Overreaction in Stock Forecasts and Prices

Alen Nosić and Martin Weber *

July 21, 2009

Abstract

We study the degree of individual and aggregate market overreaction in a dynamic experimental auction market. In 13 sessions with overall 101 students we find overreaction to new information both in stock price forecasts and transaction prices. Interestingly, market forces do not seem to help in lowering overreaction to new information in our setting. Moreover, we illustrate that subjects are not able to learn from their previous failures and thus do not correct their erroneous beliefs. Hence, overreaction in our setting remains on a stable level although subjects can at least in theory learn from other market participants or from outcome feedback. Lastly, we find first experimental evidence for a positive relation between differences of opinion and trading volume in a continuous auction market with several market participants.

Keywords: Overreaction, Underreaction, Market Experiment, Differences of Opinion, Trading Volume

JEL Classification Code: G1

1 Introduction

Behavioral finance shows that individual biases such as the disposition effect (Odean (1998a) and Weber and Camerer (1998)), hindsight bias (Biais and Weber (2009)) or overconfidence (Camerer and Lovallo (1999) and Biais et al. (2005)) can affect individual decision making. In addition, both theoretical and experimental behavioral studies argue that markets are not always fully efficient and that market forces are not fully able to correct for individual biases (see e.g. Camerer (1987), Jegadeesh and Titman (1995), or Scheinkman and Xiong (2003)). However, it is still ambiguous whether and why individual biases prevail in market outcomes and if rational explanations are more capable to explain some market anomalies. Some studies for example show that individual biases vanish totally or are significantly reduced by market forces whereas other studies illustrate that individual biases remain fairly stable or even get more pronounced on an aggregate level.

On the one hand, analyzing abstract Bayesian updating tasks Camerer (1987) and Ganguly et al. (1994 and 2000) show that probability judgment errors or biases can indeed persist in market settings. However, they show that the bias on an aggregate level is reduced. In a similar vein, Camerer et al. (1989) analyzing the hindsight bias, Kluger and Wyatt (2000) studying judgment errors in the Monty Hall problem, and Sonnemann et al. (2008) analyzing partition-dependence, show that market experience mitigates the respective bias but is not able to eliminate it completely. On the other hand, Gillette et al. (1999), Bloomfield et al. (2000), and Nelson et al. (2001) find that underreaction to signals in a coin-spin scenario shows the same extent in markets and on an individual level. Similarly, Seybert and Bloomfield (2009) find hardly any evidence for a wishful thinking bias on an individual level but strong evidence for wishful betting on an aggregate market level.

In addition, van Boven et al. (2003) and Weber and Welfens (2007) show that a bias gets smaller over repeated interactions of the same commodity or in the course of a trading round but that this learning does not generalize to interactions with a new asset or commodity. Budescu and Maciejovsky (2005) conclude that "expecting biases to disappear completely or, alternatively, to always persist are overly simplistic positions". Overall, findings in the literature indicate that the question whether markets can correct for in-

dividual biases depends on the structure of the market, the task on hand, the sort of feedback that subjects receive or the individual bias that is analyzed.

The main goal of this paper is to analyze whether individual over- or underreaction to changes in the fundamental value of a stock, i.e. to new information, affects market outcomes in a setting similar to the one empirically analyzed by Thomas and Zhang (2008). Thomas and Zhang (2008) study market reactions to earnings announcements by considering pairs of stocks from the same industry which announce earnings sequentially. More precisely, they show that the market price of a firm that has not yet announced earnings moves too far upward (downward) in reaction to good (bad) earnings reports of an early-announcing peer and is corrected when the late-announcing firm's earnings are subsequently revealed. Thus, their findings suggest that stock prices for the late announcers overreact to the information transfer from the early-announcing peer. Overreaction (underreaction) in this context means that subjects put too much (little) weight on new information. However, using empirical data it is not possible to rule out other explanations for misreaction or to analyze in more detail, how individuals react to new information about the fundamental value of a stock and whether this individual misreaction affects market variables. Misreaction in this context is simply defined as the absolute level of over- or underreaction.

Since the empirical and experimental evidence on how individual misreaction affects market outcomes is scarce, we apply the individual level approach of Biais et al. (2008) that analyzes how individuals process new information to an experimental trading market. The main features of our design are as follows: first, we give subjects the stock price charts of two related stocks (G and H) for the last six months and ask them to estimate the price of one of the two stocks (H) in six months. After having provided a best estimate for stock H subjects are able to trade this stock for two minutes in a single-unit open-book double auction market. After two minutes of trading the subjects receive additional information about the other stock (G) and are asked to update their estimates regarding the price of stock H in six months. Subsequently, subjects get the possibility to trade stock H in a single-unit open-book double auction market again for two minutes. Overall, each subject trades in 8 of these experimental rounds, consisting of two trading periods of 120 seconds and two estimation tasks.

This experimental design allows us to relate individual level overreaction to market level overreaction in an almost realistic setting, similar to the one in Thomas and Zhang (2008). However, in contrast to their empirical approach we are able to explicitly calculate the rational benchmark, and thus can rule out that risk or market microstructure effects drive results on overreaction in markets. Moreover, using this clean design we are able to quantify the exact degree of overreaction both on an individual and market level, and thus can compare findings in the individual-level-study by Biais et al. (2008) with the aggregate level study by Thomas and Zhang (2008).

Previous empirical and experimental evidence on the relation between individual level misreaction and its effect on market parameters is scarce. On the one hand, most experimental studies investigate the level of overreaction or underreaction on an individual level (see e.g. Kahneman and Tversky (1973), Griffin and Tversky (1992), Bloomfield and Hales (2002), Offerman and Sonnemans (2004), and Biais et al. (2008)). On the other hand, empirical studies are mostly providing evidence for aggregate market overreaction (see e.g. DeBondt and Thaler (1985, 1987, and 1990)), Sorescu and Subrahmanyam (2006), and Thomas and Zhang (2008)). Evidence on the question how individual misreaction to new information translates into market outcomes is scarce (see for notable exceptions Gillette et al. (1999), Bloomfield et al. (2000), and Nelson et al. (2001)) and relies almost exclusively on the classical Griffin and Tversky (1992) coin-spin scenario where subjects need to indicate if a coin is heads- or tail-biased.

Our main findings can be summarized as follows: we observe strong and persistent overreaction for individual estimates as well as for market prices following both - good and bad news. The level of overreaction in estimates is in the same range as in the individual experiment by Biais et al. (2008). Interestingly, aggregate overreaction in transaction prices is not substantially lower. Thus, market forces are not able to correct for individual biases which is in line with experimental findings in Gillette et al. (1999) who show that underreaction in their experiment is even slightly higher in markets than on an individual level. Moreover, our finding that overreaction does prevail in markets is also consistent with behavioral models arguing that individual biases affect market outcomes (Odean (1998b), Sorescu and Subrahmanyam (2006), and Biais et al. (2005)). In addition, we provide evidence that learning effects within a 120-second trading period and learning ef-

fects over the course of the experiment are hardly observable. Lastly, up to our knowledge our study is the first to find experimental evidence for the theoretically proposed positive relation between differences of opinion and trading volume (Varian (1989), Harris and Raviv (1993), and Kandel and Pearson (1995)) in experimental markets with more than two traders.

In the next section we review the related literature on misreaction and their impact on market factors and form our hypotheses. Section 3 presents our experimental design and procedure as well as some descriptive statistics. The main results of our study are reported in section 4. Section 5 concludes with a short discussion of our findings and an outlook for future research.

2 Related Literature and Hypotheses

2.1 Related Literature

The study by Biais et al. (2008) offers direct experimental evidence on the relationship between information signals, beliefs, and financial decisions. Using a novel experimental design in which subjects are asked to estimate the future price of an asset, incorporating the stock price development of a related asset as information signal, the study has three main findings: first, in this experimental environment a substantial level of overreaction seems to exist which is in line with findings in Griffin and Tversky (1992) on an individual level and Thomas and Zhang (2008) on an aggregate level. Second, analyzing the relation between individual overreaction and psychological biases the authors show that more overconfident subjects tend to overreact more heavily. Third, the authors show that individual overreaction has an impact on financial variables such as portfolio risk and portfolio efficiency.

However, it is not clear whether overreaction on an individual level translates to overreaction in a market setting. Proponents of rational approaches often argue that just a few rational subjects are sufficient to make market outcomes rational, that random mistakes cancel out or that in the long run less rational subjects learn from more rational subjects in the market (for further comments on this debate see Camerer (1987)). In the following,

we will first review the theoretical literature on individual and market overreaction and then illustrate existing empirical and experimental evidence on this issue.[1]

Theoretical models

Using various behavioral biases such as overconfidence or hindsight many behavioral models argue that individual overreaction has a substantial impact on market variables.[2] Amongst others, Daniel et al. (1998 and 2001), Odean (1998b), and Fischer and Verrecchia (1999) model financial markets with overconfident individuals. These overconfident individuals overweight the precision and overestimate the quality of a private signal that they receive. This directly results in an individual overreaction to the signal. Consequences of this individual overreaction are the wrong assessment of means, a more aggressive trading behavior, and a poor portfolio diversification. How exactly this individual overreaction affects market variables depends on the information structure of the market and on the proportion of overconfident investors in the market. However, almost all studies show that individual overreaction of some investors results in aggregate, market overreaction and has substantial effects on market demand and market prices.

In a similar vein, Biais and Weber (2007) theoretically show that hindsight biased agents overreact to new information and put too much weight on a private signal. The intuition behind is that hindsight biased agents incorrectly remember their prior expectations, and thus overweight the informational content of new information. In their CAPM-like framework individual overreaction of hindsight biased agents affects equilibrium prices and holdings.

A third strand of literature tries to reconcile two patterns that seem contradictory: the representativeness heuristic (Kahneman and Tversky, 1973) and the conservatism bias (Edwards, 1968). Griffin and Tversky (1992) argue that subjects focus too much on the strength of a new information, i.e. the degree to which it is favorable, and not enough on its weight, i.e. its statistical reliability. They use the evaluation of recommendation letters

---

[1] We have argued in section 1 that the question whether individual biases can be corrected by market forces seems to depend amongst others on the bias that is analyzed. Since we analyze over- or underreaction to new information we will only illustrate studies analyzing this sort of bias.

[2] In addition, Grinblatt and Han (2005) show that individual underreaction caused by behaviorally affected traders has also an influence on market prices in the equilibrium

as an example. Here the strength refers to how positive or warm the letter's content is and the weight refers to the credibility and knowledge of the writer. They argue that the representativeness heuristic and hence overreaction to new information prevails in situations in which subjects receive new information which is characterized by high strength and low weight. In the recommendation letter example this corresponds to a very nice recommendation letter from a person with low credibility. Barberis et al. (1998) and Sorescu and Subrahmanyam (2006) develop theoretical models that extend these findings and implications to financial markets. In line with the proposition in Griffin and Tversky (1992) both studies propose that individual overreaction is present in situations that are characterized by high strength and low weight and affects aggregate market prices and overall demand.

