

Helmut Schmidt University, Kiel Institute for the World Economy

September 2011

Overconfidence and Bubbles in Experimental Asset Markets

Forthcoming in "Journal of Behavioral Finance"

Julija Michailova
Department of Economics, Helmut Schmidt University
22043 Hamburg, Germany
E-mail: julija_michailova@yahoo.com

Ulrich Schmidt
Kiel Institute for the World Economy
24100 Kiel, Germany
E-mail: uschmidt@bwl.uni-kiel.de

Abstract

This paper investigates the relationship between market overconfidence and occurrence of stock-price bubbles. Sixty participants traded stocks in ten experimental asset markets. Markets were constructed on the basis of subjects' overconfidence: The most overconfident subjects form high overconfidence markets, and the least overconfident subjects low overconfidence markets. Prices in low overconfidence markets tend to track the fundamental asset value more accurately than prices in high overconfidence markets and are significantly lower and less volatile. Additionally we observe significantly higher bubble measures and trading volume in high overconfidence markets. Two possible explanations for these differences are analyzed: While price expectations are significantly higher in high overconfidence markets no differences in the average degree of risk aversion were detected.

Keywords: overconfidence, miscalibration, overprecision, overestimation, price bubbles, experimental asset market, risk aversion.

JEL Codes: C92, G12

1 Introduction

Since the seminal work of Smith, Suchanek and Williams (SSW [1988]) the occurrence of speculative bubbles in experimental asset markets with declining fundamental value is a well-established phenomenon. Several modifications of the design were investigated in order to detect possible causes for the emergence of bubbles and to control their size (cf. Porter and Smith [2003]). Though it has been recently shown that declining fundamental value of assets contributes to the emergence of bubbles (Kirchler, Huber, and Stöckl [2012]) behavioral characteristics of participants may also play a role. Up to date the only personal characteristic, which was found to influence the size of the bubbles in the setting of SSW [1988] was individual experience with the experiment (e.g. SSW [1988], Dufwenberg, Lindqvist and Moore [2005]). In the present paper we study the impact of another personal characteristic, namely overconfidence, on bubble formation in asset markets based on the SSW [1988] design. While overconfidence has a documented impact on trading behavior in experimental asset markets (Kirchler and Maciejovsky [2002], Biais et al. [2005], Deaves et al. [2009]), its relation to the occurrence of bubbles has not yet been addressed directly.

In the financial literature several effects are summarized under the concept of overconfidence: miscalibration, the better than average effect and illusion of control. Moore and Healy [2008] refer to these effects subsequently as to overprecision, overplacement and overestimation. Miscalibration (overprecision) is a cognitive bias that rests on the fact that people tend to overestimate the precision of their knowledge (cf. Lichtenstein et al. [1982]). Inclination of people to exaggerate their talents embodies itself in the better than average effect or overplacement (cf. Taylor and Brown [1988]). Illusion of control is linked to the exaggeration of the degree to which one can control one's fate (cf. Langer [1975]). It can also be defined as optimistic overconfidence - overestimation of the probabilities of the events that are advantageous to the subject (cf. Griffin and Brenner [2004]). The concept of overestimation of Moore and Healy [2008] includes not only Langer's illusion of control, but also overestimation of one's actual ability and performance. In this paper we assess overconfidence as miscalibration (overprecision). However, according to Moore and Healy [2008], the item-confidence paradigm, used to assess overconfidence, measures both overprecision and overestimation.

In this paper we report results of an experiment, designed to investigate the role of average overconfidence of subjects, comprising an experimental market (thereafter referred as *market overconfidence*) for the occurrence of stock-price bubbles and the emergence of other stylized facts like excessive trade volume and excessive price volatility. For this purpose we extend the basic design of SSW [1988] by a new feature, in which markets are composed on the basis of subjects' overconfidence, assessed in overconfidence measurement sessions. From these sessions we invited (i) subjects with lowest overconfidence scores and assigned them to one type of markets (referred to as *low overconfidence markets* in the sequel) and (ii) subjects with highest overconfidence scores who were assigned to a second type of markets (called *high overconfidence markets*). Within one market all subjects receive identical information so heterogeneity as e.g. in the model of Scheinkman and Xiong [2003] cannot account for bubbles in our study.

Main findings from our experiment can be summarized as follows. Consistent with theoretical analyses[1] (De Bondt and Thaler [1985], Shiller [2000], Benos [1998], Caballé and Sákovics [2003]), higher market overconfidence is accompanied by higher average market prices, higher volatility, larger deviations of prices from fundamental value, and higher trading volume. Prices in low overconfidence markets tend to track fundamental asset value more accurately than prices in high overconfidence markets, and are significantly lower. Moreover, bubble and crash patterns were observed in the aggregate prices in high overconfidence markets, whereas in low overconfidence markets no sudden drop of the aggregate market price to the fundamental value occurred.

We analyze two possible reasons why overconfidence has such a strong impact on our market outcomes. First, average degree of risk aversion might be lower in high overconfidence markets which would provide a direct rationale for higher asset prices as lower degrees of risk aversion imply higher certainty equivalents of risky assets. Second, more overconfident subjects may overestimate future prices and the probability to sell assets with profit in later rounds. The integration of a simple price forecasting task in the main experiment and a post-experimental assessment of risk attitudes reveal that more overconfident subjects indeed expect substantially higher prices whereas no significant difference of risk attitudes was observed.

The paper proceeds as follows. The next section discusses the design of our study consisting of overconfidence measurement, the experimental asset markets, and a post-experimental assessment of risk attitudes. Section 3 presents our experimental results and section 4 contains some concluding observations.

[1] For a review on overconfidence in theoretical finance please refer to Glaser, Nöth, and Weber [2004].

2 Experimental Design

2.1 Overconfidence Measurement

To measure individual overconfidence a specially developed instrument was used[2]. This instrument is a general knowledge test, consisting of 18 items. Each item has three alternative answers, of which only one is correct. Test is balanced with respect to the hard-easy effect[3] by including an equal number of questions of hard, moderate and easy difficulty levels. To avoid biases in overconfidence measures due to question content, items are not connected to economics or finance and are gender neutral. Instrument possesses good internal consistency (reliability): Cronbach's alpha for test confidence is 0.79 and for the bias score 0.68 (cf. Michailova and Katter, [2014]).

Previous economic experiments mostly assessed overconfidence through confidence interval elicitation for questions with the known numerical answers (cf. Russo and Schoemaker, [1992]). In contrast, this instrument uses different format, which is clearer to subjects and is less prone to overconfidence (cf. Klayman et al., [1999]). Namely, participants had to answer our 18 questions and for each question, had to assess how confident they were in the correctness of their answer. For this purpose they could use any number in the range from 33% (complete uncertainty) to 100% (complete certainty). There were no monetary incentives for these confidence elicitations. The under- or overconfidence of each participant was measured as a bias score (BS_{test}), which was calculated as the difference between the mean confidence level across all questions and the mean proportion of correct answers. A positive bias score represents overconfidence and a negative bias score represents underconfidence. A bias score of zero indicated an accurately calibrated person (confidence-neutral person).

$$BS_{test} = \frac{1}{N}\sum_{i=1}^{N}(c_i - a_i) \tag{1}$$

Where c_i is the confidence and a_i is the accuracy in answering item i; N is the total number of items.

