GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 7

Herding and Contrarian Behavior in Financial Markets - An Internet Experiment

Mathias Drehmann*
Jörg Oechssler**
Andreas Roider***

June 2004

*Mathias Drehmann, Bank of England
**Jörg Oechssler, Department of Economics, University of Bonn, Germany, oechssler@uni-bonn.de
***Andreas Roider, Department of Economics, University of Bonn, Germany

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Herding and Contrarian Behavior in Financial Markets - An Internet Experiment*

Mathias Drehmann† *Bank of England*

Jörg Oechssler‡ *Department of Economics University of Bonn*

Andreas Roider *Department of Economics University of Bonn*

Revised version: June 2004
First version: November 2002

Abstract

We report results of an internet experiment designed to test the theory of informational cascades in financial markets (Avery and Zemsky, *AER*, 1998). More than 6400 subjects, including a subsample of 267 consultants from an international consulting firm, participated in the experiment. As predicted by theory, we find that the presence of a flexible market price prevents herding. However, the presence of contrarian behavior, which can (partly) be rationalized via error models, distorts prices, and even after 20 decisions convergence to the fundamental value is rare. We also report some interesting differences with respect to subjects' fields of study. Reassuringly, the behavior of the consultants turns out to be not significantly different from the remaining subjects.

JEL–classification: C92, D8, G1
Key words: informational cascades, herding, contrarians, experiment, internet.

*We are grateful to Jörg Breitung, David Easley, John Morgan, Georg Nöldeke, Erik Theissen, Georg Weizsäcker, and Anthony Ziegelmeyer for suggestions, and to 3 referees and a Co-Editor for extensive and very helpful comments. We also thank seminar participants in Amsterdam, Bonn, Frankfurt, Harvard, Mannheim, Munich, NYU, WHU, Zurich, the ES Winter Meeting 2004 in San Diego, the Royal Economic Society meetings 2003 in Warwick, the EEA meetings 2003 in Stockholm, the ESA meetings 2002 in Strasbourg, and the ENDEAR meetings 2002 in Jena for lively discussions and suggestions. We thank Frauke Lammers, Susanne Niethen, and Nina Wessels for very helpful discussions throughout the project, and McKinsey & Company for financial support. We also thank Deutsche Forschungsgemeinschaft, grant OE-198/1/1. Finally, we are very grateful to Ben Greiner who did an excellent job programming the experiment.

†This paper is part of Mathias Drehmann's PhD thesis and does not necessarily reflect the opinion of the Bank of England.

‡Corresponding author: Department of Economics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany, oechssler@uni-bonn.de

1 Introduction

While the popular press often blames "lemming–like" behavior of investors for the observed exuberance in financial markets, the question whether herding actually occurs in real financial markets, remains for various reasons very difficult to resolve with field data. When investors are observed taking the same action, this may be due to herding. But it may also be caused by investors following the same information. Or, the clustering of actions may simply be incidental. In empirical work, those explanations are difficult to disentangle since the private information of investors is, in general, unobservable. In order to contribute to this question, we conducted a large–scale internet experiment based on a sequential asset market as described by Avery and Zemsky (1998). An experiment offers an opportunity to directly test herding theories since all fundamentals and the private information of agents are under control of the experimenter.

Several sources of rational herding are known to the theoretical literature. For example, when market participants' payoffs depend directly on the behavior of others, herd behavior is natural. Such payoff externalities cause herding of analysts or fund managers in models of reputational herding (e.g., Scharfstein and Stein, 1990), or herd behavior of depositors in bank runs (e.g., Diamond and Dybvig, 1983).[1] Even if such payoff externalities are absent, however, herd behavior may be observed in markets through a process of information transmission. Models based purely on informational externalities were pioneered by Bikhchandani, Hirshleifer and Welch (1992) (henceforth BHW), Welch (1992), and Banerjee (1992). An *informational cascade* is said to occur when it becomes rational to ignore one's own private information and instead follow the predecessors' decisions. Since no further information is revealed once an informational cascade has started, inefficiencies occur even though each individual is behaving rationally.

Theories of rational herding or informational cascades are, however, not directly applicable to financial markets. Market prices are a powerful mechanism that, in theory, efficiently aggregates private information of traders. In particular, Avery and Zemsky (1998) (henceforth, AZ) have shown that informational cascades cannot occur in a simple sequential asset market because a flexible market price incorporates all publicly available information. As traders receive private information, they have an informational advantage over the market maker. Thus, rational

[1] Within our experiment we also implemented treatments designed to test the effects of reputation and payoff externalities in herding decisions. At the moment, these treatments are being evaluated and will be reported in a companion paper.

traders should always follow their private signal and thereby reveal their information.[2] Note that in this class of sequential trade models, traders are only allowed to buy or sell once, and therefore classical price bubbles driven by traders, who think they can resell the asset before the bubble bursts, are not possible.[3]

In reality, herding may nevertheless occur due to the likely existence of boundedly rational traders who may be plagued by a variety of biases and follow more or less plausible rules of thumb. Imitation, trend chasing, momentum trading strategies and the like are possible alternative sources for herd behavior in financial markets. Finally, there are strategies advocated by popular guide books and analysts that should counteract herd behavior.[4] In particular, "contrarian" or "value strategies" call for buying assets with low prices relative to some fundamental value like earnings, dividends, historical prices etc. (for empirical evidence on the profitability of such strategies, see e.g., Lakonishok et al., 1994, or La Porta et al., 1998).

The internet experiment we report on in this paper was designed to address the following objectives. First, we want to test the theory of informational cascades in financial markets as introduced by AZ by taking the theory at face value. To have the potential to reject their theory we implement a design that exactly matches their theoretical setup. The main test for this theory is whether herding is actually prevented by market prices. Second, we are interested in filtering out empirical regularities that may explain possible deviations from the theory. In particular, we want to find out whether traders follow their own signal (which is rational if all others are rational too), whether they engage in herd behavior, or whether they follow contrarian strategies by trading against their signal and the market. Third, we use the variety and size of our subject pool to detect possible differences in subjects' behavior as a function of their personal characteristics. Chief among those is the question whether businessmen behave differently from college students, who are usually used in experiments. For this purpose, we conducted a control experiment with 267 consultants from McKinsey & Company, an international consulting firm.

More than 6400 subjects participated in our experiment, in which a substantial amount of

[2]AZ show that herding may occur in the presence of multi–dimensional uncertainty, even though informational cascades remain impossible. See also Cipriani and Guarino (2001) who show that informational cascades are possible if agents are heterogeneous such that gains from trade exist.

[3]For a survey of theories of rational bubbles and herding, see e.g., Brunnermeier (2001). For a survey of experimental research on bubbles in asset markets, see e.g., Sunder (1995). For recent work in this area, see e.g., Hommes et al. (2002), or Hey and Morone (2004).

[4]See e.g., the investment classic *Contrarian Investment Strategies* by Dreman (1979, 1998) or his column "The Contrarian" in Forbes Magazine.

prize money was at stake. The subject pool was exceptionally educated with more than 13% holding a Ph.D. and another 30% being Ph.D. students. Almost half the subjects were educated in natural sciences, mathematics, or engineering. The main treatments in this experiment were variants of the basic model by AZ in which a market price aggregates all publicly available information. Traders received a private signal and could observe the past history of prices and in most treatments additionally the decisions of their predecessors. The large number of participants in the experiment allowed us to introduce a variety of modifications of the basic model. For example, we explored many different combinations of a priori probabilities and signal precisions to check for robustness.

As mentioned above, in our main treatments we aimed at exactly matching AZ's theoretical setup. Thus, when setting prices, the market maker (which in the experiment was played by the computer) assumed that all subjects behaved rationally. In the course of the experiment it emerged, however, that a substantial fraction of subjects did not follow their own signal. A human market maker in a real market would presumably learn this over time and adjust his pricing behavior accordingly.[5] As a consequence, we studied two additional "error" treatments, where the market makers' price setting rule was modified to account for the fact that subjects do not always adhere to their signals. The first of these treatments is in the spirit of the literature on noise traders. In this treatment the market maker assumed that, in addition to rational traders, there is a fixed proportion of noise traders, whose decisions are uninformative. In the second "error" treatment, building on the notion of quantal response equilibrium of McKelvey and Palfrey (1995, 1998), the market maker based his pricing behavior on the assumption that the propensity to follow one's own signal might depend on the history of previous decisions.

Finally, for comparison, we studied two benchmarks. First, we conducted treatments without market prices corresponding to the basic model of BHW. Second, another important benchmark are treatments in which subjects could not only observe the decisions of their predecessors but also their private signals. In those treatments, doubts about the rationality of others cannot be an issue.

Our experiment complements a large empirical literature with field data. Beginning with Lakonishok et al. (1992) researchers were analyzing the tendency of fund managers, security analysts (Welch, 2000), or investment newsletters (Graham, 1999) to herd (for surveys, see e.g.,

[5] We thank the Co-Editor Douglas Bernheim for pointing this out.

Bikhchandani and Sharma, 2000, Hirshleifer and Teoh, 2003, or Daniel et al., 2002). However, as Hirshleifer and Teoh (2003) note, it will always be difficult to empirically disentangle the mixture of reputational effects, informational effects, direct payoff externalities, and imperfect rationality.

Following Anderson and Holt (1997) there is by now a well established experimental literature on cascade and herding models.[6] However, to our knowledge there is only one other experiment on cascades in financial markets with flexible prices which has been conducted by Cipriani and Guarino (2004). Some of our treatments are very close to their experiment, and we will comment at several places in this paper on their results.[7]

Our paper is structured as follows. In the next section we first discuss the theoretical predictions for the basic BHW model without prices and then for the AZ model with market prices. In Section 3, we describe the experimental design, in particular the different treatments, the recruitment, the characteristics of the subject pool, and the implementation on the internet.[8]

The results of the experiment are presented in Section 4. Maybe the most important result is that we find no evidence of herding or imitative behavior in the presence of a flexible market price. While this aspect is consistent with the AZ model, other theoretical predictions of the AZ model find no support in the data. Recall that the AZ model predicts that all subjects follow their private information. In the experiment this happens only in between 50 % and 70 % of cases. Clearly, such behavior yields substantial deviations of actual prices from full information prices that would be obtained if everyone behaved rationally. We find that on average actual prices are less extreme than full information prices, which implies that volatility in the actual market is lower than it should be theoretically. While we do not observe herding, we find considerable support for the existence of "contrarian" behavior. When the price of asset A is high, subjects often buy asset B even if their own private information and the decisions

[6]See e.g., Allsopp and Hey (1999), Ashiya and Doi (2001) , Bounmy et al. (1998), Goidel and Shields (1994), Huck and Oechssler (2000), Hung and Plott (2001), Krämer et al. (2000), Kübler and Weizsäcker (2004), Nöth and Weber (2003), Nöth et al. (2001), Oberhammer and Stiehler (2001), Plott and Yang (1997), Sadiraj et al. (1999), Willinger and Ziegelmeyer (1998).

[7]Cipriani and Guarino (2004) conducted a paper-and-pencil experiment with a 50-70 probability combination and 12 traders in a group. When we conducted our experiment, we were not aware of their experiment. However, it is clear that their experiment has precedence. In our notation (see Table 1) they ran treatments that are comarable to *P+D, P-D, BHW,* and an additional treatment in which two subjects acted as market makers.

[8]Conducting experiments on the internet is still novel. For first experiences, see e.g., Forsythe et al. (1992, 1999), Lucking-Reiley (1999), Anderhub et al. (2001) , Charness et al. (2001), Shavit et al. (2001), Bosch–Domènech et al. (2002) , and Güth et al. (2002). For technical issues, see e.g., Greiner et al. (2002). The internet is also used to provide a platform to run economic experiments for interactive learning (Holt, 2002).

of their predecessors favor asset A, and vice versa. Since we find that contrarian behavior can be profitable at very low or very high prices, we explore the possibility that subjects have doubts about the rationality of others and consequently mistrust their decisions. We find that error models (like the quantal response models of McKelvey and Palfrey, 1995 and 1998), which explicitly take into account the possibility of mistakes, are partly able to rationalize contrarian behavior.