Empirical and experimental evidence

Although most behavioral models argue that individual overreaction automatically impacts market variables and does not cancel out in the aggregate, the experimental and empirical evidence on this issue is not unequivocal. Most empirical and experimental studies analyze either the level of overreaction on an individual level or on an aggregate level but not simultaneously. In the following we will first review selected studies analyzing individual overreaction before we document findings in the literature on aggregate market overreaction. In the end we will present some of the very few studies analyzing both individual and aggregate market overreaction.[3]

In one of the first experimental studies on overreaction Kahneman and Tversky (1973) show that subjects tend to put too much weight on meaningless and practically irrelevant information. Another strand of the literature uses the so called coin-spin design to detect over- or underreaction on an individual level (see Griffin and Tversky, 1992). In this design subjects know that the coin that is going to be spun is either heads- or tails-biased with a prior probability of 0.5. After having observed a specific number of spins subjects are asked to report an updated probability conditional on the observed signal which is simply the number of heads and tails in the conducted spins. Findings in Griffin and Tversky (1992), Offerman and Sonnemans (2004), and Massey and Wu (2005) confirm that individual

---

[3]Since the study of Biais et al. (2008) has shown that overreaction should be prevalent in our experimental setting as the signal seems to be of relatively low weight and high strength we will only list a few exemplary studies finding empirical or experimental evidence for underreaction: Abarbanell and Bernard (1992), Hong et al. (2000), and Weber and Welfens (2007).

overreaction is present if subjects receive information with relatively high strength and low weight. A different kind of test of individual overreaction to new information has been carried out by Bloomfield and Hales (2002). They test the predictions in the theoretical model of Barberis et al. (1998) in a simple experimental environment in which subjects have to predict the next step of a random walk. They find substantial levels of individual overreaction in this setting.

The overreaction phenomenon on the aggregate market level has also been amply documented in the empirical literature. DeBondt and Thaler (1985 and 1987) find that past winners tend to be future losers and vice versa which they interpret as evidence for overreaction. Analyzing the price reaction to analysts' forecast revisions Sorescu and Subrahmanyam (2006) find evidence for the strength and weight hypothesis by Griffin and Tversky (1992). Using an analyst's ability and experience as a proxy for the weight of a signal and the dramatic nature of an event, i.e. the level of a down- or upgrade, as a proxy for the strength of a signal they test the hypotheses in Griffin and Tversky (1992). Consistent with their hypothesis, they show that for signals with relatively high strength and low weight market prices tend to overreact. Their results imply aggregate overreaction for large down- or upgrades (high strength) by inexperienced analysts from investment banks with a relatively bad reputation (low weight). The study by Thomas and Zhang (2008) which resembles our experimental design the best also finds evidence for overreaction on a market level. Analyzing subsequent earnings announcements by different firms from the same industry they document that both the price of an announcing firm and the price of a non-announcing firm from the same industry move in the same direction. However, this price response of the non-announcing firm is negatively related to its price response when it subsequently announces earnings. This result indicates that prices for subsequent announcers overreact to an early announcer's earnings and are corrected later on.

As mentioned above the empirical and experimental findings on overreaction on both - individual and aggregate level - are scarce. Using a security markets task that is closely related to the previously described coin-spin scenarios Bloomfield et al. (2000) and Nelson et al. (2001) document that individual misreaction, i.e. misreaction in stock forecasts, also translates into aggregate misreaction, i.e. misreaction in prices. However, their analyses show that underreaction in prices and forecasts is approximately the same whereas

overreaction is significantly larger in stock forecasts than in prices. Both studies use a clearinghouse market where all three individuals' in a market have to choose a linear demand schedule by setting a reservation price and a slope. Market clearing prices are determined by a computer and set such that demand equals supply in the three person economy.

Moreover, Hommes et al. (2005) find evidence for individual and aggregate overreaction in an experimental prediction market. In their setting market prices are generated by an asset pricing model with heterogenous beliefs. More specifically, in this design a computer determines market prices by taking the average beliefs of all market participants and adding some extra noise term. The authors find that in 8 out of 10 markets individual overreaction results in aggregate overreaction.

In another experimental asset market, Gillette et al. (1999) give participants public information about the liquidating dividend of an asset in every third trading period. They use both double-continuous auction markets and call markets with trading periods of 120 seconds. The trading structure of their market closely resembles ours. Their main results show that underreaction in forecasts is even larger than underreaction in market prices indicating that the standard argument that individual mistakes will cancel out and individual misreaction will be corrected by market forces does not need to apply to these kind of asset markets.

2.2 Hypotheses

Biais et al. (2008) analyze a similar setting in a static environment with no trading market and no feedback. Consistent with their findings, we expect that after observing a good signal subjects state expectations that are higher than the fundamental value and after observing a bad signal expectations that are lower than the fundamental value. As they put too much weight on these signals subjects are expected to overreact in their stock forecasts.

Moreover, we hypothesize that the results with regard to market prices should resemble empirical findings in Thomas and Zhang (2008). They analyze aggregate overreaction in prices in a dynamic empirical setting that is very similar to ours. Their main finding is

that stock prices for a late announcing firm overreact to the information signal inherent in the early announcer's earnings and that this overreaction is not corrected till the late announcing firm reveals its earnings. Thus, in our design we should observe that prices for a firm should overreact to information or signals about a related firm in the same industry.

Furthermore, in agreement with Gillette et al. (1999) and Nelson et al. (2001) we assume that misreaction (i.e. the absolute value of over- or underreaction) in market prices is not substantially smaller than misreaction in individual forecasts as individual biases are not corrected by market forces. Hypotheses 1a, 1b and 1c capture these intuitions.

Hypothesis 1a: Subjects tend to overreact to new information about related stocks when submitting stock forecasts.

Hypothesis 1b: Market prices tend to overreact to new information about related stocks.

Hypothesis 1c: Misreaction in market prices is not lower than misreaction in stock forecasts.

Besides analyzing how individuals process new information and how market prices react to new information we want to analyze the effects of overreaction in the long run. If overreaction in our experimental setting is a systematic bias then it should not diminish, even though subjects have the possibility to learn from the past and to acquire experience. Stable levels of overreaction over the course of the 120 seconds of a trading round imply that less rational subjects are not able to learn from the actions of more rational subjects. Moreover, we argue that there is no learning effect over various rounds. The level of overreaction in the first few rounds is not significantly larger than the level of overreaction in the later rounds.

The view that overreaction does not vanish with the course of the experiment is consistent with findings in Offerman and Sonnemans (2004) who show that overreaction in the coin-spin scenario does not disappear even if subjects are trained and have more experience with the task at hand.

Furthermore, psychological evidence indicates that outcome feedback is not as efficient at

lowering biases as other forms of feedback such as cognitive or task information feedback (see e.g. Benson and Önkal, 1992 or Goodwin et al., 2004). Since outcome feedback is the major feedback subjects receive in financial markets and in order to keep our design as realistic as possible we restrict our feedback to simple outcome feedback. Moreover, most other studies analyzing whether individual biases are corrected by market forces also only give subjects outcome feedback.

Consistent with findings in Offerman and Sonnemans (2004) who also analyze the role of overreaction in a dynamic setting we propose the following three learning hypotheses:

Hypothesis 2a: Learning within a round: Overreaction in stock prices remains stable in the course of a round.

Hypothesis 2b: Learning over rounds: Overreaction in stock forecasts does not diminish over the course of the experiment.

Hypothesis 2c: Learning over rounds: Overreaction in stock prices does not diminish over the course of the experiment.

In addition to analyzing the level of individual and market overreaction in an experimental asset market our design allows us to test further insights from theoretical models. According to Milgrom and Stokey (1982) even in the wake of new private information no trade should occur if agents have rational expectations. However, Varian (1989) argues that trade can be triggered by heterogenous beliefs of market participants. These heterogenous beliefs or differences in opinion appear if subjects have differing prior beliefs and/or if subjects interpret new public information differently. Related to this Harris and Raviv (1993) and Kandel and Pearson (1995) show that in speculative markets differences of opinion can explain observed high levels of trading volume. In a similar vein, Cao and Ou-Yang (2009) show that differences of opinion or disagreement about the mean of new information has an impact on trading volume in stocks but does not affect trading volume in options. All models on differences of opinion have one thing in common: the straightforward implication that trading volume is higher the more heterogenous the traders' beliefs are.

Empirical and experimental evidence on the relation between differences of opinion and

trading volume is scarce but seems to confirm the theoretical models. Antweiler and Frank (2004) analyze the effect of Internet stock message boards on trading volume in stock markets. Comparing the level of disagreement in these messages with trading volume they find evidence for the theoretical propositions. Other studies in the accounting literature (see e.g. Bamber et al, 1997) support this view and show that trading volume around earnings announcements is related to different aspects of disagreement among agents. Furthermore, Hales (2009) shows experimentally that subjects in a 2-person economy trade more aggressively if they receive more diverging signals. He argues that this aggressive trading volume can be reduced if subjects are not asked to forecast the value of a stock but the level of disagreement between agents in an economy. Hypothesis 3 captures the main intuition of the differences of opinion literature

Hypothesis 3: Trading volume is higher if subjects in a market have more differences of opinion, i.e. more diverging expectations.

3   Experimental Design and Procedure

3.1   Theoretical Framework

Similar to the framework in Biais et al. (2008) subjects in our study observe a signal ($\tilde{s}$) about the price of a stock ($\tilde{v}$). Since the signal is noisy it can be decomposed as follows:

$$\tilde{s} = \tilde{v} + \tilde{e}.$$

In our experimental setup subjects know that all random variables are identically, jointly normal, and independently distributed. Thus, using the projection theorem we can calculate the rational Bayesian benchmark:

$$E(\tilde{v}|s) = E(\tilde{v}) + \frac{cov(\tilde{v}, \tilde{s})}{var(\tilde{s})}(s - E(\tilde{s})) = E(\tilde{v}) + \delta(s - E(\tilde{s})). \tag{1}$$

δ which is equal to $\frac{cov(\tilde{v},\tilde{s})}{var(\tilde{s})}$ corresponds to the level of overreaction in this design. Thus, subjects who overweight the informational content of a signal will overestimate δ and

submit a biased forecast:

$$F = \hat{E}(\tilde{v}|s) = E(\tilde{v}) + \hat{\delta}(s - E(\tilde{s})), \qquad (2)$$

Comparing equations 1 and 2 we can derive a simple measure of overreaction:

$$\frac{F - E(\tilde{v}|s)}{s - E(\tilde{s})} = \frac{\hat{E}(\tilde{v}|s) - E(\tilde{v}|s)}{s - E(\tilde{s})} = \hat{\delta} - \delta. \qquad (3)$$

If $\hat{\delta} > \delta$ we observe overreaction as a subject's forecast to a surprisingly good signal is larger than the conditional expectation ($F > E(\tilde{v}|s)$) and smaller than the conditional expectation ($F < E(\tilde{v}|s)$) in case of a surprisingly bad signal. If on the other hand our overreaction measure is smaller than 0 we observe underreaction. If misreaction in our experiment is not systematic we should find that it should cancel out on average.[4] We are going to describe the data generating process, the experimental procedure and the calculation of the rational benchmark in more detail in the following subsections.