Overconfidence measurement sessions were performed during several economics lectures at the University of Kiel. In each of the chosen classes, students were informed that they had an

2 For a detailed description of the instrument and procedure of its development please refer to Michailova and Katter [forthcoming].

3 Hard-easy 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 et al., [1991]).

opportunity to take part in the short experiment on the voluntary basis, for which a general knowledge test had to be filled out. For this activity 15 minutes were given. Participants in each overconfidence measurement session competed for the three prizes of 30, 20 and 10 EUR, which were awarded to those who answered the most questions right[4]. Before students started with the tests, a planned market experiment was advertised. These 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 e-mail address. Information about the market experiment was presented in such a way that subjects could not anticipate the link between it and the general knowledge test. This procedure allowed us to obtain a large pool of students with their estimated bias scores and to ensure that the two stages of the experiment were perceived by participants as two non-associated experiments.

More than 200 students showed interest in the forthcoming economic experiment. A database of the interested persons included information on 221 students' name, age, nationality, subject of studies, semester and overconfidence score. Consistent with previous research, subjects in the database on average were overconfident (BS_{test}: M = 11.78, SD = 10.58). As explained above, we focused in the experiment on the least and the most overconfident subjects, whom are further on called low overconfidence and high overconfidence subjects, respectively. These students were approached via e-mail and invited to register for the main experiment. It is important to mention, that overconfidence measurement was never mentioned to subjects, and they knew neither their own degree of overconfidence, nor had they any information about overconfidence of other participants. An invitation included a list of scheduled sessions from which subjects could choose one and register for participation; however, not known by any participant, high overconfidence subjects received a different list of sessions than low overconfidence subjects.

Appendix A presents data on the bias scores of the various experimental subgroups: all participants who were in the database and all students who participated in the experimental market sessions (a subsample of those in the database). All groups seemed to be substantially

[4] Subjects were not offered reward for the accuracy in probability elicitation (confidence in correctness of an answer). A common mechanism used to incentivize probability elicitation is *quadratic-scoring rule* (QSR). Michailova and Katter [forthcoming] present two arguments for not using QSR as a payment procedure in overconfidence measurement: First, under QSR subjects who are not risk neutral are inclined to misreport their confidence (cf. Winkler and Murphy, [1970]); second, QSR is not easy for the subjects to understand (cf. Read, 2005).

overconfident, except for the participants of the low overconfidence markets. In Appendix A we also test several hypotheses of the equality of the average overconfidence of different subgroups. Most importantly for the construction of our markets, the bias score of the participants of high overconfidence markets (M = 21.33, SD = 8.26) is significantly higher than the bias score of the participants of low overconfidence markets (M = 1.06, SD = 4.03). Additionally, we were successful in excluding a gender bias, as the mean equality hypothesis is failed to be rejected for the difference between overconfidence of male versus female subjects both among all students in our database, as well as among all participants of experimental asset market.

2.2 Experimental Asset Markets

For each of our ten market sessions six participants were recruited from the set of subjects who participated in overconfidence measurement sessions. None of our subjects participated before in a similar experiment. The 60 subjects were comprised of 35 males and 25 females, aged 19 to 28 (M = 22.73, SD = 2.06), and 87% were of German nationality. Approximate time required to conduct the experiment was one hour. Subjects earned on average 390.36 ECU (10.54 EUR, SD = 197.89) in the asset market (without the reward for the forecasting activity). Men earned on average more ECUs than women, 447 ECU compared to 335 ECU. This difference is significant (Mann-Whitney Z = -2.65, $p < 0.01$, one-sided). Instructions familiarized participants with the rules of the experimental market. English translation of instructions can be found in Appendix B.

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 (low overconfidence/high overconfidence) to which they were appointed based on the results of overconfidence measurement sessions. The experiment was programmed and conducted with z-Tree (Fischbacher [2007]).

At the beginning of sessions students were given time to read the detailed instructions and ask questions. At the end of the time devoted for reading the instructions, the experimenter read out loudly the most important information. Two trial periods followed, during which students could familiarize themselves with the experimental software, and again 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 and payoff.

The design of trading rounds followed Smith, Suchanek, and Williams [1988] with slight changes in the price forecasting task, and was performed as a continuous anonymous double auction. 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. 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 prices for asset units. Each participant could purchase asset units by spending an amount of their working capital, or sell units of the inventory and thereby increase their working capital. At the end of each 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, so the expected dividend of each asset amounted to 2.4 ECU in each trading period. As the terminal asset value is zero, the fundamental value of an asset, thus, equals $n \times 2.4$ ECUs, where n is the number of trading periods remaining until the end of the experiment.

At the end of trading periods, participants were shown market summary information from the past period, 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%. 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 for the price prediction task, as used by e.g. SSW [1988], was chosen over interval estimation due to several reasons. First, overconfidence measures obtained through interval estimation 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 accurately calibrated, and underconfident. Second, this form of price prediction task enabled comparison between the two overconfidence measures: the one obtained in overconfidence measurement sessions and the other obtained in experimental asset market.

At the end of the experiment, subjects were paid in cash the amount of money corresponding to their final working capital converted at the predefined exchange rate to Euros. Final working capital (FWC) equaled:

FWC = (300 ECU starting capital) + (dividend earnings) + (stock sales revenue) - (stock purchase cost) (2)

Reward for the accuracy in predicting next period's average price was constructed to be an additional income source in order to encourage conscious engagement in the experiment. The closer the prediction was to the actual average market price, the higher was the reward[5]. The reward scheme used in the experiment was similar to the suggested by Haruvy, Lahav, and Noussair [2007][6]:

Level of Accuracy	Reward
Within 95% -105% of actual price	3 ECU
Within 87.5%-112.5% of actual price	1 ECU
Within 75%-125% of actual price	0.5 ECU

As in the overconfidence measurement via general knowledge test, subjects were rewarded only for their accuracy in predicting average price but not in the elicited probabilities that their forecast was correct. Both monetary reward and the feedback about their predictions' accuracy were used for improving the subjects' calibration in the price prediction task.

2.3 Post-experimental Measurement of Risk Attitudes

A few months after the conclusion of the main experiment we invited 50% of our subjects to an additional experiment, measuring risk attitudes with the method of Holt and Laury [2002]. From these 32 subjects 16 were high overconfidence and 16 low overconfidence subjects. Experimental design involved a choice list (see Table 1) where subjects had to choose in ten rows between two lotteries, option A and option B. Option A was a "safe" choice and paid either 3EUR or 2.40 EUR; Option B was a "risky" choice and paid either 5.78 EUR or 0.15 EUR. Subjects were asked to make ten decisions and pick one option in each of the ten rows. Options in the ten rows had equal payoffs, however, the probability of the high-payoff outcome gradually increased in steps of 10%, until it reached 100% for the tenth decision; correspondingly, probability of the low-payoff outcome has gradually decreased in steps of 10%. As in Holt and Laury [2002] a total number of safe choices was used to assess individual risk aversion. At the end of the session one row was determined randomly and each subject could play out her choice in this row for real.

---

[5] We have chosen the size of reward for a forecasting task such that it would not motivate subjects to engage in strategic manipulations of the market price, in their desire to maximize their reward for this task. Still we cannot completely rule out the possibility that it did influence behavior of some participants.