The large number of participants allows us to further conduct a number of interesting comparisons of behavior with respect to demographics, fields of studies etc.. There seems to be no significant difference between male and female subjects, or between subjects with and without college education. Subjects holding a Ph.D., however, are slightly more in line with theoretical predictions. Maybe it does not come as a surprise that when we look at selected fields of studies, physicists perform best in terms of "rationality" (i.e., performance according to theory) and psychologists worst. However, since "rational" behavior is only profitable when other subjects behave rationally as well, good performance in terms of "rationality" does not imply good performance in terms of profits. Indeed, the ranking in terms of profits is just the opposite: psychologists are best and physicists are worst. Finally, it is reassuring that the consultants in our control experiment did not behave significantly different from the subjects in the main experiment, which is important for the outside validity of our experiment. Section 5 contains a conclusion.

2 Theoretical predictions

Consider a number of investors who have to decide sequentially whether to invest in one of two assets, A or B. For simplicity each investor can only buy either one unit of asset A or one unit of asset B (sometimes we also allow for the possibility that no trade occurs). Investors are risk neutral and have the same a priori beliefs regarding the probabilities of success of the two investments. Specifically, only one asset is successful and worth 10 units at the end of the period while the other is worth 0. The successful asset is determined at the beginning of the experiment and each investors knows the a priori probability that asset S is successful, $P(S)$, $S \in \{A, B\}$.

The timing is as follows. Investors move sequentially in some exogenous order with each investor moving only once. Before deciding what to buy, each investor receives a private, informative signal a or b regarding the success of the assets. The signal's precision is $P(a|A) = P(b|B) > 0.5$, which is the conditional probability that signal $s \in \{a, b\}$ occures, given the true state is $S \in \{A, B\}$. For all investors, the signal is identically and independently distributed conditional on the true state. This is commonly known. Each investor can observe the decisions of all his predecessors.

We consider two principal versions of this model: one in which the price of the two assets is fixed (and set equal to zero for simplicity) and one in which the prices of A and B are market prices that reflect publicly available information. The version with zero prices is equivalent to the basic model studied by BHW. The model with market prices has been studied by AZ.

2.1 The BHW model

All investors can invest either in asset A or B – but not both – at zero cost.[9] Clearly, an investor with t predecessors will choose A if and only if the conditional probability that A is successful given all private and public information $P(A|H_t, s)$ is greater than $1/2$,[10] where H_t denotes the observable history of the decisions of all predecessors up to round t, and $s = a, b$ the private signal.

The difficulty lies in the interpretation of decisions of predecessors. Assuming that all predecessors are perfectly rational Bayesians, an investor who is a Bayesian himself follows his private signal and thereby reveals it, unless an informational cascade has started. If a signal can be deduced from the chosen action, it is called an imputed signal.

A cascade on asset S, an *S–cascade*, starts when an investor should buy asset S regardless of his own signal, i.e., when $P(S|H_t, s) > 1/2$, for $s = a, b$. Depending on the a priori probabilities and the signal precisions, this requires a certain number of (imputed) a or b signals. In all cases, however, the onset of a cascade depends only on the net number of signals $\Delta = \#a - \#b$ that can be imputed from the history of decisions (i.e., Δ is defined net of the signal of the current investor).

We demonstrate the calculations for our main probability combination 55-60, that is, $P(A) = 0.55$ and $P(a|A) = P(b|B) = 0.6$. The first investor should always follow his own signal since even if he receives a b-signal $P(B|b) = 1 - P(A|b) > 1/2$ holds. Hence, the signal of the first

[9] Introducing a fixed price would not alter the strategic setting.

[10] In most of our treatments, ties in expected profit cannot occur. When a tie–breaking rule is required, it is explicitly mentioned below.

player can be imputed from his action. If the first investor chooses A, the second should already disregard his own signal: even with a b signal, the second investor should choose A since

$$P(A|ab) = \frac{P(ab|A)P(A)}{P(ab|A)P(A) + P(ab|B)P(B)} = 0.55, \tag{1}$$

which is the a priori probability for A. The two signals a and b cancel out and the decision should follow the a priori probability. In this case the third player cannot impute the signal of the second player and thus faces a similar decision problem as the second player. Hence after one A, an A–cascade starts, i.e., when $\Delta \geq 1$. Likewise, it can be shown that a B–cascade must start for $\Delta \leq -2$. If all agents are rational Bayesians, a cascade is never broken once it started, and information accumulation stops (i.e., $-2 \leq \Delta \leq 1$ at all times unless errors occur).

In an experiment, one can hardly assume that all subjects are rational Bayesians, let alone that all subjects believe that all *other* subjects are rational. In particular, one has to make provisions for the fact that irrational behavior may be unambiguously observed (as when the second subject chooses B following an A by the first subject in the example above).

To account for possible non–Bayesian behavior, we assume that subjects impute signals in the following way. If there is no history of signals that can explain an observed action given that all predecessors behaved according to the Bayesian calculus explained above, we say that the action is in "obvious contradiction to Bayes' rule". If a decision is not in obvious contradiction to Bayes' rule, the imputed signal equals the decision unless a cascade has started, in which case no signal can be extracted from the respective decision. For a decision which, given the history of imputed signals, obviously violates Bayes' rule (i.e., if a cascade is broken), we considered two variants: (a) successors ignore the decision of the deviator, and (b) subjects assume that the deviator followed his private signal. As it turns out, the empirical truth lies somewhere in the middle but both assumptions yield qualitatively the same results. In the following, we only reports results based on rule (b). We say that an agent is "rational under common knowledge of rationality" (henceforth: *ruck*) if he follows Bayesian updating with respect to the imputed signal history and his own signal. In the following, it is important to remember that *ruck* is based on the assumption that all predecessors are rational. It may be perfectly rational to deviate from *ruck* if a subject thinks that his predecessors are not rational.

Of course, no ambiguity with regard to rationality arises when not only decisions but also signals of others are observable. In this case, an optimal decision is purely a matter of calculating conditional probabilities. As mentioned above, we also consider such a setting in some of our treatments.

2.2 The Avery/Zemsky model

To keep the experiment as simple as possible, we consider the simplest version of the AZ model (cf. Avery and Zemsky, 1998, Section I), which is the BHW model enriched by a flexible price. In this model, the price is set by a market maker who efficiently incorporates all publicly available information.[11] The crucial question is how the existence of a market price changes the possibilities for herding.

Let p_t denote the market price of asset A in round t and assume, as above, that a successful asset pays out 10 units in the end. Hence,

$$p_t = 10P(A|H_t). \tag{2}$$

The price of B is always equal to $10 - p_t$ since $P(A|H_t) = 1 - P(B|H_t)$.

The decision of an investor is straightforward. An investment in A is profitable in expectation if and only if

$$10P(A|H_t, s) - p_t > 0 \tag{3}$$

that is, if and only if $s = a$. Likewise, an investment in B is profitable if and only if $s = b$. In other words, each investor follows his private signal. All information is revealed, and therefore it is incorporated into the price immediately after each decision. This implies that the price is semi–strong efficient, i.e., at any point in time the price incorporates all publicly available information. The price is a martingale with respect to public information, i.e., $E(p_{t+1}|H_t) = p_t$ for all t, and one cannot take advantage of the knowledge of historical price movements to earn superior returns. As everyone follows his signal, rational herding cannot occur. Note that not trading is never optimal (unless one introduces transaction costs) because subjects always have an informational advantage over the market maker.

Again, the problem becomes more complicated when investors cannot be fully confident that their predecessors behaved rationally. Suppose instead the investor believes that (some) prior

[11] In contrast to AZ's general model which is in the spirit of Glosten and Milgrom (1985), the simple model does not incorporate uninformed traders and therefore has no bid–ask spread. With only informed traders, setting a bid–ask spread to ensure a zero profit condition for the market maker would lead to a market breakdown. However, the results of the simple model carry over to a more complex world with informed and uninformed traders and a market maker setting bid–ask spreads (see AZ, Proposition 3). As in the experiment the market maker was played by the computer, the possibility of losses was not an issue.

decisions were taken randomly. A *regression to the mean* argument implies that high prices are likely to be overvalued and low prices undervalued. Thus, it may pay for an investor to trade against the market and against his own signal. Such investors are called *contrarians* even though rational contrarians would never occur in our setting if all investors were known to be rational.

3 Experimental design

More than 6400 subjects participated in our online experiment which was available for a period of about 6 weeks in 2002 on our web site *http://www.a-oder-b.de* which is German for *a-or-b*.[12] Subjects decided in sequence and were able to observe the actual decisions of prior participants in their respective groups. In general, the group size was 20. Subjects were asked to make decisions in three independent groups. Thus, in total there were more than 19000 decisions. We call the first decision stage 1, the second stage 2, etc..

Common to all treatments are the following features. Subjects had to choose between investment opportunities A and B (in some price treatments, there was also the option of choosing neither which we label N). Only one of the two could be successful and, if so, would pay 10 "Lotto–Euros". The unsuccessful investment paid nothing. Subjects were told the a priori probabilities $P(S)$, which varied among our treatments. Furthermore, they were told that they would receive a tip by an investment banker which was reliable with a specified probability $P(s|S)$, which also varied among our treatments. Subjects were informed that all prior investors in their group had received a tip by other investment bankers and that these tips were independent of theirs (see Appendix A for a translation of the instructions). In the next subsection we introduce the details of the different treatments.

3.1 Treatments

Given the large number of participants we were able to explore a variety of different questions, information conditions, probability combinations, etc.. In this paper, we focus on treatments that are relevant to financial markets. That is, treatments which follow the basic setup of AZ where a market price exists which reflects all publicly available information. In two additional treatments, we deviate from the basic setup by assuming that subjects make errors with some

[12] Some follow-up treatments were run in January 2004.

probability, and that the market maker takes this into account. For comparison we also include two treatments without prices that follow the basic model of BHW.

Table 1 lists the main features of all treatments. Names of treatments with market prices start with P followed by *+D* when, additional to the price history, the decisions of all prior investors are observable, or *-D* when only the price history is observable. Hence, in treatments *+D* subjects could observe how decisions influenced prices. A *-N* denotes treatments in which the "no trade" option was absent, i.e., subjects were forced to buy either A or B.

Table 1: Treatments

treatment group	treatment	description
P	*P+D*	Price, all decisions observable
	P-D	Price, no decisions observable
	Pe+D	*P+D* + explicit formulation
	Pt+D	*P+D* + table to explain pricing rule
P-N	*P+D-N*	*P+D*, and N not possible
	P-D-N	*P-D*, and N not possible
	Pt+D-N	*P+D-N* + table to explain pricing rule
P-N+AS	*P+AS+D-N*	*P+D-N* + all signals observable
	P+AS-D-N	*P-D-N* + all signals observable
$P\text{-}N_{error}$	*Pf+D-N*	*Pt+D-N* + pricing rule with constant error
	$P\beta$*+D-N*	*P+D-N* + pricing rule with variable error
	BHW	Bikhchandani/Hirshleifer/Welch
	BHW+AS	*BHW* + all signals observable

Note: N denotes the option of not trading.

In our main treatments P and *P-N* we used the same pricing rule as in AZ. That is, prices were set to reflect all public information that was available at the respective points in time. The computer took the role of the market maker and assumed that all decisions were formed according to *ruck* (see (2)), i.e., that all subjects followed their own signal, unless this could be unambiguously rejected (which could only occur in case of an irrational no trade decision). In the latter case, no signal could be imputed, and hence the price remained constant (see Appendix B for a detailled description of how prices were formed in each of our treatments).

To explain the pricing rule to subjects is not a simple task. Therefore, we chose several different ways of presenting it. In the main treatment groups P and *P-N,* subjects were told in the instructions that "share prices are determined by supply and demand such that outside investors, who can observe the history of trades but not the tips given by the investment bankers, have no incentive to trade, i.e., an outside investor could not expect to profit from buying or selling one of the shares." In treatments indexed by an '*e*', subjects were told in addition that prices are conditional expected values given the history of decisions.[13]

In treatments indexed by '*t*', subjects were given a table instead of the explanation via outside investors. This table explicitly listed the price resulting from each possible net number of A–decisions by predecessors (i.e., the number of predecessors who chose A minus the number of predecessors who chose B). The instructions also contained 3 examples that demonstrated the way the price depended on the net number of A decisions.