3.2 Basic Design

The experiment consists of three main parts: an instruction phase, a trading phase, and a questionnaire. The instruction phase allows subjects to get familiar with the trading environment and to gain experience with the trading mechanism. It also gives subjects information about the payoff structure.

The trading phase consists of a sequence of 8 consecutive rounds. At the beginning of each round subjects were shown a graph illustrating the stock price movement of two related stocks G and H for the last 6 months. Subjects were told that stock price changes of both stocks at a given day i in trading round j depend upon a firm specific shock $\eta_{i,j}^{H}$ (for stock H) and $\eta_{i,j}^{G}$ (for stock G) and upon an industry shock $\epsilon_{i,j}$ which is common to both stocks. Moreover, subjects knew that these daily shocks are i.i.d., normally distributed,

---

[4]In contrast to Biais et al. (2008) we do not use *Overreaction-Beta* as additional measure of overreaction in this study. Since *Overreaction-Beta* is simply defined as the coefficient of a simple ordinary least squares regression of forecasting error onto signal for each person it is not really applicable to use it in our market design as subjects have traded only with 8 different stocks and we would need to base our analyses on regression coefficients that are based on 8 observations.

and stationary over time.[5] Then subjects were asked to provide three estimates for the price of stock H at time t = 12: a best guess as well as a lower and an upper bound. They were told to set the lower (upper) bound so that the price of stock H at time t = 12 would fall below (be above) the bound with a very low probability of 5%. Figure 1 illustrates the computer screen in the estimation task.

Insert figure 1 here

After having provided these three values subjects received an initial endowment of 1,500 monetary units and 5 shares of stock H and were allowed to start trading the stock with all other market participants for 120 seconds. Trading took place in a single-unit open-book double auction market similar to the market experiments of Plott and Sunder (1982 and 1988) and Weber and Welfens (2007). After 120 seconds trading was stopped and subjects received the stock price development of the related stock G for the second 6 months as an additional information. Incorporating this new information or signal, they were asked again to provide a best guess as well as a lower and upper bound for the price of stock H at time t = 12. Having provided these estimates the trading floor opened again and subjects could trade with each other for further 120 seconds. Figure 2 illustrates the computer screen of the trading market for the second period of a round.

Insert figure 2 here

After further 120 seconds the trading floor closed for a second time and subjects were informed about the realized price of stock H at time t = 12, the round ended and the next round started. Thus, in each trading round the trading floor is open for overall four minutes, two minutes before signal revelation and two minutes after signal revelation. Note that the 2·2 minute trading periods are exactly the same as the ones in Weber and Welfens (2007). Moreover, other studies that analyze the existence of individual biases in market settings also use two minute (Gillette et al. (1999)), three minute (Camerer (1987)), and four minute (Camerer (1987), Camerer et al. (1989), and Lei et al. (2001)) trading periods. Since some of these studies show that individual biases are reduced substantially by market

---

[5]The simulation of the price paths was similar to the one in Biais et al. (2008). The only difference was that instead of using daily shocks with a mean of 0.025 we use daily shocks with no trend, i.e. a mean of 0.

forces the length of our periods should be sufficient to offer the possibility of learning in our experimental design as well. Figure 3 summarizes the course of an experimental round.

Insert figure 3 here

Subjects know that the true price of stock H at the end of each round is determined by the sum of all daily firm H specific shocks and all industry shocks that are common to both firms in each graph. Using Bayesian updating subjects have all necessary information to calculate the conditional expected value of stock H at time t = 12 (trading day 252) in round j as well as the variance given a particular signal s explicitly. The signal in this case is simply the stock price development of the related stock (G) in the second 6 months which is given to subjects before the second turn of trading. Hence, the conditional expectation follows:

$$E(\tilde{v}_{i,j}^H \mid s) = v_{i,j}^H + E\sum \cdot[(\tilde{\epsilon}_{i,j})] + \frac{1}{2} \cdot s \tag{4}$$

Using the conditional expectation in equation 4 we can calculate the level of overreaction for each stock price forecast and each transaction price by plugging it into equation 3.

Subjects in our market are informed about the exact underlying process that generates stock prices and they see the entire order book with all purchase and sell orders as well as past transaction prices. Own orders are illustrated in red whereas orders from all other subjects are illustrated in black. Purchase orders in the experimental markets require that a subject has enough monetary units to pay the specified price as we did not allow for any short sales. Selling orders only enter the order book if a subject holds a positive number of shares. In addition, subjects can not submit a selling order which is below their own already existing purchasing order or set a purchasing order which is above their own already existing selling order.

Since our experimental market is a single-unit open-book double auction market subjects can only place one selling order and one purchasing order at a time. If subjects already have a purchasing (selling) order in the books, they can replace it by entering a new purchasing (selling) order. Market clearing happens automatically and in continuous time and transaction prices are always equal to the price of the more senior order. To control

for order effects the order in which various stocks appear is varied over different markets. After having finished all 8 rounds of trading with two trading periods each, a questionnaire starts automatically. In the questionnaire we elicit demographics such as age, gender, and line of studies.

3.3 Procedure and Descriptive Statistics

The experiment was programmed in Java and run at the University of Mannheim in November 2008. In 11 of overall 13 session we had 8 students who made up one market whereas in the other two sessions we had 6 and 7 students, respectively. Since Lei et al. (2001) and Noussair et al. (2008) also have markets with varying numbers of participants (6-8) we carry out the analyses for all 13 sessions. However, our results remain stable if we exclude the two sessions with less than 8 subjects.

Thus, we had a total of 101 students, 51 male and 50 female, who participated in our experiments. In every session we had 8 rounds of trading, each with two 120-second trading periods which gives us a total of 208 trading periods or 416 minutes of trading.[6] The average age of all subjects was 24.7 and approximately one half of the participants studied economics or business administration. The average processing time for the whole experiment including the instruction phase, the trading phase, and the questionnaire was 80 minutes.

Subjects' payment was determined as follows: subjects received a flat payment of 4 Euro for filling out the questionnaire and a variable, performance based payment for their participation in the 8 rounds of the experimental market. More specifically, the performance based payment was equal to 0.06% of the overall final wealth for all 8 rounds. The final wealth at the end of a single round was simply the amount of monetary units at the end of the round plus the number of stocks held multiplied with the realized value of the stock. Earnings averaged 13.35 Euro per subject.

Overall, we observe 4,157 buy orders and 4,282 sell orders with a total of 2,063 transac-

---

[6] In pre-tests with more than 8 rounds of two 120-second trading periods subjects indicated to us that the task was to strenuous and advised us to reduce the number of rounds. Moreover, the number of rounds is close to the one in various other studies such as Sarin and Weber (1993, 2 and 8 rounds) or Weber and Welfens (2007, 10 rounds).

tions. Interestingly, the number of trades in the first trading period of a round, i.e. before subjects received a signal, is substantially higher than the number of trades in the second trading period of a round, i.e. after subjects received the signal. In the first part of a round 1,194 transactions were conducted whereas in the second part of a round only 869 transactions were completed. These findings are in line with Weber and Welfens (2007) who also find substantially lower levels of trading after a fundamental shock. These numbers point out to a high trading volume of 11.5 transactions in the first 120-second trading period of a round, i.e. before a signal, and 8.4 transactions in the latter 120 seconds, i.e. after the signal was received.

Moreover, we find that the number of shares held by subjects at the end of a period varies substantially from a low of 0 to a high of 23. Hence, similar to other experimental asset markets (such as Plott and Sunder, 1982 and 1988 and Weber and Welfens, 2007), these findings are a first indication for relatively high levels of trading volume.

4 Results

4.1 Existence of Overreaction

The first goal of the market experiment was to detect the level of overreaction in individual stock forecasts and aggregate market prices with a design similar to Thomas and Zhang (2008) and Biais et al. (2008). Since we want to analyze the level of overreaction to a given signal we can only use observations from the second trading period of a round, i.e. from the period where subjects know the stock price of the related stock G for the entire time period. In order to analyze overreaction in stock forecasts of all traders we compute their overreaction score using equation 3 and plug in their best guess as forecasts F. Similar to Biais et al. (2008) we find overreaction to the signal about a related stock on an individual level in our experimental markets.

The upper graph in figure 4 illustrates the distribution of overreaction scores in stock forecasts for each stock and subject separately. Both the median (0.35) and the mean (0.41) score are significantly larger than zero, indicating substantial levels of overreaction. However, there is substantial variation in the level of overreaction in our sample with the

scores ranging from -1.5 to 3.2. In other words, for a majority of 599 observations we find forecasts indicating overreaction, but for 195 forecasts we also observe underreaction.[7]

The middle graph in figure 4 shows the average overreaction score in forecasts for each person. Interestingly, mean (0.41) and median (0.41) overreaction are in the same range as for each stock separately. However, almost all subjects tend to overreact on average with only 6 out of 101 subjects having a mean overreaction score of less than 0. This finding is a first hint that the level of overreaction varies heavily over different rounds but that there is a general and highly significant ($p < 0.01$) tendency for overreaction (more details on this issue will follow later in this subsection).

Analyzing the level of overreaction in market prices, a similar picture arises with a mean overreaction score in prices of 0.42 and a median score of 0.32. The lower graph in figure 4 illustrates the level of overreaction in market prices. We find that a large majority of 637 transactions are conducted at overreacting prices whereas only 232 transactions are conducted at underreacting prices. Using a Wilcoxon signed-rank test and a t-test we find that overreaction in prices is significantly larger than 0 ($p < 0.01$). However, both tests implicitly assume that transaction prices are independent. This is not the case as our experimental market is a single-unit open-book double auction market where subjects can repeatedly buy and sell the asset. To control for this we analyze the level of overreaction only for the first transaction in each trading period after subjects received a signal which gives us a total of 104 observations (13 sessions · 8 rounds). The mean (median) level of overreaction for the first transaction in each trading period is 0.43 (0.30) which is significantly larger than 0. In addition, a Mann-Whitney rank-sum test shows that the level of overreaction for the first transactions is not substantially different from the level of all other transactions ($p = 0.85$).