[6] This incentive scheme instead of a quadratic scoring rule was chosen for the sake of keeping the instructions simple (Haruvy et al. [2007]).

Option A		Option B	
1/10 of 3.00 EUR,	9/10 of 2.40 EUR	1/10 of 5.78 EUR,	9/10 of 0.15 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
3/10 of 3.00 EUR,	7/10 of 2.40 EUR	3/10 of 5.78 EUR,	7/10 of 0.15 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
5/10 of 3.00 EUR,	5/10 of 2.40 EUR	5/10 of 5.78 EUR,	5/10 of 0.15 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
7/10 of 3.00 EUR,	3/10 of 2.40 EUR	7/10 of 5.78 EUR,	3/10 of 0.15 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
9/10 of 3.00 EUR,	1/10 of 2.40 EUR	9/10 of 5.78 EUR,	1/10 of 0.15 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

Table 1: The Choice List

3 Results

Average Prices

In this section various summary statistics of the two types of the market are compared. Each session counts as one (independent) observation. Totally ten sessions were conducted, five high overconfidence and five low overconfidence markets.

Figure 1 shows that average prices[7] in high overconfidence markets are substantially higher than in low overconfidence markets. The average market price in low overconfidence markets was 33 ECUs (SD = 9.41) and 67 ECUs (SD = 16.02) in high overconfidence markets. This difference is significant (Mann-Whitney U = 0.00, $p < 0.01$, one-sided). The average fundamental value (FV) equals 19.20 ECU; depicted by the dotted line in Figure 1. Prices in both low overconfidence and high overconfidence markets are significantly higher than FV (Wilcoxon T = 1.89, $p < 0.05$, one-sided).

[7] Average prices per period averaged over 15 periods.

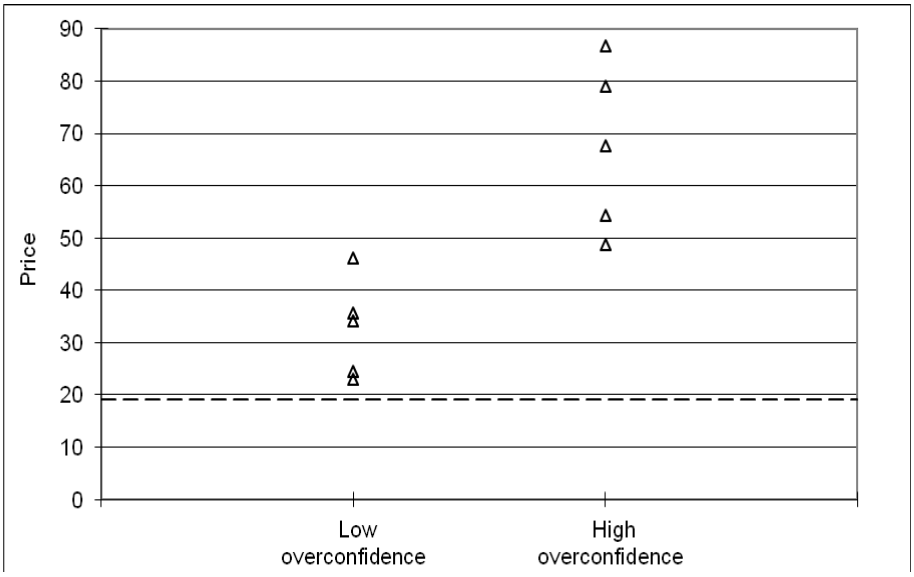


Figure 1: Average asset prices in both types of markets

Evolution of Average Prices

Figure 2 presents development of the aggregate average prices in low overconfidence and high overconfidence markets in the course of experiment. The dotted line indicates that FV diminishes by 2.4 in each of the 15 trading periods. For the graphs depicting average price development in each of the experimental markets separately see Appendix C.

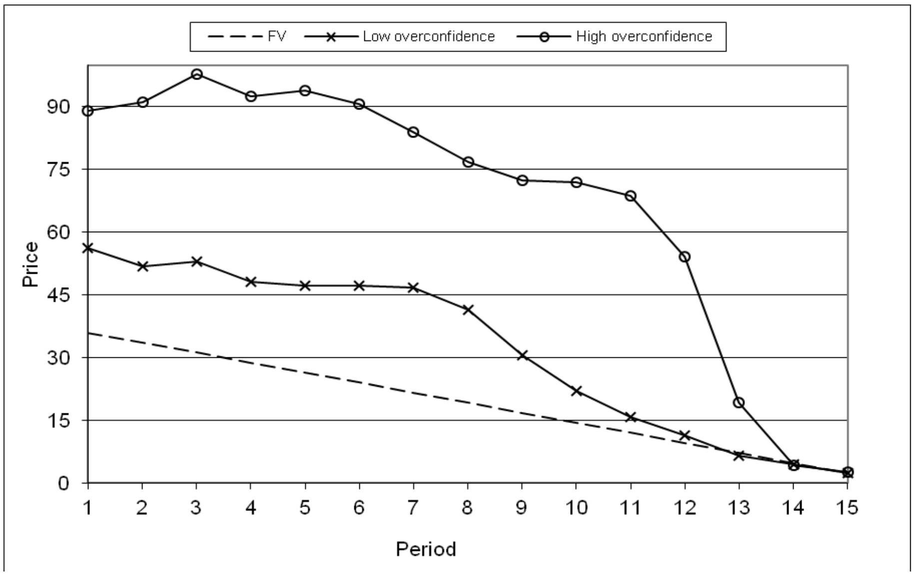


Figure 2: Development of the aggregate average market price

Visual data analysis suggests that aggregate prices deviate from FV in both types of the markets. However prices in low overconfidence markets deviate from FV to a smaller extent than prices in high overconfidence markets and tend to track FV more accurately. It can also be seen that in high overconfidence markets bubble and crash pattern is more pronounced than in low overconfidence markets, where no sudden drop of the aggregate market price to FV is observed.

Volatility

Figure 3 presents volatility in both types of markets, measured in terms standard deviation of prices[8]. A Mann-Whitney U test confirms that volatility in the high overconfidence markets is significantly higher than in the low overconfidence markets (Mann-Whitney U = 4.00, $p < 0.05$, one-sided). For both types of the market, Wilcoxon Signed Rank test enabled rejection of the null hypothesis that the volatility of prices was equal to the volatility of FV (SD = 10.73) in favor of the alternative hypothesis that volatility was higher (Wilcoxon T = 1.89, $p < 0.05$, one-sided).