The *P-N+AS* treatments were identical to the corresponding *P-N* treatments except that subjects (and the market maker) were also able to observe *all signals* of predecessors. In those treatments subjects did not need to worry about the rationality of their predecessors' actions.

If all agents act in line with theory, all of the above price treatments *P*, *P-N* and *P-N+AS* have the same theoretical prediction: everyone should follow his own signal.[14] Whether past decisions are observable or not is irrelevant since the price history fully reveals past decisions and the decision history yields no additional information. Furthermore, the no–trade option N should not alter the results because not trading is never optimal.

Up to now, we have assumed that the computer as market maker bases the price on the assumption that subjects follow their signal (as predicted by theory). However, after running most of the experiment, we observed substantial deviations from the theoretical predictions. A human market maker in a real market would presumably learn this over time and adjust his pricing rule.[15] To check whether the chosen pricing rule biased our results, we introduced two additional "error" treatments in which prices are set under the assumption that subjects sometimes deviate from the theoretical predictions. Treatment *Pf+D-N* is the same as *Pt+D-N* except that price calculations were based on the assumption that each subject chooses only with probability 0.6542 the asset that is consistent with her signal.[16,17] The second error treatment,

[13]We have also used two versions of the instructions, an old and a new version, the latter being more precise about the independence of the investment bankers' tips. See Section 4.8 for more on this.

[14]In *P-N+AS* this holds independent of the behavior of predecessors because the market maker is assumed to be able to observe the signals, and hence in these treatments actual prices equal full information prices (see Section 4.3).

[15]Interestingly, Cipriani and Guarino (2004) ran a treatment in which the market maker was played by subjects. They find no significant difference to the treatment in which the market maker followed the rule of Avery and Zemsky (1998).

[16]The probability of 0.6542 was based on the empirically observed frequency with which subjects followed their signal in treatments *P-N+AS* (in which there was no need to worry about rationality of others).

[17]Note that the market maker would arrive at the same prices if, alternatively, he would assume that the

Pβ+D-N, is similar to *Pf+D-N* except that, based on the concept of quantal response equilibrium of McKelvey and Palfrey (1995, 1998), the error probabilities were calculated separately for each history of decisions, which implied that error rates were higher for more extreme prices (see Appendix B for details as to how prices were calculated in these treatments).

In treatments $P\text{-}N_{error}$ it is rational to follow one's own signal if one assumes (which we shall do in the following) that subjects anticipate other subjects to make errors with the same probabilities that the market maker assumes when setting prices.

Finally, treatments without prices are denoted by *BHW*. The no–price treatment in which all signals of predecessors were observable is denoted by *BHW+AS.*

In the price treatments cascades should never happen regardless of the probability parameters of the model. In the no–price treatments the likelihood of cascades crucially depends on the a priori probability of the true state and the precision of the private signals. We have therefore looked at a number of different probability combinations shown in Table 2. For each probability combination, the last two columns in Table 2 give the minimum (maximum) net number of imputed signals necessary for the start of an A cascade (B cascade), respectively.

Table 2: Probabilities

a priori probability $P(A)$	signal precision $P(a\|A)$	Δ A–cascade	Δ B–cascade
55	60	1	−2
51	55	1	−2
55	80	1	−2
50*	66	2	−2
60*	60	1	−3
60	51	−9	−12
60	55	−1	−4

Note: * tie-breaking rule: follow own signal when indifferent.

Payoffs in "Lotto–Euros" were calculated as follows. If a subject chose the correct investment, he received 10 Lotto–Euros. This was the final payoff for this task in the *BHW* treatments. In the price treatments subjects received additionally an endowment of 11 for each task to avoid losses because they had to pay the market price for their investment (which could vary between 0 and 10). Thus, the payoff from each task was 11− market price + 10 (if successful).

decisions of a certain fixed proportion of subjects are uninformative (i.e., that there are noise traders), while the remaining subjects always follow their own signal. However, in our data at least it seems to be the case that the first interpretation is more appropriate (see Section 4.6 for details).

3.2 Recruiting and payment

The experiment was announced in several ads in the science section of the largest German weakly newspaper *Die Zeit,* two popular science magazines, and two national student magazines. Posters were distributed at most sciences faculties at German universities. Finally, emails were sent to Ph.D. students and postdocs in science and economics departments at 35 universities in Germany. The web site *www.a-oder-b.de* was linked to the Laboratory for Experimental Research in Economics at the University of Bonn and to the sponsor McKinsey & Company to demonstrate that the experiment had a proper scientific background and that the promised financial rewards were credible.

All payoffs in the experiment were denoted in "Lotto–Euro". Each Lotto–Euro was a ticket in a lottery to win one of our main prizes. In total there were 11 prizes of 1000 Euro each and 8 prizes of 100 Euros. The odds in those lotteries were fixed in advance and known to subjects. Each subject, when logging in on our website was told explicitly the odds per lottery ticket for winning one of our main prizes. Thus, maximizing the probability of winning one of the prizes was equivalent to maximizing the number of lottery tickets. All winners were notified by mail, and their prize money was paid through bank transfers.

In Phase I of the experiment, 1409 subjects played with high powered incentives where each of 40000 lottery tickets had an equal chance of winning one of 5 prizes of 1000 Euros. Since subjects played on average for about 15 minutes, they were making an expected hourly "wage" of 14.19 Euros, which is comparable to a very good student job and to pay in laboratory experiments. In Phase II, each of 90000 lottery tickets had an equal chance of winning one of another 5 prizes of 1000 Euros. Finally, in Phase III, 1162 subjects competed for the remaining 1000 Euros. Only in this Phase III of the experiment, where almost no monetary incentives were provided, subjects did not know how many lottery tickets were issued in the respective phase. This payment scheme was due to the fact that an unexpected large number of subjects participated in our experiment. But it also gives us the chance to test the role of incentives in such a setting.

Phase IV consists of experiments added in January of 2004, where the number of lottery tickets issued was again fixed and known to the subjects. In this Phase 320 subjects competed for 8 prizes of 100 Euros each, which amounted to an expected hourly "wage" of about 10 Euros.

Subjects is this phase played only treatment groups *Pt*, *P-N+AS,* and *P-N*$_{error}$.

Additionally, there was a control group of 267 consultants from McKinsey & Company who participated in the experiment on the same web site a couple of weeks before the start of the main experiment in 2002. The subjects of the control experiment were recruited by an internal email to all German consultants of McKinsey. Subjects knew that all other subjects were also consultants. About a third of those addressed participated. These subjects had the chance to win 8 vouchers for a nice dinner for two in a restaurant each worth 150 Euros.

3.3 Subject pool and implementation

In total, 6419 subjects finished our experiment of which 6152 subjects participated in the main experiment and 267 in the control experiment with consultants.[18] Of all subjects, 13.3% had a Ph.D. and 30.2% were Ph.D. students. All but 6.4% of the subjects had either finished a first university degree or were currently enrolled at a university. Average age was 28.3 and 28.1% of subjects were female. Considering the number of Ph.D. students and Ph.D.'s we believe we succeeded in recruiting a fairly bright subject pool. Furthermore, in contrast to most experiments in economics, our subjects came from a broad range of fields (about 50% were from natural sciences or engineering; others include business, economics, liberal arts, and law).

When arriving on our web site, subjects read a screen that introduced the general problem and the rules of the game. Subsequently, subjects were asked for some personal information, like name, mailing address, email, field of study, age etc.. Subjects were only allowed to play if all information requested was actually provided. This was also a measure to prevent subjects from playing twice: in order to win in the lottery, one had to give a correct mailing address, and the program ensured that the same name-postal code combination as well as the same email address could only play once. We also used cookies to prevent using the same computer twice.[19]

After entering the personal information, subjects were randomly placed in a currently active group,[20] and had to make their first decision. Afterwards they were randomly placed in another

[18] 835 individuals logged on but did not finish the experiment. Their decisions were not included in the history H_t since they did not face monetary incentives (payment was conditional on finishing all three stages of the experiment).

[19] It will never be possible to completely prevent clever people from playing more than once. However, we are confident that not many such attempts were successful, and given the size of the subject pool, those few probably do not matter much. It turned out that less than 5% of the participants entered an invalid email address (including unintentional typos), and these subjects did not behave significantly different (see Section 4.7).

[20] A group was *active* when it was neither full nor closed (i.e., when another subject was active in this group). We also ensured that subjects who logged on at about the same time were allocated to different treatments to

active group for the second task and then in a third group for the final task. No feedback about results was given until the subject had completed all three tasks. Even then they were only told how many "Lotto–Euro" they had won. Usually the tasks for each subject came from different treatments (except in Phase IV). Finally, we asked subjects for voluntary feedback as to how they formed their decisions, and 721 subjects sent response emails.

4 Results

For the evaluation of the results we shall consider the following 3 measures. (1) The variable *ruck* is meant to capture how well theory explains the data. In all treatments (except *P-N+AS* and $P\text{-}N_{error}$) average *ruck* is defined as the fraction of decisions that were rational under the assumption of common knowledge of rationality of predecessors.[21] (2) The fraction of cases in which subjects followed their own signal is denoted by *own.* (3) Finally, the actual market price p_t is compared to the *full information price* p_t^*, which would have resulted if the market maker could have directly observed the signals.[22]

Before we present the results it might be useful to collect the theoretical hypotheses for our various treatments (see the discussions in Sections 2 and 3.1 above). (1) In all price treatments subjects should follow their own signal ($ruck = own = 1$). (2) There should not be any difference between any of the treatments in *P* and *P-N*, regardless of whether prior decisions are observable or not (*+D* or *-D*) and whether the option N was available or not. (3) In treatment groups *P* and *P-N* actual prices p_t should match full information prices p_t^*. (4) The no trade option (N) should never be used. (5) In treatment *BHW,* subjects should follow the cascade behavior described in the last two columns of Table 2 *if* they believe that their predecessors are rational ($ruck = 1$). (6) In treatment *BHW+AS,* subjects should follow the cascade behavior of Table 2 *regardless* of what they believe about others ($ruck = 1$). (7) Different prior probabilities and signal precisions should not alter average *ruck*.

prevent "observational learning" in case two subjects sat next to each other in a computer pool.

[21]In the full information treatments *P-N+AS* common knowledge of rationality of predecessors is not required. In treatments $P\text{-}N_{error}$ average *ruck* is defined as the fraction of decisions that were rational given the assumption that the market maker's model of the behavior of one's predecessors was correct.

[22]In all but $P\text{-}N_{error}$ the full information price is equivalent to the theoretical price that would have resulted if all subjects had behaved according to *ruck.*

Table 3: Number of groups per treatment and probability combination

treatment	probability combination 50-66	51-55	55-60	55-80	60-51	60-55	60-60	total
P+D	12	6	26+15*	8	-	6	4	77
P-D	11	6	20+9*	1	-	-	3	50
Pe+D	12	6	12	6	-	6	-	42
Pt+D	-	-	9	-	-	-	-	9
P+D-N	16	6	18	-	-	-	-	40
Pt+D-N	-	-	9	-	-	-	-	9
P-D-N	17	6	18	-	-	-	-	41
P+AS+D-N	-	-	6	-	-	-	-	6
P+AS-D-N	-	-	6	-	-	-	-	6
Pf+D-N	-	-	9	-	-	-	-	9
Pβ+D-N	-	-	9	-	-	-	-	9
BHW	12	17	65+15*	8	12	6	15	150
BHW+AS	12	18	70+9*	2	12	-	12	135
total	92	65	277+48*	25	24	18	34	583

Note: As a convention the first number in x-y is the prior and the second the signal precision used in the experiment; *denotes groups in the control experiment with consultants.