Another problem for the interpretation of overreaction in transaction prices might be the existence of short selling constraints. For example, after a good signal subjects who overreact are willing to buy the stock at prices which are too high whereas subjects who react rationally are not able to drive the market prices down to the rational level by selling more than their five inventory stocks at seemingly inflated prices. We control for

---

[7]Keep in mind that both overreaction in forecasts and overreaction in prices can only be calculated for the second two minute trading period in each round, i.e. after the signal was revealed.

the short selling problem using two approaches: first, similar to the non-independency problem above, we show that overreaction for the first transaction in each trading period is not smaller than for all observations ($p = 0.85$). The fact that the short selling constraint is not binding in the first transaction after a signal revelation and that overreaction is still present is a first indication that without short selling constraints overreaction would not abate. Second, we exclude all markets in which at least one subject hits the boundary and ends up with 0 assets in his / her portfolio. Analyzing the level of overreaction for markets where the short selling constraint is not binding we find a significantly positive mean overreaction score in prices of 0.5 and a median score of 0.41. A Mann-Whitney rank-sum test shows that the level of overreaction in prices is even larger than the level of overreaction in markets where the short selling constraint is binding ($p < 0.01$). Thus, we believe that short selling constraints cannot explain the substantial levels of overreaction in our markets.

Overall, both overreaction in forecasts and in prices are in the same ballpark as the level of overreaction in Biais et al. (2008) who find average median overreaction scores of 0.33 and 0.37. Similar to their results, we also observe large heterogeneity of overreaction, with some subjects providing almost rational estimates and trading at rational prices but with a majority of subjects overreacting to the new information. Moreover, our results support empirical findings on overreaction in Thomas and Zhang (2008) in a clean experimental design, where we can control for other possible explanations. Thus, both hypothesis 1a and hypothesis 1b are supported by our results.

Insert figure 4 here

Next, we turn to hypothesis 1c and analyze whether market forces help reducing the level of misreaction, i.e. over- or underreaction. Figure 5 illustrates the median level of misreaction in forecasts and market prices for each stock separately. Comparing the level of misreaction in forecasts and prices for each stock separately we find that self-regulating forces of the market do not seem to help in lowering misreaction of subjects. For 7 out of 8 stocks median misreaction in prices is even larger than the median misreaction in stock forecasts. Thus, the standard argument that self-regulating forces of markets will correct for erroneous individual beliefs and hence result in less overreaction does not apply in our

experimental market. We have illustrated previously, that there is a large heterogeneity of overreaction with some subjects having almost perfectly rational expectations. Thus, we can rule out the argument that market forces do not help in lowering the individual bias due to the fact that there are no traders in the market with rational expectations. This finding is in line with Gillette et al. (1999) who show for a different kind of task that individual misreaction in a continuous double auction market is slightly smaller than aggregate market misreaction which supports hypothesis 1c.

In the following we want to analyze in more detail why market forces or more specifically traders with rational estimates fail to lower misreaction. The first 3 columns of table 1 show the median overreaction score in stock forecasts of subjects who bought stocks ($Median-OR^{Buyer}$), subjects who sold stocks ($Median-OR^{Seller}$), and all subjects, regardless whether they traded the stock or not ($Median-OR^{All}$), for each stock separately. Column 4 shows the median overreaction score derived from prices ($Median-OR^{Prices}$) broken down by stocks. The results confirm our previous assumption that the level of overreaction varies heavily with the stock analyzed. For most stocks subjects tend to overreact to the signal, however, for two stocks (stock 3 and 5) most subjects show the tendency to underreact.[8]

The results in table 1 also suggest that some sort of "winner's curse" might explain the finding that misreaction in market prices is not smaller than misreaction in forecasts. For stocks with a good signal which are marked by a $^+$ sign in table 1, buyers are excessively optimistic and keep prices too high whereas for stocks with a bad signal which are marked by a $^-$ sign in table 1, sellers are excessively pessimistic and keep prices too low. In other words, after good signals buyers heavily overreact and are willing to pay a price which is

---

[8]We try to analyze why overreaction is present for some stocks and why underreaction is present for some other stocks. As we varied the order in which stocks were presented, ordering effects cannot explain our finding. According to Griffin and Tversky (1992) subjects should overreact to signals with high strength and low weight and underreact to signals with low strength and high weight. In our experimental setup the weight of a signal is fixed due to the fact that the stock prices depend upon a common industry shock and the correlation between the two prices should be 0.5 on average. The strength of a signal is equal to the percentage change in the stock price of stock G. Hence, overreaction should be higher the more extreme and lower the less extreme the signal is. However, our results show that for the two stocks with the lowest (stock 1) and the highest (stock 8) signal the level of overreaction is somewhere in the middle. Furthermore, even if we assume that the weight is not fixed but is inferred from the correlation between the two stocks in the graph the pattern of overreaction and underreaction cannot be explained by differences in signal strength and weight. Thus, further research is required to analyze what factors influence the level of overreaction and underreaction exactly.

way above the rational benchmark and even above the average forecast whereas after a bad signal sellers heavily overreact and are willing to sell at a price which is way below the rational benchmark and even below the average forecast. Thus, buyers - after a good signal - and sellers - after a bad signal - suffer from what we call a "winner's curse" and cause transactions to be executed at overreacting prices.

<div align="center">**Insert figure 5 here**</div>

<div align="center">**Insert table 1 here**</div>

4.2   Learning to Overreact Less

In addition to finding evidence for the existence of individual and market overreaction we are interested in analyzing the effects of overreaction in the long-run. More precisely, we would like to know if subjects are able to learn to overreact less with more trading experience. Subjects gain trading experience in two ways: first, within the course of a round as they are able to learn from the bids and asks of other subjects and second, over rounds as they receive outcome feedback at the end of each round. In the following, we will first analyze learning within a round and then learning over rounds.

Figure 6 shows the average (upper graph) and the median (lower graph) overreaction scores for each trading second in the second trading period of a round. To measure whether there is a learning effect within a round we estimate a simple partial adjustment model similar to the one in Camerer (1987), Camerer et al. (1989), and Weber and Welfens (2007):

$$OR_t = \alpha + \beta \cdot OR_{t-1} \tag{5}$$

Hence, we calculate the level of average overreaction in second t as a function of average overreaction in the previous second. As also can be seen in the graphs there are hardly any learning effects and the level of overreaction remains fairly stable over time. Moreover, in line with Camerer (1987) we estimate the degree of equilibrium bias $OR^{Equilibrium}$ consistently by the estimator $\hat{OR} = \frac{\hat{\alpha}}{(1-\hat{\beta})}$, where $\hat{\alpha}$ and $\hat{\beta}$ denote ordinary least squares

estimators of α and β from equation 5. Fitting with an ordinary least squares regression results in $\hat{\alpha} = 0.456$ and $\hat{\beta} = -0.112$ with $\hat{\beta}$ not being significantly different from 0. Furthermore, the equilibrium bias $\hat{OR} = \frac{\hat{\alpha}}{(1-\hat{\beta})} = 0.41$ and thus substantially larger than 0.

To analyze findings on learning within a round in more depth we split the 120-second trading period into six 20 seconds intervals and calculate three overreaction scores for each interval separately. Table 2 reports $Overreaction^{Mean}$ and $Overreaction^{Median}$ which are simply the mean and median overreaction in each time span and $Overreaction^{Proportion}$ which is the proportion of transactions that were carried out at overreacting prices in a given time span. A series of binomial tests which analyzes whether most transactions were carried out at overreacting prices indicates that overreaction is prevalent in each time interval. The same results emerge if we use Wilcoxon signed-rank tests.

Moreover, the numbers in the respective time intervals seem to be very similar and closely related. Using a series of Mann-Whitney rank-sum test we find that no single difference between two time spans is significant indicating that learning effects are hardly existent within a round and that the level of overreaction is stable over time. Hence, we do not observe any learning effects within a round, consistent with hypothesis 2a.[9]

Besides learning within a round subjects could also gain experience over the course of the experiment and learn from the outcome feedback that is provided to them. For a more detailed test, we report means and medians of overreaction as well as the proportion of overreacting forecasts or transactions for each trading round in table 3. Looking at the results we find very large differences between single rounds[10], however, we are not able to detect a significant trend. Our results that overreaction does not disappear with learning effects are stable even if we control for financial expertise which is proxied by the number of finance courses a subject attended or other demographic variables.

Overall, the findings in this subsection confirm our hypotheses 2a, 2b, and 2c. Overreaction remains stable even though subjects acquire more experience and have the possibility to learn both from the actions of other subjects and from outcome feedback. This is in

---

[9]Our results remain stable if we split the 120-second trading period in 3, 4, or 10 time spans of equal length.

[10]Note, that we control for order effects by varying the sequence in which graphs were presented.

line with results in Offerman and Sonnemans (2004) who show that overreaction in the coin-spin scenario is present even if subjects receive extensive training. It is also in line with findings in Bloomfield et al. (2000) and Kraemer and Weber (2004) who show that expertise is of no help in lowering overreaction. Similarly, Benson and Önkal (1992) and Goodwin et al. (2004) point out that outcome feedback is hardly effective at reducing biases. Moreover, the results in Thomas and Zhang (2008) who analyze a similar setting to ours empirically also suggests that this kind of overreaction to earnings announcements of related firms is a stable construct in real-world markets.

<center>**Insert figure 6 here**</center>

<center>**Insert table 2 here**</center>

<center>**Insert table 3 here**</center>

4.3 Differences of Opinion and Trading Volume

As reported in subsection 3.3 subjects seem to be engaged in relatively high levels of trading volume with on average 9.9 transaction in each 120-second trading period. In this experimental setup two testable explanations for the observed high trading volume can be brought forward: first, differences in risk attitudes between traders and second, heterogenous beliefs about the value of a stock, i.e. differences of opinion.

To show whether trading volume can be explained by risk sharing motives we compare the trading volume, i.e. the number of shares traded, in each trading period with the level of differences in risk attitudes between buyer and seller ($Differences\ Risk\ Aversion^{MaxMin}$). Risk Aversion was simply measured for each person on a five point Likert scale with the endpoints "1 = high risk aversion" and "5 = very low risk aversion". $Differences\ Risk\ Aversion^{MaxMin}$ is simply defined as the difference between the most risk averse and the least risk averse subject in each session.[11] Running a clustered least square

---

[11]Alternatively, we also compute $Differences\ Risk\ Aversion^{Std.}$ which is the standard deviation of all subjects' risk attitude measures in each session and relate it to the varying measures of differences of opinion. The results of the following analyses are essentially the same.

<center>23</center>

regression of the level of trading volume on the differences in risk attitudes between the two traders we are not able to find a significant effect. Differences in risk attitudes cannot explain the observed high levels of trading volume (see regression 1 in table 4).