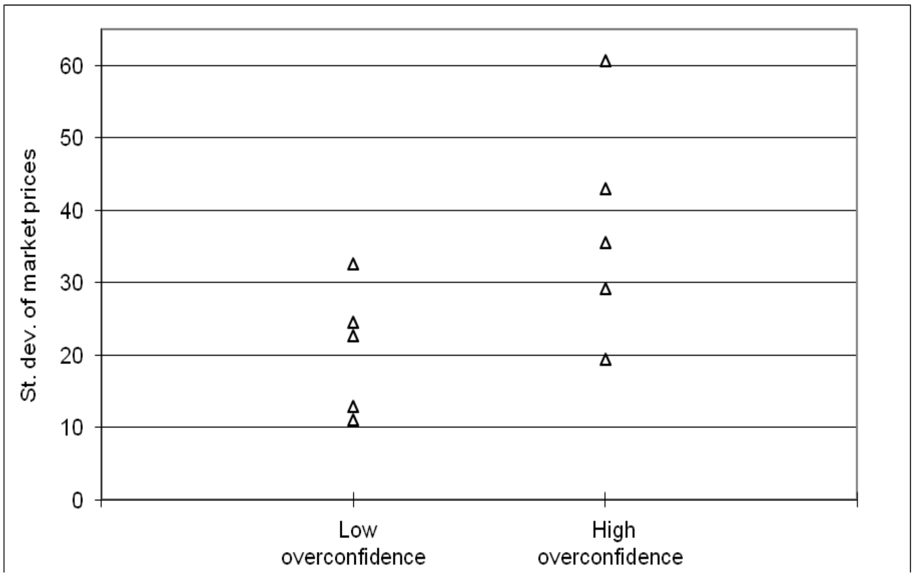


Figure 3: Volatility of asset prices in both types of markets

Trading Activity

According to the No-Trade Theorem by Milgrom and Stokey [1982] 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. Figure 4 shows that this result does not hold for our experimental markets, even though there was no private information at all in our design. Wilcoxon Signed Rank test of the hypothesis that turnover (calculated as number of assets traded in one period divided by the total number of assets, i.e. 18) equals zero is rejected for both markets in favor of the alternative hypothesis that turnover is significantly higher than zero (Wilcoxon T = 1.90, $p < 0.05$, one -sided).

[8] Standard deviation per period averaged over 15 periods.

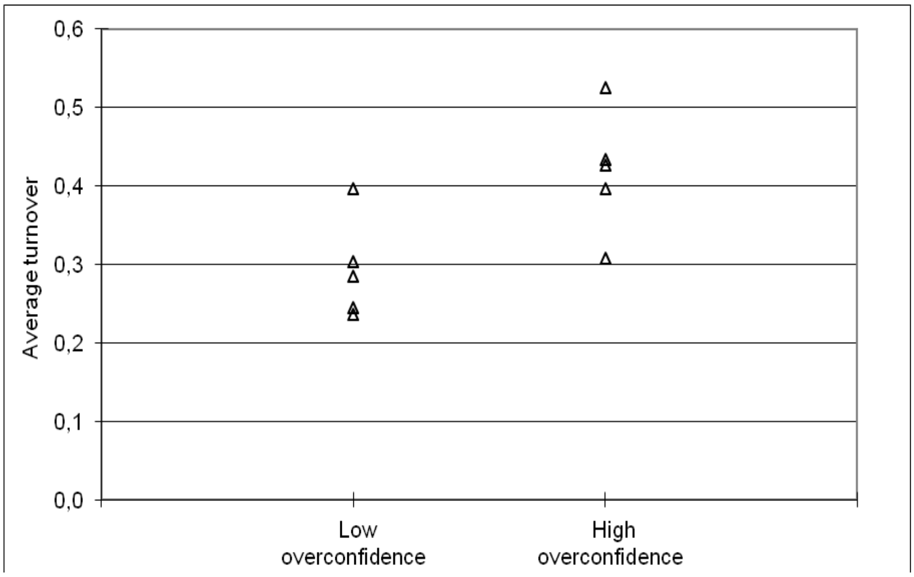


Figure 4: Average trading activity (turnover) in both types of markets

Trading activity in low overconfidence markets is lower than in high overconfidence ones: average market turnover in low overconfidence sessions is 28% (5 units of the asset) and 44% (8 units of the asset) in high overconfidence sessions. This difference is significant (Mann-Whitney U = 1.50, $p < 0.05$, one-sided). This result is in line with previous research that found overconfidence to be connected to higher trading volume (Odean [1999], Statman, Thorley, and Vorkink [2006], Deaves et al. [2009]).

Evolution of the joint average market turnover in five low overconfidence and five high overconfidence markets is shown in Appendix D. 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 an end-game effect.

Price Forecasting

Average forecasts for low overconfidence and high overconfidence markets are shown in Figure 5. It is obvious that forecasts in the high overconfidence markets (M = 69.80, SD = 16.75) are higher than in the low overconfidence markets (M = 32.20, SD = 8.32). According to a Mann-Whitney U test this difference is significant (Mann-Whitney U = 0.00, $p < 0.01$). Higher price forecasts in high overconfidence markets are not entirely driven by higher realized prices as already in the first period forecasts in high overconfidence markets (M = 90.20, SD = 40.24) exceed those in low overconfidence markets (M = 51.37, SD = 17.53) significantly (Mann-Whitney U = 3.00, $p < 0.05$, one-sided).

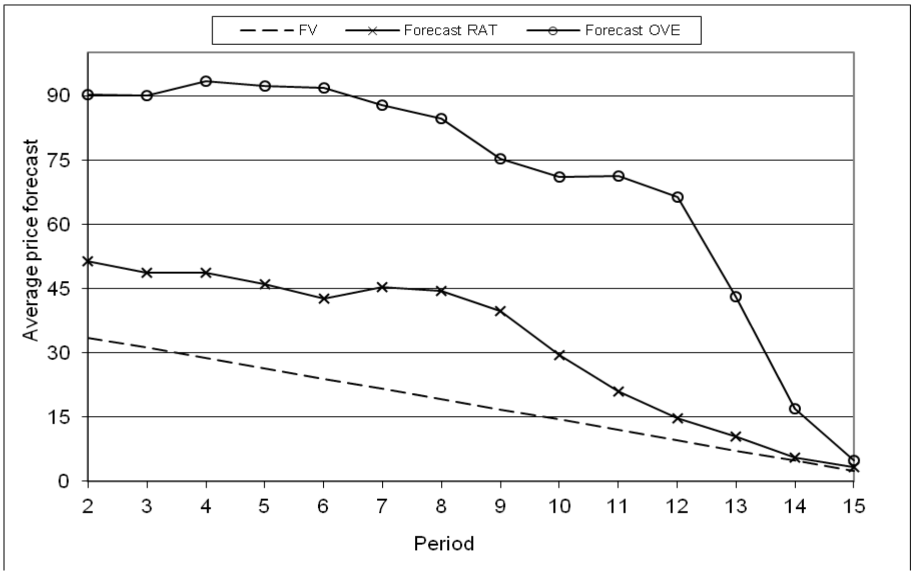


Figure 5: Price forecasting in both types of markets

Market bias score from the price forecasting task ($BS_{forecasting}$) was calculated for each session separately, based on the same methodology as in the general knowledge task. First, for each subject, the proportion of correct forecasts across all periods was subtracted from the mean confidence level across all periods. Second, individual bias scores were averaged across all participants in that market:

$$BS_{forecasting} = \frac{1}{M}\frac{1}{N}\sum_{j=1}^{M}\sum_{i=1}^{N}\left(c_{ij} - a_{ij}\right), \tag{3}$$

where M is the number of forecasting periods; N is the number of participants in the market; c_{ij} is the confidence in forecasting average price in period i of a participant j; a_{ij} is the accuracy in forecasting average price in period i of a participant j.

Market overconfidence measure obtained via overconfidence test (BS_{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 we can assume that both constructs measure the same phenomenon. This result also suggests that overconfidence is a robust phenomenon in our sample.