4.1 Preliminary data analysis

Table 3 gives the number of groups we conducted in our experiment, separately for each combination of treatments and probabilities. As stated above, the group size was usually 20 (it was 10 for treatment *BHW+AS* and the control experiment with consultants). Since reporting results for each variation would be tedious, we checked first which variants of our treatments could be grouped together and pooled. We did this by comparing treatments with respect to the variables *ruck* and *own* both by non–parametric MWU tests and logit regressions taking each group as one observation. The following summarizes the results of those tests:

Incentives. The phase of the experiment (recall that incentives were different in Phases I through IV) had no significant effect. This implies that – at least in this experiment – incentives seem to matter little as compared to the intrinsic motivation of subjects to perform well.[23]

Stage. The stage of the task (whether a task was the first, second, or third a subject performed in) did not matter. This shows that learning effects do not play a significant role. This is in line with our expectations, given that subjects did not receive any feedback until the

[23] Camerer and Hogarth (1999) provide a survey of studies which look at the effects of monetary incentives.

end of the game.

Instructions. Presenting the pricing rule through a table (t) does not change the results as there are no significant differences between *P+D* and *Pt+D*, and between *P+D-N* and *Pt+D-N*, respectively. Also, a more explicit formulation for the price formation process (e) was irrelevant which, given that a large fraction of the subject pool has a mathematical background, is reassuring and indicates that subjects understood how prices were formed. The two versions of the instructions (see Footnote 13) made no significant difference.

Observability of decisions. The observability of decisions of predecessors did not matter which implies that subjects understood how their predecessors' decisions are reflected in the price history.

With the exception of $P\text{-}N_{error}$, we will, therefore, only report results on the treatment group level (see Table 1) because treatments within these treatment groups did not significantly differ from each other.[24] In the following, we will present the results from the main experiment. Results from the control experiment with consultants are reported in Section 4.8.

4.2 Summary statistics

In this subsection, we present summary statistics on the main variables of interest, *ruck* and *own*. Recall that theory predicts *ruck* = 1 for all treatments. Table 4 shows average *ruck* for 7 treatment groups and the probability combinations 55-60 and 50-66.[25]

Table 4: Average *ruck*

prob. comb.	*P*	*P-N*	*P-N+AS*	*Pf+D-N*	*Pβ+D-N*	*BHW*	*BHW+AS*
55-60	.54	.65	.65	.68	.70	.66	.72
	(.12)	(.11)	(.15)	(.10)	(.07)	(.11)	(.14)
50-66	.59	.68	-	-	-	.78	.76
	(.11)	(.09)				(.09)	(.12)

Note: Standard deviations in parentheses.

In the *P-N* treatments average *ruck* is 65% and 68%, respectively, i.e., roughly two-thirds of subjects in these treatments acted in line with theory. In the *P-N+AS* treatments and the two error treatments *Pf+D-N* and *Pβ+D-N* average *ruck* is between 65% and 70%.[26] None of

[24] Differences across treatment groups are discussed below.

[25] Since in the *BHW* treatments the observed *ruck* varies with different probability combinations, we compare treatments only for our main probability combinations 50-66 and 55–60. Table 15 in Appendix C displays the average values of *ruck* and *own* across all probability combinations.

[26] Binomial tests do not allow to reject the hypothesis that for any observed price the actual error probability in

the treatments without the N option (*-N)* are significantly different at any conventional level according to pairwise non–parametric MWU–tests. In the P treatments *ruck* is substantially and significantly lower, probably because an additional mistake can be made, namely not to trade.[27] In fact, N was chosen on average 19% of the time in the P treatments, which was never *ruck*.[28] Among subjects who did buy A or B in treatment P average *ruck* was similar to the -N treatments, namely 66% (71%) in probability combination 55-60 (50-66). Average *ruck* in *BHW* reaches 66% and 78%, respectively. Even in *BHW+AS,* where there is no need to worry about the rationality of others, *ruck* reaches only 72% and 76%, respectively.

A curious pattern appears with respect to the percentage of subjects who followed their own signals (*own*). In the 55-60 and 50-66 *BHW* sessions 75% of subjects (see Table 15 in Appendix C) follow their own signal (often in contrast to what they should do).[29] On the other hand, in the price treatments only between 54% and 70% of subjects follow their own signal, when following the own signal is always *ruck*.

Finally, we can take a more detailed look how *ruck* varies for different probability combinations. Table 5 shows *ruck* levels for treatment *P.* Taking our main treatment 55-60 as base, only 50-66 shows a significant difference at the 5% level according to MWU tests.[30] Hence, results from the price treatments seem to be fairly robust across probability combinations.

Table 5: *ruck* in treatment P across probability combinations

Prob. combinations	55-60	50-66	51-55	55-80	60-55	60-60
	.54	.59	.56	.53	.51	.60
	(.12)	(.11)	(.13)	(.11)	(.13)	(.10)

Note: Standard deviations in parentheses.

Pf+D-N is equal to the ex-ante assumed probability of 0.6542. In the price calculation in treatment $P\beta$*+D-N* we used an error rate $\beta = 0.5552$ (derived from treatment *P-N+AS*). Estimating such an error rate ex-post from the data of $P\beta$*+D-N* (see Section 4.6 and in particular Footnote 44) yields 0.471, which is not significantly different from 0.5552.

[27]Risk aversion cannot really account for this behavior since we employed a binary lottery procedure which – at least theoretically – should induce risk neutrality (see Roth and Malouf, 1979, or Berg et al., 1986).

[28]These findings are roughly in line with Cipriani and Guarino (2004). In their flexible price treatment (with probability combination 50-70) they report that 65% (22%) of subjects acted rationally (did not trade) which is comparable to our treatments P 50-66 where the respective numbers are 59% and 24%.

[29]In equilibrium (when all subjects are *ruck*) *own* should be about 60% (see again Table 15 in Appendix C).

[30]Differences across probability combinations look similar in treatment *P-N* but not all combinations were played for this treatment group.

4.3 Actual versus full information prices

The key question with respect to the efficient market hypothesis is whether prices accurately reflect the information in the market. Clearly, informational efficiency presupposes that individual traders act rationally on their information and, as we have seen in the last subsections, this is not always the case in our experiment. Consequently, in this subsection we want to look at how strongly prices are distorted relative to the full information benchmark. We define the *full information price* p_t^* as the price that would have resulted if the market maker could have observed the private signals.

While in the long–run the full information price will converge to the true value of the asset, this is not necessarily the case in the short or medium–run. To judge how well actual prices incorporate available information, we first look at convergence of final prices (i.e., prices after the decision of the last player) to full information final prices. Our data show that final prices are rarely close to the final full information price. In *P-N* the final price deviates by at most 20% from the final full information price in only 25% of groups (pooled over all probability combinations). The option of not trading in treatment *P* makes things worse since now only 18% of end prices deviate by at most 20% from final full information prices. In the error treatments *Pf+D-N* and *Pβ+D-N* convergence is even rarer as in both of these treatments only 11% of final prices are within 20% of final full information prices.

Table 6 depicts the ratio of the final full information price of asset A to the actual final price of asset A, (p^*_{T+1}/p_{T+1}), separately for states where asset A is successful and where asset B is successful. An important observation is that full information prices tend to be more extreme than actual prices. In other words, actual prices undershoot: when state A (B) is true, the actual final price of A is on average too low (high).

Table 6: Ratios of full information final prices to actual final prices

treatment	if A is successful	if B is successful
P	1.38	0.46
P-N	1.42	0.57
Pβ+D-N	0.98	0.40
Pf+D-N	1.11	0.51

Note: Pooled over all probability combinations

In a next step, in Figure 1 we consider the empirical distribution of actual prices (upper panels) and full information prices (lower panels) in treatments P-N and P-N_{error} pooled over all periods and probability combination 55-60.[31] In general, full information prices are concentrated around the a priori price. The distribution of actual prices matches this feature well but the distributions are even more concentrated. Extreme prices are rarely observed. It is interesting to see that even though a high fraction of subjects behave seemingly irrationally, the volatility in the actual prices is less than the volatility in the full information prices.

- Insert Figure 1 here -

That deviations between the full information and actual price can be severe is illustrated in the left panel of Figure 2 which shows the average deviation $(p_t^* - p_t)$ across full information prices p_t^* (not only final ones). The graph clearly indicates that actual prices undershoot and are less extreme than the theoretical prediction.

- Insert Figure 2 here -

Finally, over time deviations become more severe. The right panel of Figure 2 shows that the average relative distance between full information price and actual price increases.[32] Note that by construction of the price mechanism deviations in the early rounds must be small. More interestingly, there is no inverted U-shape form which would indicate convergence over time. In the next subsection we shall consider possible explanations for those deviations.

4.4 Possible explanations: imitation and contrarians

An obvious candidate explanation for observed deviations from the full information prices is herd behavior. As explained in Section 2.2, rational herding is impossible in our framework. However, if subjects indeed behave like "lemmings", they would imitate prior decisions and thus produce herd–like behavior. We find, however, that imitation plays no significant role.

The first evidence against imitation stems from inspecting Figures 1 and 2, which show that on average actual prices are less extreme than full information prices. Yet, imitation would predict the opposite. If, for example, an early investor in asset A induced later investors to

[31] Actual (full information) prices include the decision (signal) of the current subject. The figures for treatment P and for other probability combinations look similar. Treatment P-N+AS is omitted as in this treatment actual prices are equal to full information prices.

[32] The observed divergence over time may explain the higher levels of convergence reported by Cipriani and Guarino (2004) who consider groups of 12 subjects.

Table 7: Logit analysis: choice of A in price treatments

	P-N	*Pf+D-N*	*Pβ+D-N*	*P-N+AS*
price of A	-0.371^{**}	-0.837^{**}	-0.330^{*}	-0.377^{**}
	(0.040)	(0.323)	(0.137)	(0.068)
dhint	1.569^{**}	1.540^{**}	1.802^{**}	1.720^{**}
	(0.112)	(0.334)	(0.345)	(0.321)
pred	-0.016	0.068	0.132^{*}	-0.065
	(0.033)	(0.086)	(0.064)	(0.068)
Observations	1659	180	180	240
-2 Log-Likelihood	1985.31	216.60	210.43	269.77
R^2 (Nagelkerkes)	0.230	0.221	0.256	0.295

Note: ** significant at 1%-level; * significant at 5%-level; standard deviations in parentheses; all probability combinations used; all regressions include a constant and, in case of *P-N*, dummies for probability combinations.

buy A even though they got a b signal, this would quickly drive the actual price above the full information one. Likewise, an early investor in B should drive down the price of A more quickly than the full information price since imitation would yield more buyers of B than justified by private information.

To check more rigorously for imitation on the individual level, we analyze whether subjects' choices are influenced by the decisions of their direct predecessors. If a subjects is more likely to choose A after (say) three subjects immediately before him also chose A, we would take that as an indication for imitative behavior. Let *pred* measure the number of direct predecessors who chose an identical action. The variable is positive if this action was A and negative if the action was B.[33] In Table 7, we report results from logit regressions for all -*N* treatments in order to test whether the 0-1 variable "choice of A" is influenced by *pred*.[34] Besides *pred*, the following explanatory variables were used: (i) Price of A, (ii) *dhint*, which is a dummy variable that was 1 if the private information of the subjects in question was a and 0 otherwise, and (iii) a constant and, in case of *P-N*, dummies for the various probability combinations.

The price of A and the dummy for private information *dhint* have the expected signs and

[33] Alternatively, we included dummies if the number of A or B predecessors is greater than a certain threshold. This change was inconsequential.

[34] In treatments P we ran ordered logits with the endogenous variable taking the values 1 (choice of A), 0 (no trade), and -1 (choice of B), and obtained qualitatively the same results as in *P-N*. Since, from a theoretical point of view, N should never be chosen, it is not perfectly clear whether A, B and N can indeed be ordered. Hence, we also ran multinomial logits which also qualitatively yielded the same results. The independence of irrelevant alternatives underlying the multinomial model is fulfilled.

are significant. However, in all but one case, the coefficients for the imitation variable *pred* are not significantly different from zero. It seems that in the presence of a flexible market price, imitation cannot explain our data.[35]

If imitation is not the right story to explain the deviations between the full information and actual price, what is? We suggest a story based on contrarian behavior which can be justified by a regression to the mean argument. As explained above, we say that a subject is a *contrarian* if he trades against his signal and against the market, or equivalently if he receives an a (b) signal at a price for A which is strictly above (below) the a priori price (i.e., 10 times the a priori probability for A) and deviates from his signal to buy B (A) instead. Such contrarian behavior can only be (ex-post) rational if the trader is convinced that prior traders irrationally drove the price to an extreme.