The second argument that is often brought forward to explain high levels of trading are differences of opinion. Amongst others, Varian (1989), Harris and Raviv (1993) or Kandel and Pearson (1995) propose that excessive trading volume can be explained by heterogenous beliefs among market participants. Since we elicited beliefs of all market participants in each period in the estimation phases we are able to analyze this theoretical proposition in our experimental setup.

Using forecasts (best guesses) that were submitted by all subjects in the estimation part of the experiment before each 120-second trading period we construct three distinct differences of opinion measures. Our first measure ($Differences\ of\ Opinion^{MaxMin}$) is simply the difference between the most optimistic, i.e. highest best guess, and the most pessimistic, i.e. lowest best guess, in each 120-second trading period. The second measure is the standard deviation of all subjects' best guesses ($Differences\ of\ Opinion^{Std}$) before each 120-second trading period (see Morse et al. (1991) and Bamber et al. (1997)). Our last measure of differences of opinion, $Change\ in\ Dispersion$, is also adopted from Bamber et al. (1997). It measures the change in the standard deviation of forecasts in each trading round before and after signal revelation. More specifically, it is defined as the standard deviation of beliefs in the second trading period of a round minus the standard deviation of beliefs in the previous 120-second trading period of a round. For the first two measures we have 208 different observations (13 $sessions \cdot 8\ rounds \cdot 2\ trading\ periods\ per\ round$) and for the third measure we have 104 observations (13 $sessions \cdot 8\ rounds$).

The upper graph in figure 7 illustrates the relation between $Differences\ of\ Opinion^{MaxMin}$ and the number of shares traded in each of the 208 periods. The fitted line of a clustered least square regression and the positive Spearman correlation of 0.23 (Wilcoxon signed-rank test: $p < 0.01$) are a first indication that trading volume is positively related to differences of opinion. A similar picture emerges if we take a look at the middle graph in figure 7. $Differences\ of\ Opinion^{Std}$ is positively related to trading volume (Spearman's rho = 0.21 and $p < 0.01$). The lower graph in figure 7 illustrates the relation between $Change\ in\ Dispersion$ and $Change\ in\ shares\ traded$. The variable

Change in shares traded is defined as the relation between the number of shares traded in the second period of a round (i.e. after signal revelation), divided by the number of shares traded in the first period of a round. Similar to the results in the two upper graphs we observe a positive relation (Spearman's rho = 0.20 and $p = 0.04$).

However, one problem of correlation analysis is that we cannot account for non-independent residuals over sessions. Hence, we additionally run ordinary least squares regressions in which we cluster our observations over sessions for each difference of opinion measure separately and additionally control for differences in risk attitudes. The result in equation (2) supports our graphical findings as the coefficient of *Differences of Opinion^{MaxMin}* is significantly positive. If the difference between the most optimistic and the most pessimistic forecast is 50, then the number of shares traded in a market rises by 2.45 compared to a market with perfectly homogenous beliefs. Interestingly, *Differences Risk Aversion^{MaxMin}* cannot explain trading volume. Similar results emerge if we analyze the relation between *Differences of Opinion^{Std}* and trading volume. Our second measure of difference of opinion is also positively related to the number of shares traded in each period.

Regression (4) in table 4 analyzes the relation between *Change in Dispersion* and the change in the number of shares traded in the second period of a round compared to the first period of a round. Since we only have one observation for each round the number of observations drops to 104 (13 *sessions* · 8 *rounds*). The positive coefficient of *Change in Dispersion* indicates that the larger the standard deviation in forecasts in the second period of a round, compared to the first one, the larger the fraction of shares traded in the second period of a round compared to the first one. Again, *Differences Risk Aversion^{MaxMin}* are not able to explain trading volume.

Overall, the findings on the positive relation between differences of opinion and trading volume are consistent with first experimental evidence in simple 2-subject call markets (see Hales (2009)) and support our hypothesis 3. Interestingly, not the differences in risk attitudes but solely the differences of opinions are significantly related to trading volume. In addition, one could argue that finding evidence on the theoretically proposed relationship between differences of opinion and trading volume is a further indication that our data seems to be quite reasonable.

Insert figure 7 here

Insert table 4 here

5 Conclusion

This paper extends the individual-level-study by Biais et al. (2008) to a simple experimental trading market. We analyze if subjects are able to update their beliefs according to Bayes rule or if they misreact when they receive new information about a stock and consequently if market prices overreact. Consistent with findings in Biais et al. (2008) subjects in our setting overreact to new information on an individual level. Additionally, we find evidence for aggregate overreaction in market prices, consistent with Thomas and Zhang (2008). Interestingly, consistent with propositions in theoretical models (e.g. Odean (1998b) and Biais and Weber (2007)) and findings in the experimental literature (Gillette et al. (1999)) individual misreaction translates into market outcomes as misreaction in transaction prices is not lower than in individual estimates.

Furthermore, we analyze if overreaction both on an individual and market level diminishes over time, i.e. if subjects are able to learn from the actions of other subjects or from the outcomes in previous rounds. Our results indicate that learning effects can neither be observed within a two minute trading period nor over various rounds. Hence, overreaction in this setting seems to be a stable construct. This finding is consistent with Offerman and Sonnemans (2004) who illustrate individual overreaction to be persistent in their coin-spin design even though subjects receive extensive training. In addition, evidence in psychology indicates that outcome feedback, exactly the sort of feedback subjects receive in our experiment, is not very effective at reducing biases (Benson and Önkal (1992) and Goodwin et al. (2004)).

Lastly, we are able to provide experimental evidence for a positive relation between differences of opinion and trading volume. Although theoretical evidence on this relation is prevalent (Varian (1989), Harris and Raviv (1993), and Kandel and Pearson (1995)) empirical and experimental studies have mainly ignored this relation. A notable exception is the study by Hales (2009) who shows that trading volume is larger if two traders receive

more differing signals in a simple two-trader call market. We extend his findings to a more realistic environment by showing that trading volume is also larger if the disagreement about future stock price among all market participants is larger in a continuous auction market.

Further research should analyze if different sorts of feedback, such as cognitive or task information feedback, could help in lowering both individual and market biases. Previous findings on the role of different sorts of feedback on the level of overconfidence has shown that in particular outcome feedback is not helpful in lowering biases. In addition, it could be fruitful to study the high differences in the level of overreaction between stocks in more detail. Varying the degree of signal strength and weight or the length of the forecasting period might help in determining reasons for the large heterogeneity of over- and underreaction.

References

Abarbanell, J. S., and V. L. Bernard, 1992, "Tests of Analysts' Overreaction/Underreaction to Earnings Information as an Explanation for Anomalous Stock Price Behavior," *Journal of Finance*, 47, 1181–1207.

Antweiler, W., and M. Z. Frank, 2004, "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *Journal of Finance*, 59, 1259–1294.

Bamber, L. S., O. E. Barron, and T. L. Stober, 1997, "Trading Volume and Different Aspects of Disagreement Coincident with Earnings Announcements," *The Accounting Review*, 72, 575–597.

Barberis, N., A. Shleifer, and R. Vishny, 1998, "A Model of Investor Sentiment," *Journal of Financial Economics*, 49, 307–343.

Benson, P. G., and D. Önkal, 1992, "The Effects of Feedback and Training on the Performance of Probability Forecasters," *International Journal of Forecasting*, 8, 559–573.

Biais, B., D. Hilton, K. Mazurier, and S. Pouget, 2005, "Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market," *Review of Economic Studies*, 72, 287–312.

Biais, B., A. Nosić, and M. Weber, 2008, "Overreaction and Investment Choices: An Experimental Analysis," *Working Paper*.

Biais, B., and M. Weber, 2007, "Hindsight Bias and Investment Performance," *Working Paper*.

——— , 2009, "Hindsight Bias, Risk Perception, and Investment Performance," *Management Science*, 55, 1018–1029.

Bloomfield, R., and J. Hales, 2002, "Predicting the Next Step of a Random Walk: Experimental Evidence of Regime-Shifting Beliefs," *Journal of Financial Economics*, 65, 397–414.

Bloomfield, R., R. Libby, and M. W. Nelson, 2000, "Underreactions, Overreactions and Moderated Confidence," *Journal of Financial Markets*, 3, 113–137.

Budescu, D. V., and B. Maciejovsky, 2005, "The Effect of Payoff Feedback and Information Pooling on Reasoning Errors: Evidence from Experimental Markets," *Management Science*, 51, 1829–1843.

Camerer, C. F., 1987, "Do Biases in Probability Judgement Matter in Markets? Experimental Evidence," *American Economic Review*, 77, 981–997.

Camerer, C. F., G. Loewenstein, and M. Weber, 1989, "The Curse of Knowledge in Economic Settings: An Experimental Analysis," *Journal of Political Economy*, 97, 1232–1254.

Camerer, C. F., and D. Lovallo, 1999, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 89, 306–318.

Cao, H. H., and H. Ou-Yang, 2009, "Differences of Opinion of Public Information and Speculative Trading in Stocks and Options," *Review of Financial Studies*, 22, 300–335.

Daniel, K. D., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor Psychology and Security Market Under- and Overreactions," *Journal of Finance*, 53, 1839–1885.

——— , 2001, "Overconfidence, Arbitrage, and Equilibrium Asset Pricing," *Journal of Finance*, 56, 921–965.

DeBondt, W. F. M., and R. H. Thaler, 1985, "Does the Stock Market Overreact?," *Journal of Finance*, 40, 793–805.

——— , 1987, "Further Evidence on Investor Overreaction and Stock Market Seasonality," *Journal of Finance*, 42, 557–581.

——— , 1990, "Do Security Analysts Overreact?," *American Economic Review*, 80, 52–57.

Edwards, W., 1968, *Conservatism in Human Information Processing*. in: Formal Representation of Human Judgement, edited by: B. Kleinmuntz.

Fischer, P. E., and R. E. Verrecchia, 1999, "Public Information and Heuristic Trade," *Journal of Accounting and Economics*, 27, 89–124.

Ganguly, A. R., J. H. Kagel, and D. V. Moser, 1994, "The Effects of Biases in Probability Judgements on Market Prices," *Accounting Organizations and Society*, 19, 675–700.

——— , 2000, "Do Asset Market Prices Reflect Traders' Judgment Biases?," *Journal of Risk and Uncertainty*, 20, 219–245.

Gillette, A. B., D. E. Stevens, S. G. Watts, and A. W. Williams, 1999, "Price and Volume Reactions to Public Information Releases: An Experimental Approach Incorporating Traders Subjective Beliefs," *Contemporary Accounting Research*, 16, 437–479.

Goodwin, P., D. Önkal-Atay, M. E. Thomson, A. C. Pollock, and A. Macaulay, 2004, "Feedback-Labelling Synergies in Judgmental Stock Price Forecasting," *Decision Support Systems*, 37, 175–186.

Griffin, D., and A. Tversky, 1992, "The Weighing of Evidence and the Determinants of Confidence," *Cognitive Psychology*, 24, 411–435.