Figure 6 indicates that on average the bias score from the price forecasting task was higher in high overconfidence than in low overconfidence markets. On average overconfidence in price prediction task differed between the two types of market by 10 units ($BS_{forecasting}$ in low overconfidence markets: M = 50.08, SD = 8.96; $BS_{forecasting}$ in high overconfidence markets: M = 60.31, SD = 5.02). $BS_{forecasting}$ value in high overconfidence markets is significantly larger than $BS_{forecasting}$ in low overconfidence markets (Mann-Whitney U = 4.00, $p < 0.05$, one-sided).

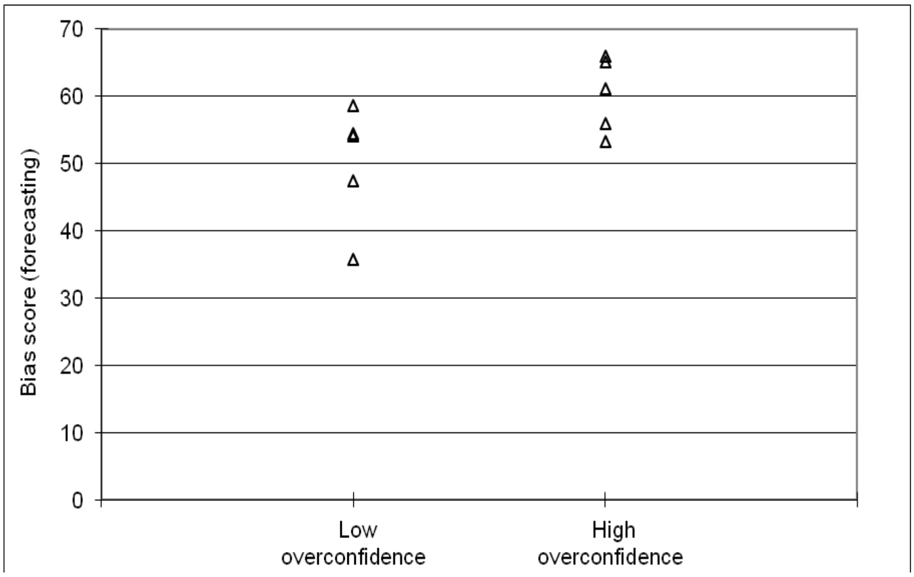
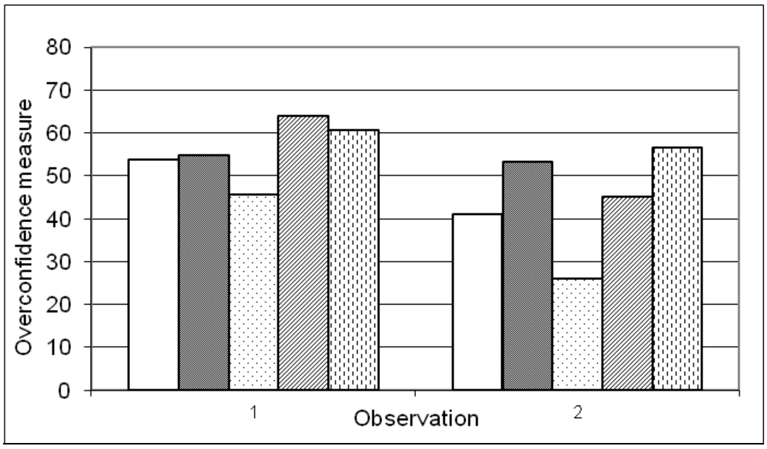


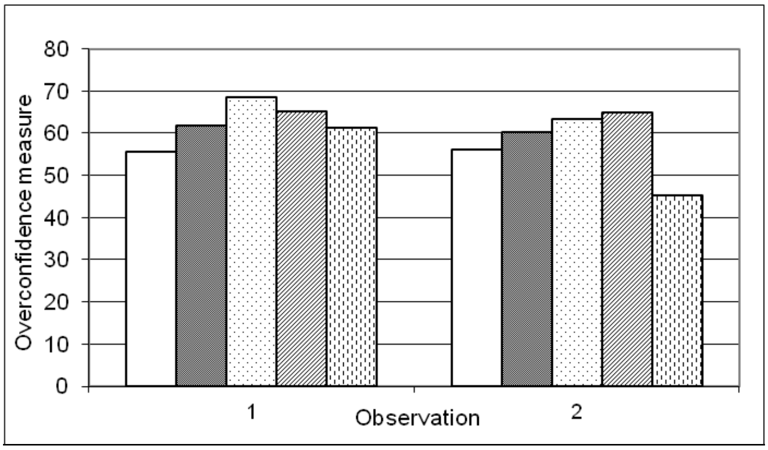
Figure 6: Average overconfidence in both types of markets

Risk Aversion

Experimental results present evidence that on average subjects were risk averse with 5.66 "safe" choices (SD = 1.82). Low overconfidence subjects took on average 5.81 safe choices (SD = 1.42), and high overconfidence subjects 5.50 safe choices (SD = 2.19). Statistical tests, conducted to check whether more overconfident subjects were also more risk loving, detected no significant difference between the two groups of players, neither in terms of their average number of "safe" choices (Mann-Whitney Z = 0.320, p = 0.749, two-sided), nor in their variation (Siegel-Tukey test Z = 0.47, p = 0.642, two-sided). Correlation coefficient between risk aversion, measured as the number of "safe" choices, and individual overconfidence, measured as general knowledge test bias score and the bias score from forecasting task, implies no linear relationship between them (BS_{test}: Spearman's Rho(30) = -0.095, p = 0.303, one-sided; $BS_{forecasting}$: Spearman's Rho(30) = 0.199, p = 0.137, one-sided).



a.



b.

Figure 7: Comparisons of BS(2-8) and BS(9-15): a. low overconfidence markets; b. high overconfidence markets

Evolution of the Bias Score

To investigate whether overconfidence is reduced by the end of the game, data of the price prediction task were divided into two time intervals of seven periods 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 7 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 last seven last periods. Wilcoxon Signed Ranks test confirms that BS(2-8) is significantly higher than BS(9-15) ($Z = -2.43$, $p < 0.01$, one-sided). This finding could serve as an explanation why bubbles mostly burst close to the end of the experiment.

Bubble Measures

From the previous analysis we obtained evidence, that although prices, volatility and turnover in low overconfidence (L.O.) markets are significantly below those in high overconfidence (H.O.) markets, they are still much higher than initially hypothesized. In other words, low overconfidence markets might also be prone to bubbles, but of a smaller magnitude. To analyze this issue, we calculate several measures of the magnitude of bubbles that were used by previous authors (e.g. Porter and Smith [1995], Noussair and Tucker [2006], Dufwenberg et al. [2005], Stöckl et al., [2010]). These measures are: Hassel-R^2, Price Amplitude (APL), Normalized Absolute Deviation (NAD), Normalized (Average) Price Deviation (NPD), Velocity, Realtive Absolute Deviation (RAD), and Relative Deviation (RD)[9]. Table 2 reports the values of the measures by session and market type.