Figure 3 gives evidence in favor of contrarian behavior. It shows average frequencies of *ruck* for treatment *P-N* and probabilities 55-60 given that the subject has received an a-signal. The higher the price, the more likely subjects are to trade against their signals and the lower *ruck*.[36] Average *ruck* drops from around 85% at low prices to around 50% at high prices. Contrarian behavior is also strongly suggested by Figure 2.

- Insert Figure 3 here -

These findings are confirmed in our regression analysis (see again Table 7). High prices for A significantly lower the probability of choosing asset A in all regressions.

4.5 Does it pay to be a contrarian?

While Section 4.4 provides evidence for contrarian behavior, the question arises whether it is a good idea to be a contrarian (at least in the treatments other than our benchmark treatments *P-N+AS*, where the answer is always no)? There is a paradox here since contrarian behavior is profitable if and only if prices are overvalued, but if there are many contrarians, prices are not overvalued. The current and the next subsection will address this issue.

First, we turn to the question whether contrarian behavior is actually profitable. A first aggregate look at the data suggests that this is not the case. Table 8 lists average profits

[35] A multinomial logit regression of treatment P confirms that there is also no imitation with respect to the no–trade decision N.

[36] Other treatments and probability combinations yield qualitatively similar figures (for both possible signal realizations).

(excluding the fixed payment of 11 in the price treatments) of subjects depending on whether subjects were *ruck* or not. In all cases, *ruck* yields higher average profits. The last column of Table 8 contains Pearson correlation coefficients between *ruck* and profits. All correlation coefficients are positive and significantly different from zero at the 1% (10%) level in treatments *P* and *P-N* (*Pf+D-N*).[37]

Table 8: Profits and *ruck*

	average profits of subjects classified as		
treatment	$ruck = 1$	$ruck = 0$	corr. coeff.
P	1.3	−0.5	0.21**
P-N	1.0	−0.4	0.13**
Pβ+D-N	0.1	−1.1	0.10
Pf+D-N	0.6	−0.8	0.13*

Note: Pooled over all probability combinations; profits exclude fixed payments; ** significant at 1%-level; * significant at 10%-level.

It seems plausible that contrarian behavior might be sensible if actual prices are extreme. To check this, we compare the profit of a contrarian with the *counterfactual* profit he would have received had he played according to theory. Since p_t =*(10- price of B)*, it is straightforward to show that the counterfactual payoff is always (-1) times the actual profit. If we pool over all probabilities and consider prices that deviate by more than two from the a priori price p_0, we get the results presented in Table 9. While in treatment *P* contrarians always do worse than their counterfactual rational subjects, in treatment *P-N* being a contrarian was profitable for prices that are more than two away from p_0.[38] In *P-N* for extreme prices a contrarian is actually more successful than a counterfactual, *ruck* subject would be.[39]

4.6 An error model to explain contrarian behavior

When can contrarian behavior be optimal? Let us assume that some traders behave like noise traders, in particular, they choose A and B more or less randomly. This implies that whenever the actual price is very high or very low, it is likely that this was driven by noise traders. In

[37]This also holds in *BHW* and *BHW+AS*. On average, *ruck* players earned 6.1 in *BHW* and *BHW+AS* whereas others earned 4.4 in *BHW* and 3.4 in *BHW+AS*.

[38]In $P\text{-}N_{error}$ there were hardly any observations of potential contrarians for extreme prices, which is why these treatments are not shown in Table 9. For intermediate prices contrarians earn negative profits in these treatments.

[39]One possible explanation for the difference between *P* and *P-N* is that prices in *P-N* have a higher variance (3.1 rather than 2.6 in *P* pooled over all probability combinations) which could make it more attractive to be a contrarian. We thank a referee for pointing that out to us.

Table 9: Contrarians' average profits in different price regions

	$p_t - p_0 \geq 2$	$2 \geq p_t - p_0 \geq -2$	$p_t - p_0 \leq -2$
P	-0.3	-0.7	-0.6
P-N	0.9	0.0	0.9

Note: Pooled over all probability combinations; profits exclude fixed payments; p_0 is the a priori price.

other words, whenever the actual price is extreme, the full information price is likely to be less extreme. Vice versa, whenever the full information price is extreme, the actual price is likely to be less extreme, which is simply a regression to the mean argument. Given that traders anticipate the random behavior of noise traders, they should be contrarians because a low price for A yields a buying opportunity even if the own private information is favoring B.[40] Given that there are some contrarians among the traders, actual prices will on average be less extreme than theoretical ones, just as observed in our data (see Figure 1).

As discussed in the introduction, in this subsection, we explicitly incorporate the possibility that subjects could be aware of the fact that others make mistakes. The error model is based on the concept of quantal response equilibrium of McKelvey and Palfrey (1995, 1998). In particular, it is assumed that subjects choose the correct action according to a logit function that depends on the difference in payoffs between the two alternatives.[41] If the difference is large, the correct action is chosen with high probability. If it is small, mistakes are more likely. Analogous to Anderson and Holt (1997), we estimate error parameters β_t for treatment P-N recursively for each round $t = 1, ..., 20$ taking into account that predecessors have reacted to possible earlier errors (for details of the estimation and for all coefficient estimates see Appendix D).[42] The coefficients for the β_t range between 0.58 and 0.08 and all but one of the 20 coefficients are significant at the 5% level.[43]

As it turns out, if one pools over all probability combinations average *ruck* in the error model (denoted by *ruck*-β) and *ruck* under the standard model are both equal to 0.67.[44] We

[40]This does not hold in the error treatments where the market maker, when setting the price, takes into account that not all subjects follow their signal.

[41]As in treatment P each subject has three possible choices we do not estimate an error model for this treatment.

[42]The above error model assumes a full level of reasoning. See Kübler and Weizsäcker (2004) for an estimation with various degrees of reasoning.

[43]In order to make our coefficients comparable to those of Anderson and Holt (1997) one has to multiply them by approximately 5 to account for the larger range in payoffs in our experiment (see Appendix D for details). Nevertheless, our estimates are somewhat smaller.

[44]We have analyzed two additional error models. First, in an iterative process, we estimated a constant parameter β across all rounds: in each iteration, expected profits were calculated by using the error parameter

believe that this is due to the fact that in our case errors do not change the optimal actions for most prices even though expected profits do change. This can be seen in the left panel of Figure 4 which shows average expected profits of buying asset A given an a signal depending on the assumptions with respect to errors of predecessors (treatment *P-N,* 55-60). In the standard model, where agents hold the belief that all agents follow their signal, it is always optimal to follow one's signal as well. Not so with the error model. The left panel of Figure 4 shows that at low or moderate prices for asset A, agents should optimally follow an a signal but at high prices contrarian behavior is optimal. Given the left panel of Figure 4, it is not surprising that average *ruck* at moderate prices is similar under both definitions of rationality. Only at very high and very low prices *ruck* increases in *P-N* if one allows for the possibility of errors (see the right panel of Figure 4).

- Insert Figure 4 here -

These results have implications for the discussion about "noise traders" in market microstructure models. As discussed earlier, in a model where there are only informed traders and where a market maker sets a bid-ask spread, the no–trade theorem would apply. To avoid this, most market microstructure models introduce noise traders ad hoc.[45]

First, our results suggest that "noise" seems to emerge automatically due to the irrationality of some of the traders. This becomes especially obvious in our *P-N+AS* treatments where, pooled over all probability combinations, average rationality is still only 65%. This observation together with the fact that even under the error model *ruck* does not exceed 70%, suggests that the problem is not so much extracting the relevant information from the decision of predecessors but rather processing it correctly. Second, our data provide support for the hypothesis that *each* agent decides rationally with a certain probability as opposed to the hypothesis that some of the subjects always decide rationally while others always decide irrationally. This can be seen by considering the 80 subjects who played three times *P-N+AS*, which does not require

obtained in the previous estimation, and we observed convergence of β after approximately 20 iterations. Second, we also analyzed a simpler error model in which the error probability was assumed to be constant, i.e., history-independent. The results from both of these modified error model are qualitatively the same as those reported above.

[45]To justify the presence of noise traders it has been argued that (rational) noise traders act due to liquidity or hedging needs (see e.g. Ausubel, 1990) or due to incentives arising from optimal delegation contracts for portfolio managers (see e.g., Dow and Gorton, 1997). Following Black (1986) there is a strand of the literature in which noise traders are seen as traders that trade on noise as if it was information or as agents who just act randomly.

assumptions with respect to the rationality of predecessors. Table 16 depicts how many of these subjects made zero, one, two, or three rational decisions. A two–sided Kolmogorov–Smirnov test reveals that this distribution is not significantly different (at any conventional level) from a distribution which would result if each of these subjects had always decided rationally with probability 0.6542, which is the average value of *ruck* in *P-N+AS*.

Table 10: Frequency of *ruck* when playing *P-N+AS* three times

	0	1	2	3
frequency in actual data	0.063	0.213	0.425	0.300
expected if rational with prob. 0.6542	0.041	0.235	0.444	0.280

4.7 The influence of personal characteristics

With the large number of participants in our experiment, it is possible to investigate the behavior of a variety of subgroups. For the price treatments, Table 11 compares average *ruck* of male and female subjects, subjects who are current or former students ("college"), subjects who have never attended college ("no college"), subjects holding a Ph.D., and current Ph.D. students. To test for differences, we ran logit regressions at the individual level to explain the rationality variable *ruck* by dummies for the above subgroups. Controlling for the duration of play, the age of the subject, and its position in the group, we find that neither the sex of the subject nor the college dummy are significant at conventional levels. However, there is evidence that the 13.3% Ph.D.'s in our subject pool have a significantly higher *ruck* (with a p–value of 0.068).[46]

Table 11: *ruck* of subgroups in price treatments

male	female	college	no college	Ph.D.	Ph.D. students	all
0.60	0.60	0.61	0.55	0.64	0.61	0.60

Note: Data from treatment groups *P, P-N, P-N+AS*, and $P\text{-}N_{error}$ pooled over all probability combinations.

As a large part of our subject pool had some university education, we asked subjects for their major field of study as it should be interesting whether there are differences in behavior. Figure 5 depicts average *ruck* and profits (excluding fixed payments) in the price treatments for

[46]Controlling for the field of studies reveals that this effect is mainly driven by Ph.D.'s from the sciences, medicine and engineering. Surprisingly, there is no significant Ph.D. effect in economics. We also included a dummy to reflect whether a subject had provided an invalid email address (see Footnote 19). This dummy is not significant.

selected fields of study. As expected, with respect to *ruck* physicists, but also economists, perform above average. Business students are slightly below average and psychologists performed worst. However, if one looks at profits, the ranking is almost exactly reversed. Physicists and economist do much worse than psychologists. A high *ruck* is not necessarily a virtue if others are irrational. This suggests that sometimes an intuition for the possibly irrational behavior of others seems to be more important than being able to apply Bayes' rule. While we should not overemphasize those results (recall that profits may vary from -10 to $+10$ and standard deviations are rather large), we do find one interesting behavioral difference between the two extreme groups, psychologists and physicists. Define a *potential contrarian situation* as a situation when the price for asset S is strictly higher than the a priori price of S and the subject receives a signal in favor of S, where $S \in \{A, B\}$. It turns out that in potential contrarian situations, psychologists act more often like actual contrarians than physicists do (in 56% of potential contrarian situations versus 37% for physicists). And this is not due to the fact that psychologists are more likely to trade against their signal in general. In fact, in non–contrarian situations, psychologists trade against their signal in only 14% of cases (versus 22% for physicists).