Grinblatt, M., and B. Han, 2005, "Prospect Theory, Mental Accounting, and Momentum," *Journal of Financial Economics*, 78, 311–339.

Hales, J., 2009, "Are Investors Really Willing to Agree to Disagree? An Experimental Investigation of How Disagreement and Attention to Disagreement Affect Trading Behavior," *Organizational Behavior and Human Decision Processes*, 108, 230–241.

Harris, M., and A. Raviv, 1993, "Differences of Opinion Make a Horse Race," *Review of Financial Studies*, 6, 473–506.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden, 2005, "Coordination of Expectations in Asset Pricing Experiments," *Review of Financial Studies*, 18, 955–980.

Hong, H., T. Lim, and J. C. Stein, 2000, "Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies," *Journal of Finance*, 55, 265–295.

Jegadeesh, N., and S. Titman, 1995, "Overreaction, Delayed Reaction and Contrarian Profits," *Review of Financial Studies*, 8, 973–993.

Kahneman, D., and A. Tversky, 1973, "On the Psychology of Prediction," *Psychological Review*, 80, 237–251.

Kandel, E., and N. D. Pearson, 1995, "Differential Interpretation of Public Signals and Trade in Speculative Markets," *Journal of Political Economy*, 103, 831–872.

Kluger, B. D., and S. B. Wyatt, 2000, "Are Judgment Errors Reflected in Market Prices and Allocations? Experimental Evidence Based on the Monty Hall Problem," *Journal of Finance*, 59, 969–997.

Kraemer, C., and M. Weber, 2004, "How Do People Take into Account Weight, Strength and Quality of Segregated vs. Aggregated Data? Experimental Evidence," *Journal of Risk and Uncertainty*, 29, 113–142.

Lei, V., C. N. Noussair, and C. R. Plott, 2001, "Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality," *Econometrica*, 69, 831–859.

Massey, C., and G. Wu, 2005, "Detecting Regime Shifts: The Causes of Under- and Overreaction," *Management Science*, 51, 932–947.

Milgrom, P., and N. Stokey, 1982, "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.

Morse, D., J. Stephan, and E. K. Stice, 1991, "Earnings Announcements and the Convergence (or Divergence) of Beliefs," *The Accounting Review*, 66, 376–388.

Nelson, M. W., R. Bloomfield, J. W. Hales, and R. Libby, 2001, "The Effect of Information Strength and Weight on Behavior in Financial Markets," *Organizational Behavior and Human Decision Processes*, 86, 168–196.

Noussair, C. N., G. Richter, and J.-R. Tyran, 2008, "Money Illusion and Nominal Inertia in Experimental Asset Markets," *Working Paper*.

Odean, T., 1998a, "Are Investors Reluctant to Realize Their Losses?," *Journal of Finance*, 53, 1775–1798.

———— , 1998b, "Volume, Volatility, Price, and Profit: When All Traders Are Above Average," *Journal of Finance*, 53, 1887–1934.

Offerman, T., and J. Sonnemans, 2004, "What's Causing Overreaction? An Experimental Investigation of Recency and the Hot-hand Effect," *Scandinavian Journal of Economics*, 106, 533–553.

Plott, C. R., and S. Sunder, 1982, "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models," *Journal of Political Economy*, 90, 663–698.

——— , 1988, "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085–1118.

Sarin, R., and M. Weber, 1993, "The Effect of Ambiguity in Market Setting," *Management Science*, 39, 602–615.

Scheinkman, J. A., and W. Xiong, 2003, "Overconfidence and Speculative Bubbles," *Journal of Political Economy*, 111, 1183–1219.

Seybert, N., and R. Bloomfield, 2009, "Contagion of Wishful Thinking in Markets," *Management Science*, 55, 738–751.

Sonnemann, U., C. F. Camerer, C. R. Fox, and T. Langer, 2008, "Partition-Dependent Framing Effects in Lab and Field Prediction Markets," *Working Paper*.

Sorescu, S., and A. Subrahmanyam, 2006, "The Cross Section of Analyst Recommendations," *Journal of Financial and Quantitative Analysis*, 41, 139–168.

Thomas, J., and F. Zhang, 2008, "Overreaction to Intra-Industry Information Transfers?," *Journal of Accounting Research*, 46, 909–940.

van Boven, L., G. Loewenstein, and D. Dunning, 2003, "Mispredicting the Endowment Effect: Underestimation of Owners' Selling Prices by Buyer's Agents," *Journal of Economic Behavior & Organization*, 51, 351–365.

Varian, H. R., 1989, *Differences of Opinion in Financial Markets*. in: Financial Risk: Theory, Evidence, and Implications, edited by: C.C. Stone, Boston: Kluwer.

Weber, M., and C. F. Camerer, 1998, "The Disposition Effect in Securities Trading: An Experimental Analysis," *Journal of Economic Behavior & Organization*, 33, 167–184.

Weber, M., and F. Welfens, 2007, "How Do Markets React to Fundamental Shocks? An Experimental Analysis of Underreaction and Momentum," *Working Paper*.

Table 1: Misreaction in prices vs. misreaction in forecasts

Column (1) to (3) of this table report the median overreaction scores in the forecasting task for buyers ($Median-OR^{Buyer}$), sellers ($Median-OR^{Seller}$), and all subjects in the market ($Median-OR^{All}$) separately for each stock. Column (4) reports the median overreaction score in transaction prices ($Median-OR^{Price}$) separately for each stock. $^+$ ($^-$) signs in the first column indicate that subjects received a good (bad) signal.

Stock	(1) $Median-OR^{Buyer}$	(2) $Median-OR^{All}$	(3) $Median-OR^{Seller}$	(4) $Median-OR^{Price}$
1^+	0.377	0.123	0.123	0.174
2^-	0.317	0.317	0.485	0.400
3^-	-0.206	-0.011	0.072	-0.178
4^+	1.290	1.290	1.290	1.189
5^+	0.110	-0.070	-0.162	-0.089
6^-	0.165	0.286	0.527	0.527
7^+	1.072	1.072	1.072	1.239
8^+	0.315	0.315	0.315	0.334

Table 2: Learning within a round

This table reports mean ($Overreaction^{Mean}$) and median ($Overreaction^{Median}$) overreaction scores in transaction prices as well as the proportion of transactions at overreacting prices for six time intervals of twenty seconds. Overreaction is calculated for every transaction separately and afterwards aggregated for each twenty seconds trading period.

Seconds	$Overreaction^{Mean}$	$Overreaction^{Median}$	$Overreaction^{Proportion}$
-120 to -101	0.354	0.296	68.14
-100 to -81	0.451	0.315	71.85
-80 to -61	0.411	0.322	69.33
-60 to -41	0.429	0.325	77.27
-40 to -21	0.403	0.315	75.17
-20 to -0	0.454	0.372	76.16

Table 3: Learning over rounds

This table reports mean and median overreaction scores in forecasts ($OR^{Estimates}$) and transaction prices (OR^{Prices}) as well as the proportion of overreacting forecasts and prices for each trading round separately.

	$OR^{Estimates}$			OR^{Prices}		
	Mean	Median	Prop.	Mean	Median	Prop.
Round 1	0.256	0.315	78	0.088	0.296	70.71
Round 2	0.567	0.485	79	0.702	0.485	91.76
Round 3	0.076	0.072	59	0.134	0.045	54.62
Round 4	0.673	0.712	76	0.724	0.547	83.05
Round 5	0.503	0.292	67	0.270	-0.071	40.35
Round 6	0.145	0.165	73	0.116	0.021	56.47
Round 7	0.757	0.678	89	0.937	0.906	99.07
Round 8	0.353	0.377	79	0.384	0.377	88.73

Table 4: Trading volume and differences of opinion vs. differences in risk attitudes

This table presents results on the relation between trading volume and differences of opinion as well as differences in risk attitudes using clustered least squares regressions (number of clusters is equal to 13). Trading volume in regressions (1) to (3) is simply the number of shares traded and in regression (4) the relation of shares traded in the second 120 trading seconds of a round divided by the number of shares traded in the first 120 trading seconds of a round. Our first measure of differences of opinion is $Differences\ of\ Opinion^{MaxMin}$ which is simply the difference between the most optimistic and the most pessimistic forecast in each 120-second trading period. $Differences\ of\ Opinion^{Std}$ is simply the standard deviation of all forecasts in each 120-second trading period. $Change\ in\ Dispersion$ is defined as the difference in the variation of expectations between the second and the first 120-second trading period in each round. Changes in the level of risk aversion $Differences\ Risk\ Aversion^{MaxMin}$ are simply the difference between the most risk averse and least risk averse subject in each trading market. We report regression coefficients and p-values in parentheses.* indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level.