All bubble measures calculated for low overconfidence sessions are statistically significantly smaller than the ones obtained from high overconfidence sessions (see Appendix E.I). These results demonstrate that although bubbles in low overconfidence markets are not completely eliminated, they are less severe in comparison to the bubbles in high overconfidence markets. Correlation coefficients between bubble measures and the two measures of overconfidence (BS_{test} and $BS_{forecating}$) are large and significant (see Appendix E.II), implying that the size of the bubble measures increases with the increase in market overconfidence.[10]

[9] *The Hassel-R^2* (Haessel, 1978) measures goodness-of-fit between average market price per period and fundamental asset value. Hassel-R^2 can take values in the interval [0, 1]; it converges to 1 if trading prices converge to fundamental values. *The Normalized Price Deviation* and the *Normalized Absolute Deviation* show whether stocks were overpriced or underpriced relative to the fundamental value. NPD (NAD) is calculated by summing up all (absolute) deviations of market contract prices from fundamental value and dividing this sum by the total number of stocks in the market. *The Price Amplitude* is the maximum value of the shift of average contract price from the fundamental value for an experimental session. Higher price amplitudes imply greater bubbles, and larger swings in the market price of the asset relative to fundamental value. *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 measure is connected to the volume of trade: the higher is the velocity, the higher is the volume of trade. *Relative Absolute Deviation* measures the average level of asset mispricing and *Absolute Deviation* of asset overvaluation in the market

[10] This is in line with Ackert et al. [2009] who found irrationality (i.e. probability judgment errors) to be correlated with magnitude and frequency of price bubbles.

Session	Market type	Hassel-R^2	NPD	NAD	APL	Velo-city	RAD	RD
1	H.O.	0.58	9.14	9.31	1.69	4.61	0.50	0.50
2	H.O.	0.54	24.91	24.94	2.25	5.94	0.72	0.72
3	H.O.	0.41	38.26	38.38	2.87	7.89	0.44	0.43
4	H.O.	0.29	13.01	13.20	1.32	6.50	0.47	0.46
5	H.O.	0.89	25.87	25.96	3.33	6.39	0.30	0.30
6	L.O.	0.94	0.98	1.02	0.30	3.67	0.05	0.05
7	L.O.	0.91	5.75	6.13	1.09	4.56	0.10	0.09
8	L.O.	0.57	1.77	3.41	0.67	5.89	0.08	0.06
9	L.O.	0.94	9.59	9.92	1.67	4.28	0.37	0.36
10	L.O.	0.81	3.78	4.10	1.15	3.56	0.36	0.34

Table 2: Bubble Measures in the Single Sessions

Table 3 presents average values of Normalized Absolute Deviation[11] and the Amplitude from our experimental treatments alongside with the values from the paper of SSW [1988]. On average values from the low overconfidence markets treatment lie below values obtained by SSW [1988]; thus there is evidence of the smaller deviations from the fundamental value in low overconfidence markets.

	NAD	Velocity	Amplitude
High overconfidence markets	2.24	6.27	2.29
Low overconfidence markets	0.49	4.39	0.98
Smith, Suchanek, and Williams [1988]	5.68	4.55	1.24

Table 3: Comparison of Average Bubble Measures

4 Conclusions

In this paper results of an experiment, designed to investigate the role of market overconfidence for the occurrence of stock-prices bubbles have been reported. The design of

[11] For the comparison of NAD measure from our experiment to those of SSW [1988], it has to be divided by ten. The reason is that, their study used an expected dividend value of 0.24 ECU; in our experiment it is 2.40 ECU.

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. In the experiment two types of markets are constructed: low overconfidence and high overconfidence market.

Our results refine differences between market outcomes in the experimental treatments and suggest the existence of the connection between market overconfidence and market outcomes. Although all traders in our study have identical information we observe that trading activity in low overconfidence markets is significantly above zero; however it is significantly below trading activity in the high overconfidence markets. Our results show very clearly that higher market overconfidence is accompanied by the higher average market prices and larger deviations of the prices from fundamental value. Although average prices in both types of markets significantly exceed the fundamental value, prices in low overconfidence markets tend to track the fundamental asset value more accurately than prices in high overconfidence markets, and are significantly lower. Moreover, bubble and crash patterns were observed in the aggregate price in high overconfidence markets, whereas in low overconfidence markets no sudden drop of the aggregate market price to the fundamental value occurred. Volatility of the prices and trade volume proved to be significantly lower in low overconfidence markets.

We analyze two possible channels through which overconfidence could be projected onto asset prices in our markets. First, more overconfident subjects may be less risk averse and value possession of risky assets higher. More overconfident subjects may overestimate future prices and the probability to sell assets with profit in later rounds. While we do not detect significant differences in the average degree of risk aversion on both markets, high overconfidence subjects have substantially higher price forecasts than low overconfidence ones, suggesting that the second explanation is the driving force behind high prices in the high overconfidence markets.

Several studies showed that professional investors may be particularly prone to overconfidence. For instance Glaser, Langer and Weber [2005, 2007] demonstrated that professional traders are more overconfident than students. Menkhoff, Schmeling and Schmidt [2010] found investment advisors to be substantially miscalibrated, despite their high degree of professionalism. Barber and Odean [2001] suggest that overconfidence motivates mutual fund managers to trade excessively. Also, empirical evidence exists that investment activity of the firm is correlated with CEO overconfidence (cf. Liu and Taffler [2008], Sautner and Weber [2009]). Given these observations, our experimental results may have serious implications for financial markets in the real world.

References

Ackert, L. F., Charupat, N., Deaves, R., and Kluger, B. D. “Probability Judgment Error and Speculation in Laboratory Asset Market Bubbles.” *Journal of Financial and Quantitative Analysis*, 44, (2009), pp. 719–744.

Barber, B., and Odean, T. “Boys Will Be Boys: Gender, Overconfidence and Common Stocks Investments.” *Quarterly Journal of Economics*, 116, (2001), pp. 261–292.

Benos, A. “Aggressiveness and Survival of Overconfident Traders.” *Journal of Financial Markets*, 1, (1998), pp. 353–383.

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. “Judgmental overconfidence, self-monitoring and trading performance in an experimental financial market.” *Review of Economic Studies*, 72, (2005), pp. 287–312.

Caballé, J., and Sákovics, J. “Speculating Against an Overconfident Market.” *Journal of Financial Markets*, 6, (2003), pp. 199–225.

Cohen, J. *Statistical Power Analysis for the Behavioural Sciences (2nd ed.)*. New Jersey: Lawrence Erlbaum, 1988.

Daniel, K., Hirshleifer, D., and Subrahmanyam A. “Investor Psychology and Security Market Under-And Overreactions.” *Journal of Finance*, 53, (1998), pp. 1839–1885.

De Bondt, W. F. M., and Thaler R. “Does the Stock Market Overreact?” *The Journal of Finance*, 40, (1985), pp. 793–805.

Dufwenberg, M., Lindqvist, T., and Moore, E. “Bubbles and Experience: An Experiment”. *The American Economic Review*, 95, (2005), pp. 1731–1737.

Fischbacher, U. “z-Tree: Zurich Toolbox for Ready-made Economic Experiments.” *Experimental Economics*, 10, (2007), pp. 171–178.

Gigerenzer, G., Hoffrage, U., and Kleinbölting, H. “Probabilistic Mental Models: A Brunswikian Theory of Confidence.” *Psychological Review*, 98, (1991), pp. 506–528.