- Insert Figure 5 here -

4.8 Are consultants different?

A common (and justified) critique against experiments in economics is that, with few exceptions, they rely on a subject pool consisting only of economics students. While this can be changed with relatively little effort, the reliance on students in general is often dictated by financial and practical constraints. This makes it all the more important that outside validity is checked when one has the rare opportunity to conduct experiments with professionals in business. We were able to conduct the experiment on the same platform with 267 McKinsey consultants.[47] Do their results differ?

Originally we used a different text version in the control experiment.[48] To verify whether consultants behaved differently, all four combinations of general subject pool / consultants and old text / new text were played in treatment $P{+}D$ and probability combination 55-60 (which

[47]The groups size in this control experiment was always limited to 10 and sometimes 5. However, testing revealed that this has no significant effect.

[48]The new text was more precise about the independence of the investment bankers' tips. In the general phase, the two different text versions did not yield significantly different results, and hence the groups with old text were always included in the previous results.

Table 12: *ruck* of general subject pool and consultants

	general subject pool	consultants
old text	.53	.54
	(.09)	(.17)
new text	.56	.58
	(.14)	(.12)

Note: Treatment *P+D*; probability combination 55-60; standard deviations in parentheses; number of groups: general/old text: 8, consultants/old text: 9, general/new text: 18, consultants/new text: 6.

was the only probability combination played in the control treatment with consultants). Table 12 shows average *ruck* for these combinations. For both text versions the consultants have slightly higher values of *ruck* but none of the differences is significant at the 5% level of a MWU–test (two–sided). Thus, the control group of consultants does not behave significantly different from our general subject pool.

5 Conclusion

In this paper, we present results of a large scale internet experiment based on a sequential asset market with privately informed traders. Avery and Zemsky (1998) predict that in such markets herd behavior should not be observed because all trade decisions are immediately incorporated into the market price which, consequently, reflects all public information. And indeed, as predicted, we do not find evidence for herding or imitative behavior in our experiment. However, in contrast to theory, subjects do not always follow their private information but frequently act as contrarians, i.e., they trade against the market and their own signal. To explain this behavior, we study an error model which allows for the possibility that subjects have doubts about the rationality of others and consequently mistrust their decisions. The error model is able to rationalize contrarian behavior at relatively low or high prices. In fact, a very successful strategy in the price treatments would have been to follow one's signal for moderate prices and to be a contrarian for extreme prices. However, the fact that contrarian behavior occurs even in our error treatments *Pf+D-N* and *Pβ+d-N*, and in treatments *P-N+AS* suggests that it can not be fully explained by an optimal response to errors.

Our experiment complements a large empirical literature on herding. Our results confirm the finding of the bulk of this literature that herd behavior driven by informational externalities does not seem to be an important force in financial markets. To the contrary, one could even argue that the observed contrarian behavior, which we find sometimes to be profitable, has a stabilizing effect as it implies that agents tend to differentiate their investments from those of their predecessors. Of course, this does not rule out herding in financial markets based on explanations other than purely information based ones, as for example reputation concerns, payoff externalities, etc.. To disentangle these factors is an important task for future empirical and experimental research.

References

Allsopp, L. and J. D. Hey, "Two Experiments to Test a Model of Herd Behavior," *Experimental Economics*, 2000, *3(2)*, 121–136.

Anderhub, V., R. Müller, and C. Schmidt, "Design and Evaluation of an Economic Experiment Via the Internet," *Journal of Economic Behaviour and Organization*, 2001, *46(2)*, 227–47.

Anderson, L. R. and C. A. Holt, "Information Cascades in the Laboratory," *American Economic Review*, 1997, *87(5)*, 847–862.

Ashiya, Masahiro and Takero Doi, "Herd Behavior of Japanese Economists," *Journal of Economic Behavior and Organization*, 2001, *46*, 343–346.

Ausubel, L.M., "Partially-Revealing Rational Expectations Equilibria in a Competitive Economy," *Journal of Economic Theory*, 1990, *50*, 93–126.

Avery, Christopher and Peter Zemsky, "Multidimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, 1998, *88(4)*, 724–48.

Banerjee, A. V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 1992, *107(3)*, 797–817.

Berg, J. E., A. D. Lane, J. W. Dickhaut, and J. R. O'Brien, "Controlling Preferences for Lotteries on Units of Experimental Exchange," *Quarterly Journal of Economics*, 1986, *101*, 281–306.

Bikhchandani, S. and S. Sharma, "Herd Behavior in Financial Markets: A Review," *IMF Policy Working Papers WP/00/48*, 2000.

____ , D. Hirshleifer, and I. Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 1992, *100(5)*, 992–1026.

Black, Fischer, "Noise," *Journal of Finance*, 1986, *41(3)*, 529–543.

Bosch-Domenech, A., J. Montalvo, R. Nagel, and A. Satorra, "One, Two, (Three), Infinity, ...: Newspaper and Lab Beauty-Contest Experiments," *American Economic Review*, 2002, *92(5)*, 1687–1701.

Bounmy, K., J. Vergnaud, M. Willinger, and A. Ziegelmeyer, "Information Externalities and Learning with Sequential Interactions," in Cohendet, Llerena, Stahn, and Umbhauer, eds., *The Economics of Networks: Interaction and Behaviours*, Springer-Verlag, 1998, pp. 291–305.

Brunnermeier, Markus, *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding*, Oxford: Oxford University Press, 2001.

Camerer, C. and R. Hogarth, "The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework," *Journal of Risk and Uncertainty*, 1999, *19(1-3)*, 7–42.

Charness, G., E. Haruvy, and D. Sonsino, "Social Distance and Reciprocity: The Internet Vs. The Laboratory," *mimeo*, 2001.

Cipriani, Marco and Antonio Guarino, "Herd Behavior and Contagion in Financial Markets," *mimeo, New York University*, 2001.

___ **and** ___ , "Herd Behavior in a Laboratory Financial Market," *mimeo, University College London*, 2004.

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

Diamond, D. and P. Dybvig, "Bank Runs, Deposit Insurance and Liquidity," *Journal of Political Economy*, 1983, *91(3)*, 401–19.

Dow, J. and G. Gorton, "Noise Trading, Delegated Portfolio Management, and Economic Welfare," *Journal of Political Economy*, 1997, *105(5)*, 1024–1050.

Dreman, D., *Contrarian Investment Strategy*, New York: Random House, 1979.

___ , *Contrarian Investment Strategies: The Next Generation*, New York: Simon and Schuster, 1998.

Forsythe, R., F. Nelson, G. Neumann, and J. Wright, "Anatomy of an Experimental Political Stock Market," *American Economic Review*, 1992, *82*, 1142–1161.

___ **, T. Rietz, and T. Ross**, "Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets," *Journal of Economic Behavior and Organization*, 1999, *39*, 83–110.

Glosten, Lawrence and Paul Milgrom, "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 1985, *14(1)*, 71–100.

Goidel, R. K. and T. G. Shields, "The Vanishing Marginals, the Bandwagon, and the Mass Media," *The Journal of Politics*, 1994, *56(3)*, 802–810.

Graham, J.R., "Herding Among Investment Newsletters: Theory and Evidence," *Journal of Finance*, 1999, *54*, 237–268.

Greiner, B., H. Jacobsen, and C. Schmidt, "The Virtual Laboratory Infrastructure for Online Economic Experiments," *mimeo*, 2002.

Güth, W., C. Schmidt, and M. Sutter, "Fairness in the Mail and Opportunism on the Internet - a Newspaper Experiment on Ultimatum Bargaining," *German Economic Review, forthcoming*, 2003.

Hey, J. and A. Morone, "Do Markets Drive Out Lemmings or Vice Versa?," *Economica, forthcoming*, 2004.

Hirshleifer, D. and S.H. Teoh, "Herd Behavior and Cascading in Capital Markets: A Review and Synthesis," *European Financial Management*, 2003, *9(1)*, 25–66.

Holt, Charles, "Webgames and Strategy: Recipes for Interactive Learning," *mimeo*, 2002.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. Van de Velden, "Expectations and Bubbles in Asset Pricing Experiments," *CeNDEF Working Paper Series WP-02-05*, 2002.

Huck, S. and J. Oechssler, "Informational Cascades in the Laboratory: Do They Occur for the Right Reasons?," *Journal of Economic Psychology*, 2000, *21*, 661–671.

Hung, A. A. and C. R. Plott, "Information Cascades: Replication and an Extension to Majority Rule and Conformity Rewarding Institutions," *American Economic Review*, 2001, *91(5)*, 1508–20.

Krämer, C., M. Nöth, and M. Weber, "Information Aggregation with Costly Information and Random Ordering: Experimental Evidence," *mimeo*, 2000.

Kübler, D. and G. Weizsäcker, "Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory," *Review of Economic Studies*, 2004, *71(2)*, 425–441.

Lakonishok, J., A. Shleifer, and R. Vishny, "The Impact of Institutional Trading on Stock Prices," *Journal of Financial Economics*, 1992, *32*, 23–43.

___, ___, and ___, "Contrarian Investment, Extrapolations, and Risk," *Journal of Finance*, 1994, *49*, 1541–1578.

Lucking-Reiley, D., "Using Field Experiments to Test Equivalence Between Auction Formats: Magic on the Internet," *American Economic Review*, 1999, *89(5)*, 1063–1080.

McKelvey, R.D. and T.R. Palfrey, "Quantal Response Equilibria for Normal-Form Games," *Games and Economic Behavior*, 1995, *10*, 6–38.

___ and ___, "Quantal Response Equilibria for Extensive Form Games," *Experimental Economics*, 1998, *1*, 9–41.

Nöth, M. and M. Weber, "Information Aggregation with Random Ordering: Cascades and Overconfidence," *Economic Journal*, 2003, *113*, 166–189.

___, C. F. Camerer, C. R. Plott, and M. Weber, "Information Aggregation in Experimental Asset Markets: Traps and Misaligned Beliefs," *mimeo*, 2001.

Oberhammer, C. and A. Stiehler, "Does Cascade Behavior in Information Cascades Reflect Bayesian Updating? An Experimental Study," *mimeo*, 2001.

Plott, C., J. Wit, and W. Yang, "Parimutuel Betting Markets as Information Aggregation Devices: Experimental Results," *Economic Theory*, 2003, *22*, 311–351.

Porta, R. La, J. Lakonishok, A. Shleifer, and R. Vishny, "Good News for Value Stocks: Further Evidence on Market Efficiency," *Journal of Finance*, 1997, *52(2)*, 859–874.

Roth, A. and M. Malouf, "Game-Theoretic Models and the Role of Information in Bargaining," *Psychological Review*, 1979, *86*, 574–594.

Sadiraj, K., A. Schram, and S. Van Wijnbergen, "Rational Behavior Vs. Herd Behavior: Theoretical and Experimental Analysis," *mimeo*, 1999.

Scharfstein, David and Jeremy Stein, "Herd Behavior and Investment," *American Economic Review*, 1990, *80(3)*, 465–79.

Shavit, T., D. Sonsino, and U. Benzion, "A Comparative Study of Lotteries Evaluation in Class and on the Web," *Journal of Economic Psychology*, 2001, *22*, 483–491.

Sunder, Shyam, "Experimental Asset Markets: A Survey," in John Kagel and Alvin Roth, eds., *Handbook of Experimental Economics*, Princeton: Princeton University Press, 1995, pp. 445–500.

Welch, I., "Sequential Sales, Learning, and Cascades," *Journal of Finance*, 1992, *47(2)*, 695–732.

——, "Herding Among Security Analysts," *Journal of Financial Economics*, 2000, *58*, 369–396.

Willinger, M. and A. Ziegelmeyer, "Are More Informed Agents Able to Shatter Information Cascades in the Lab?," in Cohendet, Llerena, Stahn, and Umbhauer, eds., *The Economics of Networks: Interaction and Behaviours*, Springer, 1998, pp. 291–305.