|	(1) Shares traded	(2) Shares traded	(3) Shares traded	(4) Change in shares traded
$Differences\ of\ Opinion^{MaxMin}$		0.049		
		(0.002)***		
$Differences\ of\ Opinion^{Std}$			0.128	
			(0.003)***	
$Change\ in\ Dispersion$				0.017
				(0.003)***
$Differences\ Risk\ Aversion^{MaxMin}$	1.541	1.494	1.496	0.101
	(0.157)	(0.163)	(0.173)	(0.145)
Constant	6.600	4.566	4.768	0.602
	(0.011)**	(0.058)*	(0.056)*	(0.002)***
Observations	208	208	208	104
Adjusted R-squared	0.058	0.093	0.082	0.083
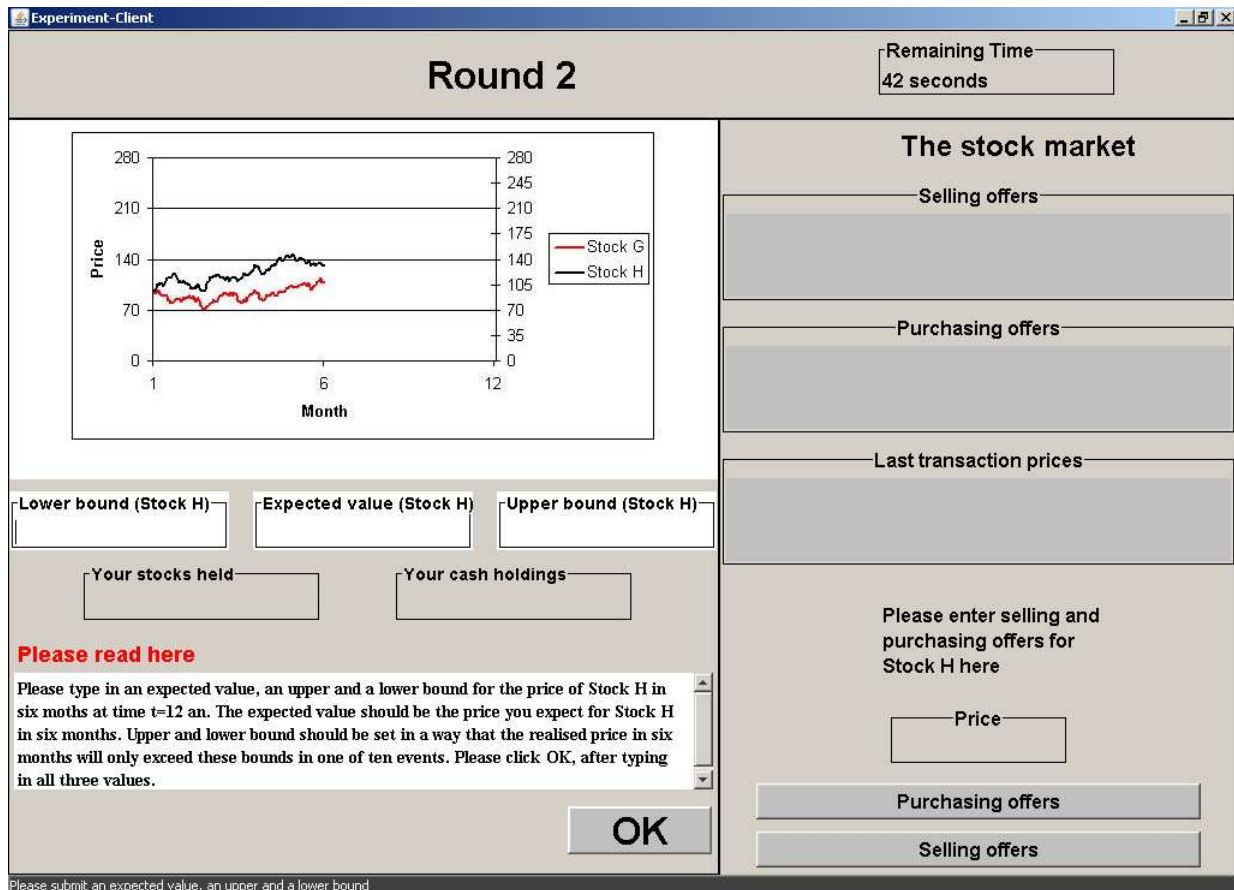
Figure 1: Estimation screen

The figure illustrates the screen of the estimation task in the first period of round 2. In the upper left-hand part the price development for the two stocks G and H in the first 6 months is displayed. In the three white boxes below the graph subjects can submit their three estimates (lower bound, best guess, and upper bound) for the price of stock H at t = 12. The remaining time for this task is shown in the upper right-hand part. The trading boxes on the right-hand side of the graph are in the estimation task inactive and are going to be explained in the following.
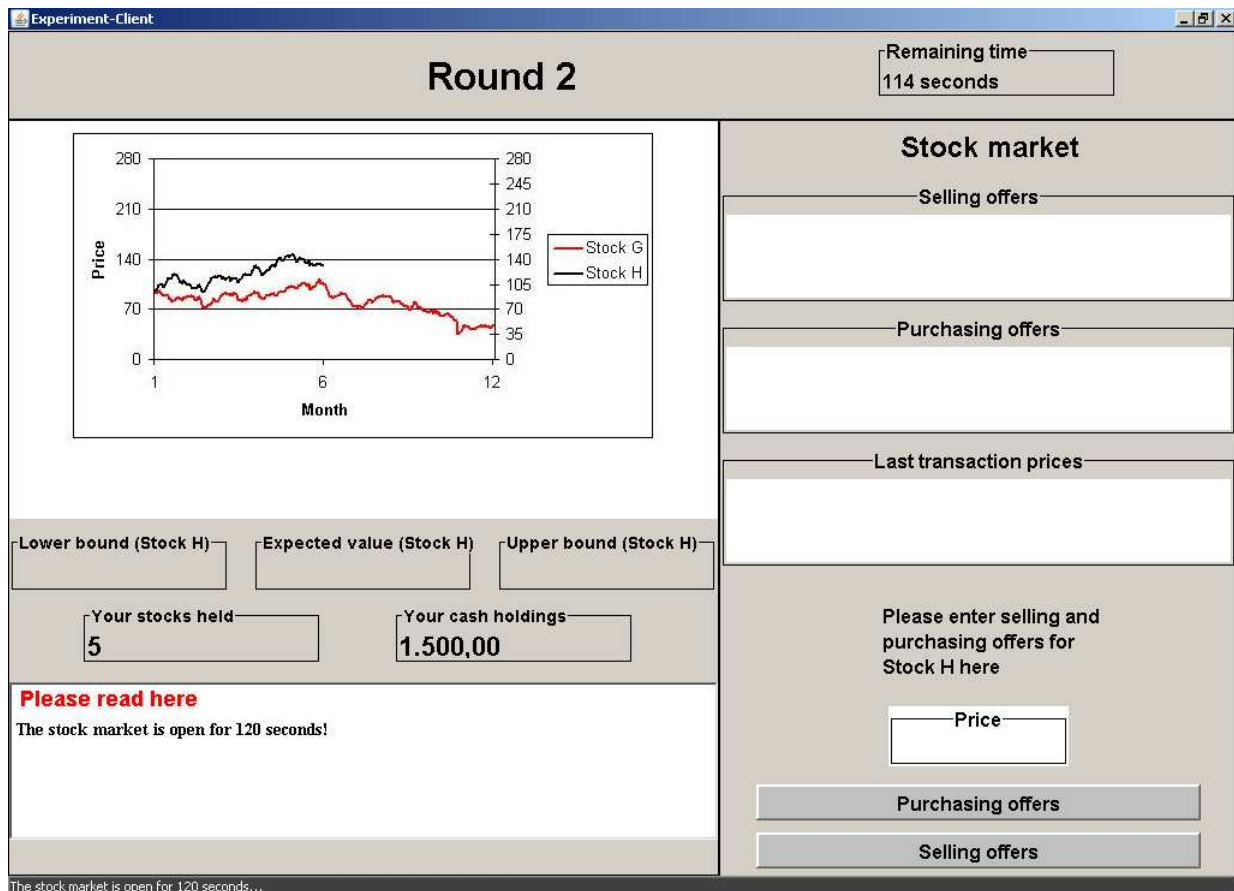
Figure 2: Trading screen

The figure illustrates the screen of the trading task in the second period of round 2. In the upper left-hand part the price development for stock H in the first 6 months and the price development of stock G over all 12 months is displayed. The three estimation boxes are inactive and the values subjects had previously submitted are not shown anymore. The trading boxes on the right-hand side are activated and show all selling offers, purchasing offers, and the last transaction prices. Purchasing offers are sorted highest to lowest, selling offers lowest to highest, and transaction prices are sorted by time. A subject's own selling or purchasing orders are illustrated in red whereas all remaining orders are displayed in black. To submit an order subject's simply had to type in a price and click on the button "Submit a selling order" and "Submit a purchasing order", respectively.
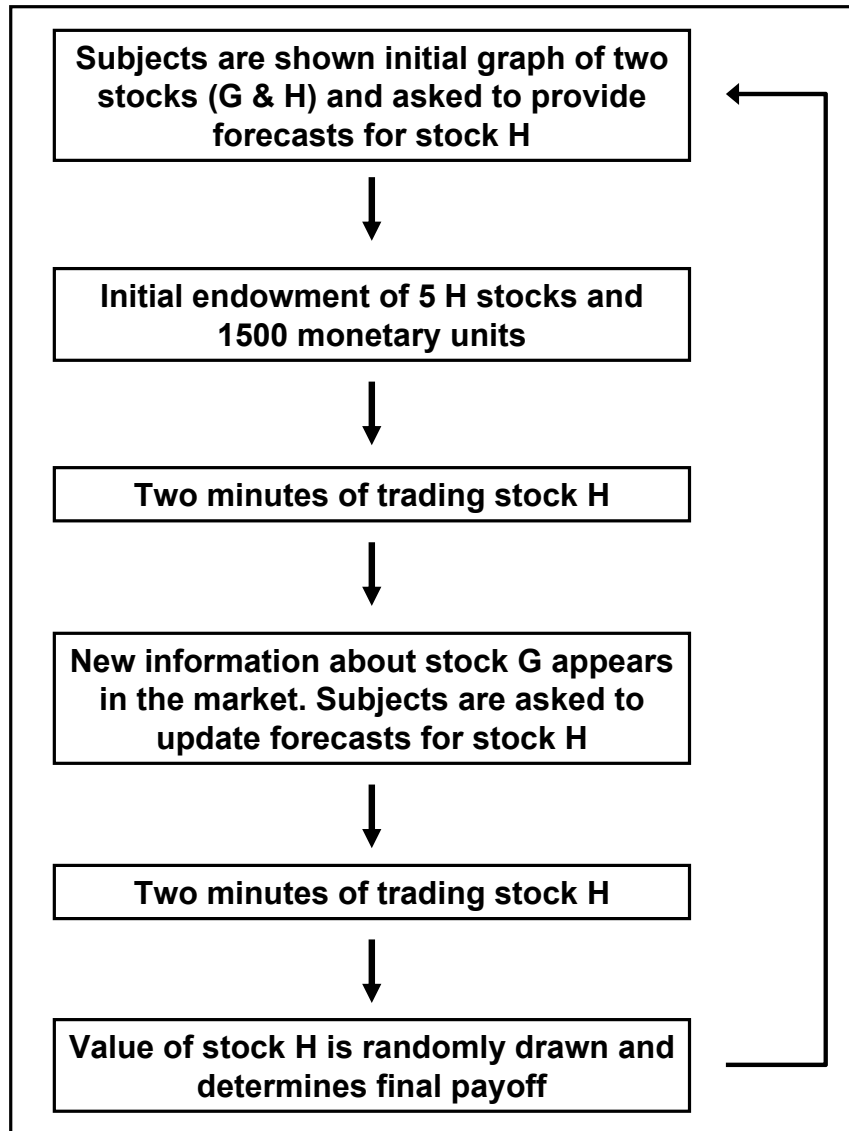
Figure 3: Course of the experiment