Glaser, M., M. Nöth, and Weber, M. “Behavioral Finance.” In D. Koehler, and N. Harvey, eds., *Blackwell Handbook of Judgment and Decision Making*. Malden Mass.: Blackwell, 2004, pp. 527–546.

Glaser, M., Langer, T., and Weber, M. “Overconfidence of professionals and lay men: Individual differences within and between tasks?” Working paper, University of Mannheim, April 2005.

Glaser, M., and Weber, M. “Overconfidence and trading volume.” *The Geneva Risk and Insurance Review*, 32, (2007), pp. 1–36.

Griffin, D., and Brenner, L. “Perspectives on probability judgment calibration.” In D. Koehler, and N. Harvey, eds., *Blackwell Handbook of Judgment and Decision Making*. Malden Mass.: Blackwell, 2004, pp. 177–199.

Haruvy, E., Lahav, Y., and Noussair, C. “Traders’ Expectations in Asset Markets: Experimental Evidence.” *American Economic Review*, 97, pp. 1901–1920.

Kirchler, E., and Maciejovsky, B. “Simultaneous Over- and Underconfidence: Evidence from Experimental Asset Markets.” *Journal of Risk and Uncertainty*, 25, (2002), pp. 65–85.

Kirchler, M., Huber, J., and Stöckl, T. “Thar She Bursts – Reducing Confusion Reduces Bubbles.” *American Economic Review*, 102, (2012), pp. 865–883.

Klayman, J., Soll., J. B., Gonzáles-Vallejo, C., and Barlas, S. “Overconfidence: It Depends on How, What, and Whom You Ask.” *Organizational Behavior and Human Decision Processes*, 79, (1999), pp. 216–247.

Langer, E. “The Illusion of Control.” *Journal of Personality and Social Psychology*, 32, (1975), pp. 311–328.

Liu, Y., and Taffler, R. “Damned out of their own mouths: CEO overconfidence in M&A decision making and its impact on firm performance.” Working paper, University of Edinburgh, 2008.

Lichtenstein S., Fischhoff B., and Phillips L.D. “Calibration of Probabilities: The State of the Art to 1980.” In D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment under uncertainty: heuristics and biases*, New York: Cambridge University Press, 1982, pp 306–334.

Michailova, J., and Katter, J.K.Q. “Quantifying Overconfidence in Experimental Finance”. *International Journal of Behavioural Accounting and Finance*, 4, (2014), pp. 351–367.

Milgrom, R., and Stokey N. “Information, Trade and Common Knowledge.” *Journal of Economic Theory*, 26, (1982), pp. 17–27.

Menkhoff, L., Schmeling, M., and Schmidt, U. “Overconfidence, Experience, and Professionalism: An Experimental Study.” *Journal of Economic Behavior and Organization*, 86, (2013), pp. 92–101.

Moore, D.A, and Healy P.J. “The Trouble with Overconfidence.” *Psychological Review*, 115, (2008), pp. 502–517.

Noussair, C., and Tucker, S. “Futures Markets and Bubble Formation in Experimental Asset Markets.” *Pacific Economic Review*, 11, (2006), pp. 167–184.

Odean, T. “Do Investors Trade Too Much?” *American Economic Review*, 89, (1999), pp. 1278–1298.

Porter, D., and Smith, V. “Futures Contracting and Dividend Uncertainty in Experimental Asset Markets.” *Journal of Business*, 68, (1995), pp. 509–541.

Russo, J. E., and Schoemaker, P. J. H. “Managing Overconfidence.” *Sloan Management Review*, 33, (1992), pp. 7–17.

Sautner, Z., and Weber, M. “How do managers behave in stock option plans? Clinical evidence from exercise and survey data.” *Journal of Financial Research*, 32, (2009), pp. 123–155.

Scheinkman, J. A., and Xiong, W. “Overconfidence and Speculative Bubbles.” *Journal of Political Economy*, 111, (2003), pp. 1183–1219.

Shiller, R. J. *Irrational Exuberance*. Princeton, N.J.: Princeton Univ. Press, 2000.

Smith, V. L., Suchanek G. L., and Williams A. W. “Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets.” *Econometric*a, 56, (1988), pp. 1119–1151.

Statman, M., Thorley, S., and Vorkink, K. “Investor Overconfidence and Trading Volume.” *Review of Financial Studies*, 19, (2006), pp. 1531–1565.

Stöckl, T., Huber, J., and Kirchler, M. “Bubble Measures In Experimental Asset Markets.” Experimental Economics, 13, (2010), pp. 284–298.

Taylor, S. E., and Brown, J.D. “Illusion and Well-Being: A Social Psychological Perspective on Mental Health.” *Psychological Bulletin*, 103, (1988), pp. 193–210.

APPENDIX A: BIAS SCORES OF THE EXPERIMENTAL SUBGROUPS (in brackets p-values of group mean differences)

Database					
OBS	Group	Mean	SD	Min	Max
201	All	11.78	10.57	-11.33	43.50
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)	--	--	--
Experimental Sessions					
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	High overconfidence	21.33	8.26	10.17	43.50
30	Low overconfidence	1.06	4.03	-5.89	6.78
	Male vs. female diff.	2.13 (0.81)	--	--	--
	High ove. vs. low ove. diff.	20.27 (0.00)	--	--	--
High ove. markets	Male vs. female diff.	-0.65 (0.64)	--	--	--
Low ove. markets	Male vs. female diff.	0.68 (0.76)	--	--	--

APPENDIX B: TRANSLATION OF INSTRUCTIONS

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*1/4+0.8*1/4+2.8*1/4+6*1/4)

Before each trading period information on potential income from holding your assets till the end of the experiment (15^{th} 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.

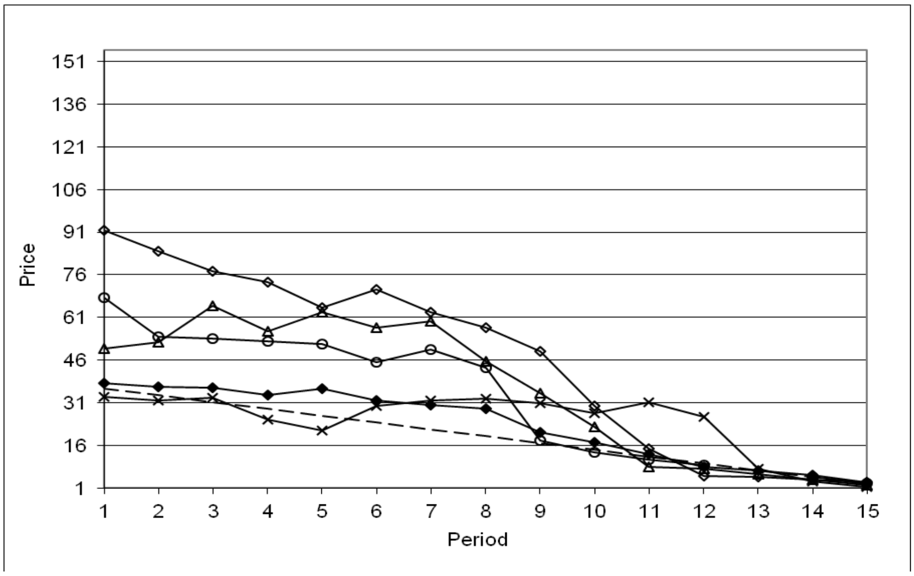
In the gap marked "Confidence level" you have to write how confident you are that your price forecast is correct! You can use any number between 0% and 100% to express your confidence, that your forecast is correct. Thus 0% means that you completely do not believe that your forecast can be true, and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

This is the end of the instructions!