Figures

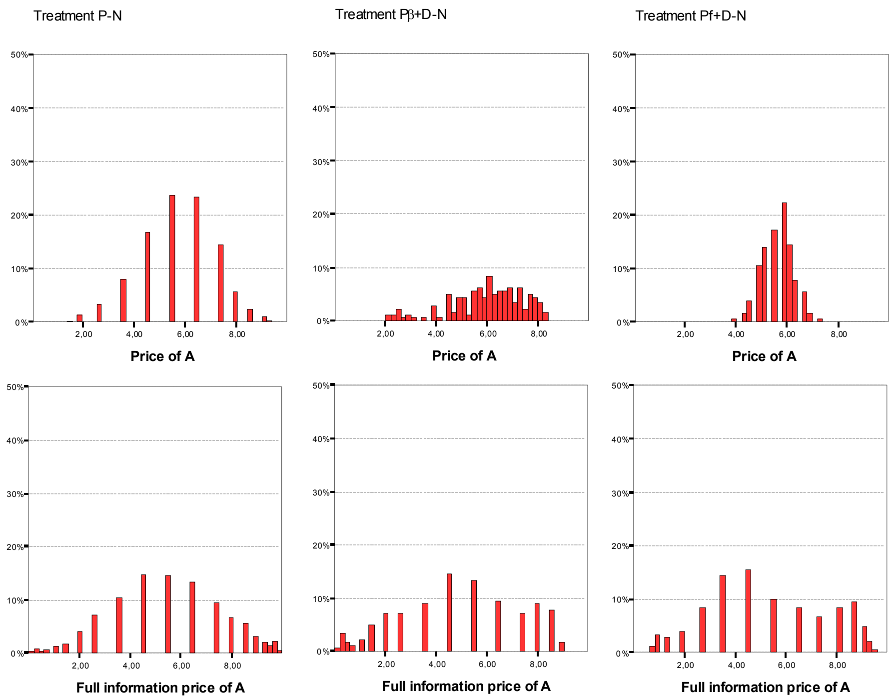

Figure 1: Distribution of actual prices and full information prices over all periods (treatments *P-N*, *Pβ+D-N* and *Pf+D-N*, probability combination 55-60)

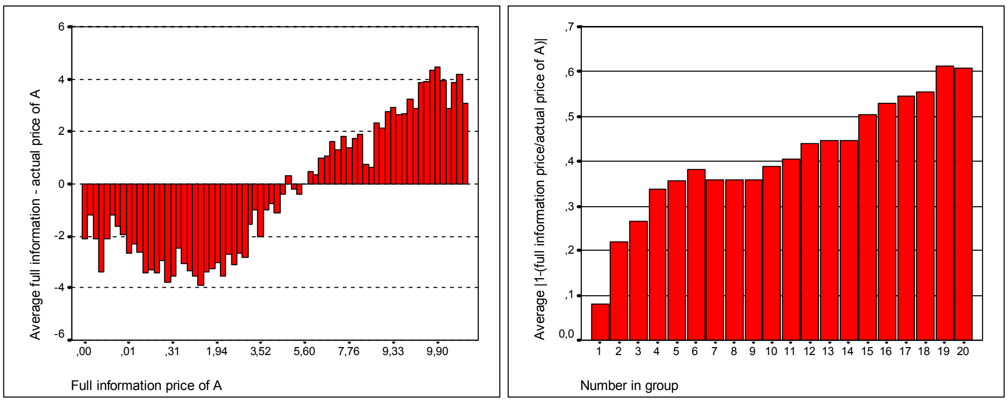

Figure 2: Average price deviation (full information minus actual price of A) given a full information price (left panel) and average relative distance between full information price and actual price of A given the number in the group (right panel) (treatment *P-N*, pooled over all probabilities)

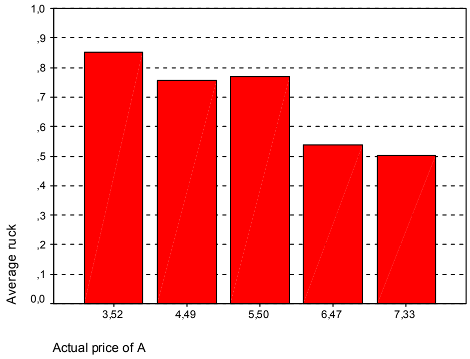

Figure 3: Average *ruck* given an a signal across actual prices of A (treatment P-N, probability combination 55-60, more extreme prices than those shown are omitted due to a low number of observations.)

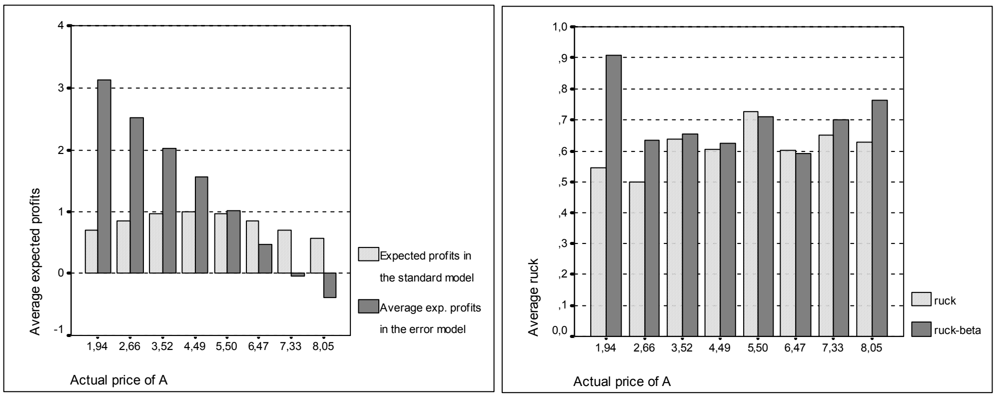

Figure 4: Average expected net profits of buying asset A given an a signal (left panel) and average ruck (right panel) in the standard model and in the error model across prices (treatment P-N, probability combination 55-60)

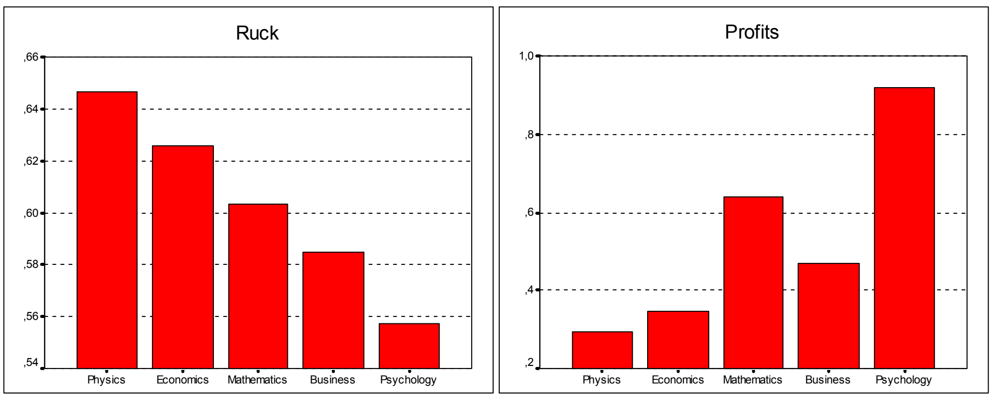

Figure 5: *ruck* (left panel) and profits (right panel) for selected fields of study (treatments P, P-N, P-N+AS, and P-N_{error}, all probability combinations, profits exclude fixed payments)

Appendix (to be published on our website only)

Appendix A: Instructions

Once connected to our website *www.a-oder-b.de*, there was first a general overview on the experiment (screen 1 below). Then, subjects where asked to provide some personal information (screen 2 below). Only if all information was provided, subjects were allowed to continue and learn their player number as well as the monetary incentives in the current phase of the experiment (screen 3 below). Note that the number of lottery tickets and the prizes mentioned below relate to Phase I of the experiment. Subsequently, the actual experiment began and Screen 4 below provides an example of the first of three stages (treatment BHW). Screen 5 below provides an example of a price treatment played in the second stage (treatment $P{+}D$). Stage 3 had the same basic structure, and therefore we omit an example of this stage.

Subjects also had at all times the option of opening a pop–up window that contained a summary of the main features of the setup. All phrases emphasized in this translation were also emphasized in the original web page.

At the end of this Appendix A, we additionally provide translations of the descriptions of the pricing rules in treatments *P-N+AS*, *Pt*, *Pf+D-N*, and $P\beta$*+D-N*.

Screen 1: Introduction

A game-theoretic experiment Are you a good decision-maker? We challenge you! Professor J. Oechssler together with the "Laboratorium for Experimental Research in Economics" at the University of Bonn aims to test various scientific theories through the online–experiment "A-or-B". Financial support is provided by the consultancy McKinsey & Company.

Attractive prizes By participating in the experiment you support the scientific work of the University of Bonn. At the same time you participate in a lottery for a total of 5,000 Euros which are distributed among 5 of the participants. The more thorough your decisions are, the greater your chances of winning. Of course you will also need some luck. The game takes approximately 15 minutes.

The experiment The experiment consists of three rounds. In every round you'll be assigned to a group and you - as well as every other member of your group - will have to take an investment decision. Without background knowledge the decision would be pure speculation. However, all players in a group will receive tips by investment bankers. Each group member gets a tip from a different investment banker. The investment bankers are experienced but can't make perfect predictions. The reliability of the tip is the same for every investment banker. As additional information, each player can observe the decisions of his predecessors in his group.

For each correct decision you will earn a predetermined amount of Lotto–Euros. After the third round, the Lotto–Euros you earned will be converted into lottery tickets on a one-to-one basis. Hence, the better your investment decisions, the higher your chances of winning. The experiment ends on June 7, 2002. The winners of the lottery will be notified after June 16, 2002 via ordinary mail. Now, let's begin the experiment!

Screen 2: Request of personal information

Welcome to the online-experiment "A-or-B". Please note that you can only play once. Before the game starts, we would like to ask you for some personal information. Of course, the results of the game will be kept separately from your personal information and will be analyzed anonymously. The mail address is only needed to notify the winners. Information on your field of studies, age, sex, etc. are only used for scientific purposes. Detailed information regarding data protection may be found here [Link].

[Data entry fields for last name, first name, address, email, student status, field of studies, year of studies, Ph.D. status, age and sex]

Screen 3: Player number and incentives

Thank you for providing the requested information. Your player number is: [player number]. Your player number, the number of lottery tickets you won, and additional information regarding the experiment will be automatically send to your email address after you have completed the experiment.

In this phase of the experiment, a total of 40,000 lottery tickets will be distributed, and 5 participants can win 1000 Euros each. Every lottery ticket has the same chance of winning.

Screen 4: Stage 1

You have to make an important investment decision: there are two risky assets (A and B). Only *one* asset will be successful and pay out *10 Lotto-Euros* (LE). The other asset will yield no profit at all. The successful asset was determined randomly before the first player of this group played. Hence, *the same asset* is successful for all players in your group. *Without additional information* you can rely on the fact that in *55%* of cases *asset A* is successful while in *45%* of cases *asset B* is successful.

Each participant in your group faces the same problem as you do: he has to choose between the assets and receives a tip from his respective investment banker. The reliability of the tips is the same for all investment bankers, and the tips of the investment bankers are independent of each other. The tip of each investment banker is correct in *60% of the cases*, i.e. in 100 cases where asset A (respectively B) is successful, in 60 cases the investment banker gives the correct tip A (respectively B) while in 40 cases the tip is not correct. The tip of your investment banker is: [B]

While each participant only knows the tip of his own investment banker, you - as every player in your group - can observe the decisions of the respective predecessors. Which players are assigned to which group is random and will differ from round to round. You are the [4th] investor in this group. One after another, your predecessors have made the following decisions:

Investor no.	1	2	3
Decision	B	A	B

What do you choose? [A] or [B].

Was the decision difficult? Independent of your decision, what do you think is the probability of A being the successful asset? [ ] %.
After the third round you'll find out whether your decision was correct. Let's move on the next round.

Screen 5: Stage 2

In this round you receive an endowment of 11 Lotto-Euro. The basic structure remains the same as in round 1. (In case you want to review the central features of round 1 please click [here].) This time you have to decide in which of the two risky shares (A or B) you want to invest. Only one share will yield a profit of *10 Lotto-Euro*, the other one will be worthless. *Share A* is successful in *55%* of all cases, *share B* in *45%* of all cases. As in round 1 the successful share was determined by chance before the first player of this group played.