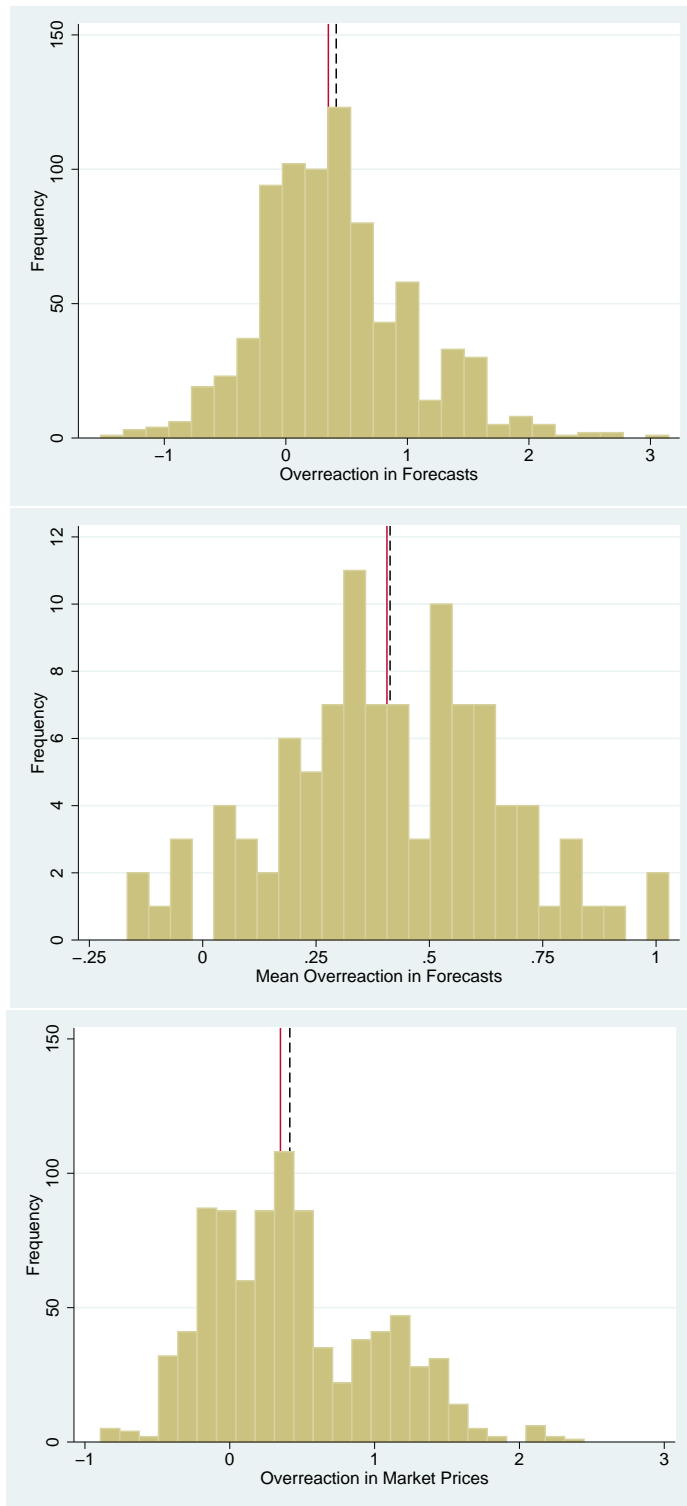The figure shows the course of each round in the market experiment.

Figure 4: Overreaction histograms

The figure shows the level of overall overreaction in stock forecasts (upper graph), the level of average overreaction for each person in stock forecasts (middle graph), and the level of overreaction in market prices (lower graph). Mean values are indicated by the dotted black line and median values by the continuous red line.
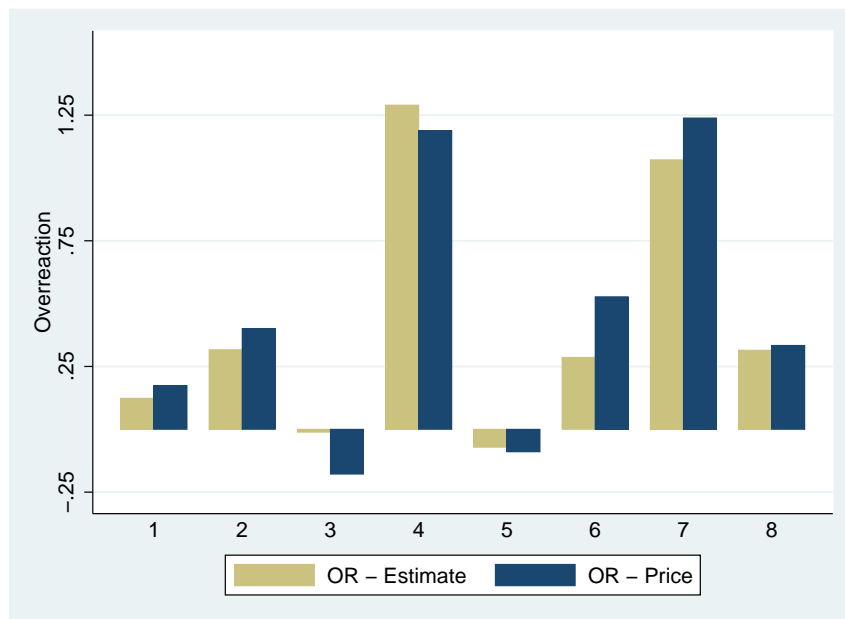
Figure 5: Overreaction prices vs. overreaction forecasts

The figure illustrates the level of overreaction in transaction prices (OR - Price) and in estimates (OR - Estimate) for all 8 stocks separately. A positive value on the y-axis indicates overreaction whereas a negative value indicates underreaction.
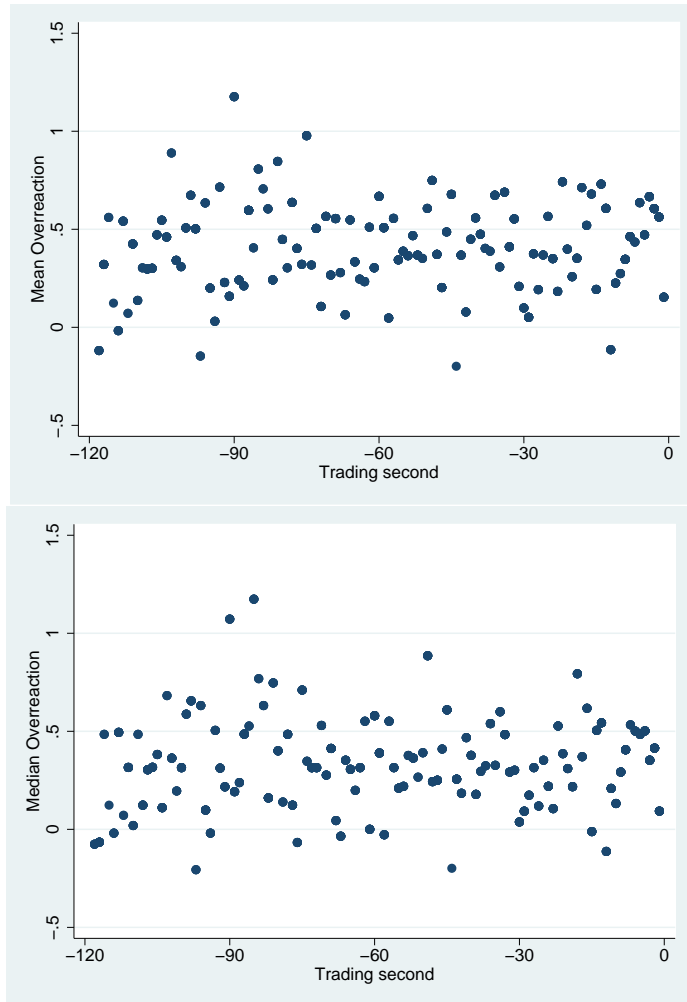
Figure 6: Learning within a round

The upper graph shows average overreaction scores in transaction prices over all rounds and the lower graph shows median overreaction scores in transaction prices over all rounds, for each trading second separately.
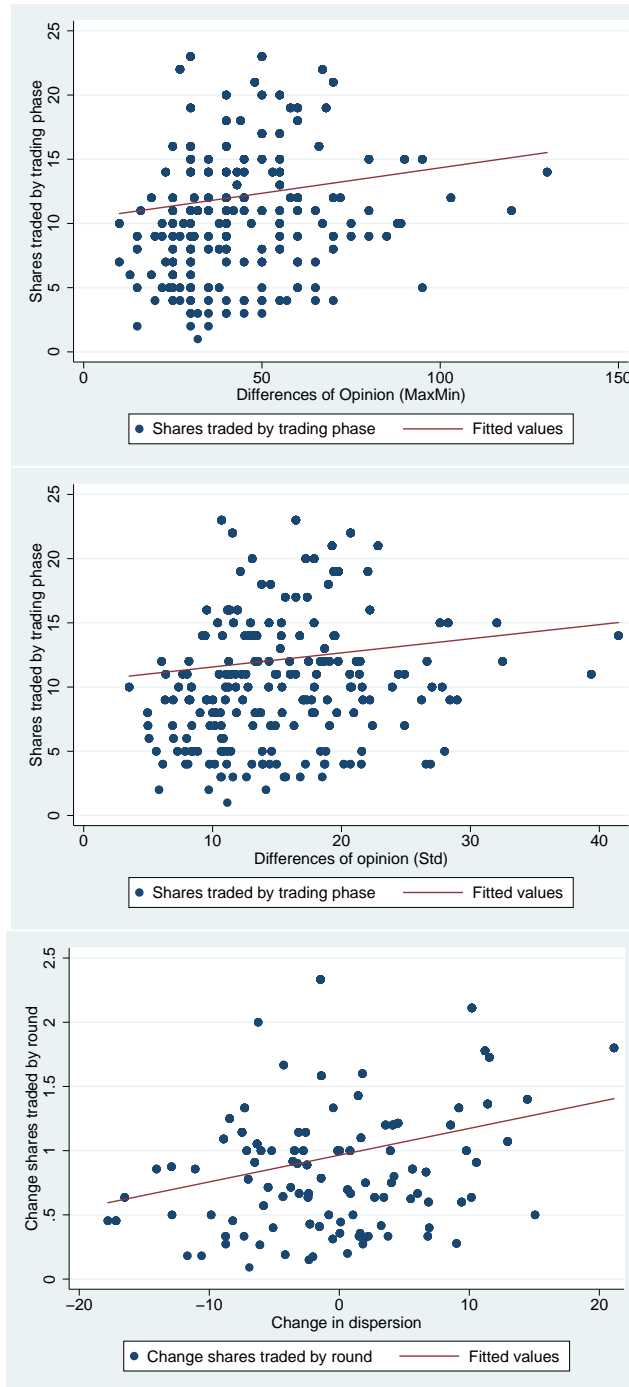
Figure 7: Differences of Opinion and Trading Volume

The two upper graphs illustrate the relation between number of shares that were traded in each 120-second trading period and two distinct measures for differences of opinion. The differences of opinion measure in the upper graph is simply the absolute difference between the most optimistic forecast and the most pessimistic forecast of subjects (*Differences of Opinion (MaxMin)*) in each 120-second trading period. Similarly, the differences of opinion measure in the middle graph is the standard deviation of all forecasts (*Differences of Opinion (Std.)*) in each 120-second trading period. The lower graph illustrates for each session and stock separately, how a change in the variation of expectations (*Change in Dispersion*) is related to a change in trading volume.