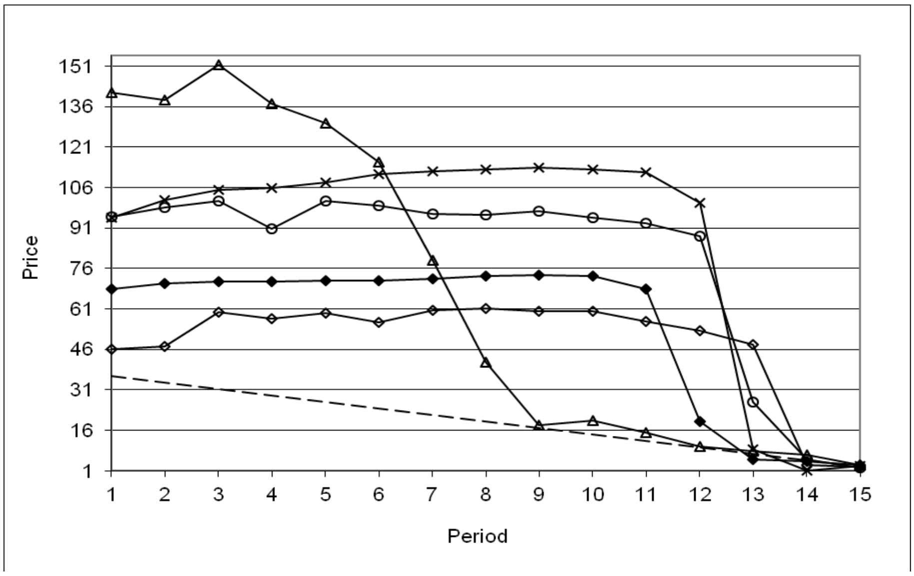
If you have a question that was not fully answered by the instructions please raise your hand and ask the experiment monitor before proceeding.

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

APPENDIX C: Development of the average market price (a. Low overconfidence markets, b. High overconfidence markets; dotted line represents FV)

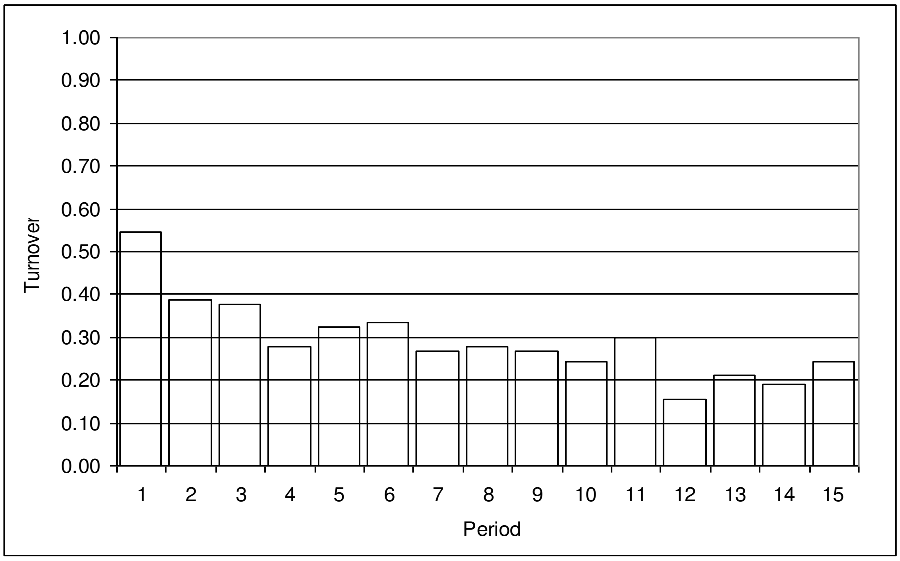


a.

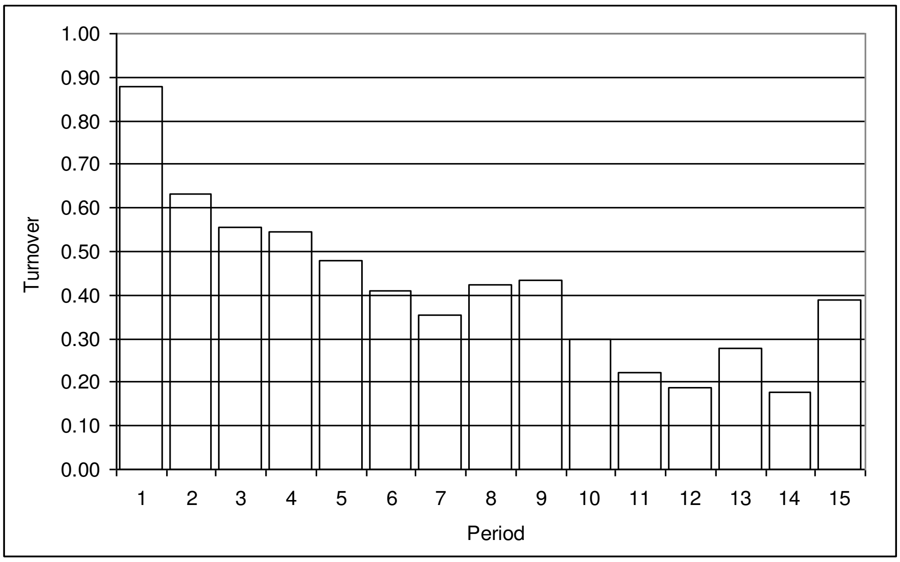


b.

APPENDIX D: JOINT AVERAGE TURNOVER DEVELOPMENT (a. Low overconfidence markets, b. High overconfidence markets)



a.



b.

APPENDIX E

E.I: TEST OF THE DIFFERENCES BETWEEN BUBBLE MEASURES IN TWO TREATMENTS

Hassel R^2	Mann-Whitney U = 3.00 ($p < 0.05$, one-sided)
NPD	Mann-Whitney U = 1.00 ($p < 0.01$, one-sided)
NAD	Mann-Whitney U = 1.00 ($p < 0.01$, one-sided)
Velocity	Mann-Whitney U = 1.00 ($p < 0.01$, one-sided)
Amplitude	Mann-Whitney U = 1.00 ($p < 0.01$, one-sided)
RAD	Mann-Whitney U = 2.00 ($p < 0.05$, one-sided)
RD	Mann-Whitney U = 2.00 ($p < 0.05$, one-sided)

E.II: Spearman's Rho Correlation Coefficient between Bias Scores and Bubble Measures

	BS_{test}	$BS_{forecasting}$
Hassel R^2	-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 ($p < 0.05$, one-sided)
Amplitude	0.661 ($p < 0.05$, one-sided)	0.515 ($p < 0.1$, one-sided)
RAD	0.636 ($p < 0.05$, one-sided)	0.661 ($p < 0.05$, one-sided)
RD	0.636 ($p < 0.05$, one-sided)	0.661 ($p < 0.05$, one-sided)