In contrast to round 1, you - as every player in the group - have to pay the current share price if you decide to invest in a share. Share prices are determined by supply and demand such that outside investors, who can observe the history of trades but not the tips given by the investment bankers, have no incentive to trade, i.e. an outside investor could not expect to profit from buying or selling one of the shares [only in treatments *e*: ... because the price of share A (B) is equal to the conditional expected value of A (B) given the decisions of all your predecessors]. The role of outside investors is played by the computer.

As in round 1, every participant receives a tip from his investment banker which is correct in 60% of all cases. This time, your investment banker recommends: [A]

The current price of share A is 6.47 LE. The current price of share B is 3.53 LE. The profit in this round is given by:
your endowment (11 LE)
- price of the respective share
+ stock profit (10 or 0 LE)

You can also decide not to invest. In this case, you just keep your endowment of 11 LE.

Like your predecessors, you can *observe* the *price history* and the *history of decisions* in your group. You are the [2nd] investor in this group. Your predecessors in this group were facing the same problems as you, and one after another they have purchased the shares shown below at the price valid at that point in time:

Investor no.	1
Decision	A
Price of A	5.50
Price of B	4.50

What do you choose? [A], [B] or [No trade].

Independent of your decision, what do you think is the probability of A being the successful asset? [ ] %.
After the third round you'll find out whether your decision was correct.

Description of the respective pricing rules in other treatments

Recall that in the treatments to be described below (i) subjects did not have the option not to trade (and hence the fifth paragraph of Screen 5 above was omitted in these treatments), and (ii) these treatments were played in all three rounds.

Treatments *P-N+AS*

In the instructions for these treatments, subjects were notified in the sixth paragraph of Screen 5 above that also the history of tips was observable. In addition, the following replaced the second paragraph of Screen 5 above:

You - as every player in the group - have to pay the current share price if you decide to invest in a share. Share prices are determined by supply and demand such that outside investors, who can observe all the tips that your predecessors have received from their respective investment bankers (but cannot observe the tip you have received), have no incentive to trade, i.e., an outside investor could not expect to profit from buying or selling one of the shares. The role of outside investors is played by the computer.

Treatments *Pt*

In the *Pt* treatments the following replaced the second paragraph of screen 5 above:

You - as every player in the group - have to pay the current share price if you decide to invest in a share. To be precise, the share price depends on the difference (#A-#B) between the number of predecessors who bought A and the number of predecessors who bought B. The following table lists the share price for all possible differences (#A-#B):

$\#A - \#B$	0	1	2	3	4	5	6	7	8	9
share price A	5.50	6.47	7.33	8.05	8.61	9.03	9.33	9.54	9.69	9.79
share price B	4.50	3.53	2.67	1.95	1.39	0.97	0.67	0.46	0.31	0.21
$\#A - \#B$	0	-1	-2	-3	-4	-5	-6	-7	-8	-9
share price A	5.50	4.49	3.52	2.66	1.95	1.39	0.97	0.67	0.45	0.31
share price B	4.50	5.51	6.48	7.34	8.05	8.61	9.03	9.33	9.55	9.69

$\#A - \#B$	10	11	12	13	14	15	16	17	18	19
share price A	9.86	9.90	9.94	9.96	9.97	9.98	9.99	9.992	9.994	9.996
share price B	0.14	0.10	0.06	0.04	0.03	0.02	0.01	0.008	0.006	0.004
$\#A - \#B$	-10	-11	-12	-13	-14	-15	-16	-17	-18	-19
share price A	0.21	0.14	0.09	0.06	0.04	0.03	0.02	0.012	0.008	0.006
share price B	9.79	9.86	9.91	9.94	9.96	9.97	9.98	9.988	9.992	9.994

Three examples shall illustrate the above table. Suppose you were the fifth participant in a group.

If three of your predecessors chose share A and one share B, $\#A - \#B$ would be 2, in which case the price of share A would be 7.33, and the price of share B would be 2.67.

However, had three of your predecessors chosen share B and only had chosen share A, $\#A - \#B$ would be -2. The share price of A would be 3.52 and the price of share B would be 6.48.

In case two of your predecessors had chosen A and two had chosen B, $\#A - \#B$ would be 0. The share price of A would be 5.50 and that of share B would be 4.50. These are also the prices faced by the first participant of each group.

Treatment *Pf+D-N*

The pricing rule in this treatment was described in exactly the same way as in treatments *Pt* above with the only difference that the share prices in the table and in the examples were adapted to take account of the assumption that subjects followed their signal only with a certain probability (see Appendix B below).

Treatment $P\beta$*+D-N*

In treatment $P\beta$*+D-N* the following replaced the second paragraph of screen 5 above:

You - as every player in the group - have to pay the current share price if you decide to invest in a share. The share price is set by the computer. The computer sets the price in a way of a seller who maximizes his profit. It is assumed that the following holds:

1. The seller knows the above described basic structure of the situation.
2. When the seller sets the prices at which he is willing to trade, he can observe only the decisions of the predecessors in the respective group. In particular, the seller does not know any of the tips of the investment bankers.
3. However, the seller has ample experience. Based on the behavior of participants in similar earlier experiments he is trying to estimate which influence the history of prices, the earlier decisions of participants and the tips of investment bankers have on the decisions of participants.
4. The seller is facing competition by many other sellers who possess the same properties as himself, and who also want to trade in shares A and B.

Appendix B: Pricing rules

In this appendix we explain in more detail how prices were set in each of our price treatments. The computer served as market maker. Prices were set based on the observable history of decisions – not based on the history of signals (except in treatment group *P-N+AS* where prices were based on the history of signals since they *were* observable in this treatment group). At the end of this example we provide examples to illustrate the various pricing rules. In all treatments the price of A encountered by the first player was $10 \cdot P(A)$.

1. *P* and *P-N*: In all treatments in these treatment groups it was assumed that all observed decisions were formed according to *ruck*, i.e., that all subjects followed their own signal. The only possible observable deviation from this is a no–trade decision in treatment *P* (in violation of *ruck)*. In this case the out-of-equilibrium belief was that a deviation could have been committed by subjects with either signal. Hence, no signal could be imputed, and the price remained unchanged. Given this assumption, it is straightforward to show (by a simple application of Bayes' rule) that the price p_t only depends on the net number n_t^D of A decisions up to date t, i.e., the number of decisions to buy A minus the number of decisions to buy B observed by the market maker:

$$p_t \equiv 10P(A|H_t) = 10P(A|n_t^D) = \frac{P(A)}{P(A) + (1 - P(A)) \cdot P(a|A)^{-n_t^D} \cdot (1 - P(a|A))^{n_t^D}}. \quad (4)$$

2. *P-N+AS* and full information price p_t^*: In both treatments in treatment group *P-N+AS* it was assumed that the market maker can observe the history of signals. In analogy to the reasoning above, the price in this case only depends on the net number n_t^{AS} of a signals up to date t, i.e., the number of a signals minus the number of b signals observed by the market maker:

$$p_t \equiv 10P(A|n_t^{AS}) = \frac{P(A)}{P(A) + (1 - P(A)) \cdot P(a|A)^{-n_t^{AS}} \cdot (1 - P(a|A))^{n_t^{AS}}}. \quad (5)$$

   Our theoretical benchmark (the full information price p_t^*) is calculated in exactly the same way, and hence does not depend on the specific treatment under consideration.

3. *Pf+D-N*: In contrast to treatment *P-N* in both error treatments the market maker and subjects take into account that subjects may make mistakes. In *Pf+D-N* the price was set as in treatment *P-N* with the only difference that it was assumed that subjects followed their signal only with probability 0.6542 (which is the empirical frequency of *own* in treatment group *P-N+AS*). Given this assumption it is straightforward to show that:

$$p_t \equiv 10P(A|H_t) = 10P(A|n_t^D) = \frac{P(A)}{P(A) + (1 - P(A)) \cdot \widehat{P}(a|A)^{-n_t^D} \cdot (1 - \widehat{P}(a|A))^{n_t^D}}, \quad (6)$$

   where $\widehat{P}(a|A) \equiv 0.6542 \cdot P(a|A) + (1 - 0.6542) \cdot (1 - P(a|A))$.

4. *Pβ+D-N*: Again, in this treatment the market maker and subjects take into account that subjects do not always follow their signals. However, in contrast to *Pf+D-N*, in the present treatment the market maker does not assume that the probability of following one's signal is independent of the history of decisions. Rather, based on the concept of quantal response equilibrium of McKelvey and Palfrey (1995, 1998), the market maker assumes that, for a given signal, the probability with which a subject chooses to buy asset

A is given by a logit function that depends on the difference in payoffs between the two alternatives (A or B):

$$P(D_i = A|H_t, s_i) = \frac{1}{1 + e^{-\beta(\pi_i^A - \pi_i^B)}}, \tag{7}$$

where it is assumed that, due to his experience in market making, the market maker knows the error rate β (for as to how β was derived see below). Hence, if the payoff difference is large, the correct action is chosen with high probability. If it is small, mistakes are more likely.

Pricing rule: Given (7), the market maker sets prices in the following way. For the first player the market maker states a price equal to $10 \cdot P(A)$. Since for the first player the payoff difference from choosing A rather than B only depends on $P(A)$ and the realization of the signal, the calculation of the price that the second player faces is straightforward: the market maker uses (7) to calculate $P(s_1 = a|H_1)$, then calculates $P(A|H_1)$ (see (9) below), and sets the price equal to $p_2 = 10P(A|H_1)$.

$P(A|H_1)$ is then taken as the prior for the second player. Using this prior and the second player's expected payoff difference $\pi^A - \pi^B$ as a function of the second player's signal, the market maker calculates a new $P(A|H_2)$ and a new price $p_3 = 10P(A|H_2)$. $P(A|H_2)$ is then the prior for the third player, and so on.

The error rate β was estimated from the data of treatments *P-N+AS* where subjects did not need to worry about the behavior of their predecessors. As in these treatments the expected payoff differences $\pi_i^A - \pi_i^B$ only depended on the (observable) signal history, for all players, we calculated the expected payoff differences from the signal histories and estimated (7) with a logit regression. We obtained $\beta = 0.5552$ (standard deviation = 0.09), and based the price calculations in treatment *Pβ+D-N* on this value.

Examples

To illustrate the pricing rules Tables 13 and 14 contain examples how prices were calculated for a given history of up to 5 decisions and signals in our main probability combination 55-60.

Table 13: Price formation: an example without the no-trade option

Investor #	1	2	3	4	5	6	
Signal	a	a	b	a	b	a	
Decision	A	B	B	A	A	B	
P-N	5.50	6.47	5.50	4.49	5.50	6.47	5.50
P-N+AS	5.50	6.47	7.33	6.47	7.33	6.47	7.33
Pf+D-N	5.50	5.80	5.50	5.19	5.50	5.80	5.50
Pβ+D-N	5.50	5.99	5.52	5.03	5.53	6.03	5.56
Full information price p_t^*	5.50	6.47	7.33	6.47	7.33	6.47	7.33

Note: Probability combination 55-60.

Table 14: Price formation: an example with the no-trade option

Investor #	1	2	3	4	5	6	
Signal	a	a	b	a	b	a	
Decision	A	N	B	A	A	N	
P	5.50	6.47	6.47	5.50	6.47	7.33	7.33
Full information price p_t^*	5.50	6.47	7.33	6.47	7.33	6.47	7.33

Note: Probability combination 55-60.

Appendix C: Average *ruck*, *own*, and *own**

The following table displays average values of *ruck* and *own* for all employed probability combinations. For treatments *BHW* and *BHW+AS*, as a benchmark, the table additionally displays the average values of *own**, which is the fraction of subjects that would have followed their own signal if all subjects had behaved rationally.

Table 15: Average *ruck*, *own*, and *own**

prob.	P	P-N	P-N+AS	Pf+D-N	Pβ+D-N		BHW			BHW+AS	
comb	ruck	ruck	ruck	ruck	ruck	ruck	own	own*	ruck	own	own*
55-60	.54	.65	.65	.68	.70	.66	.75	.59	.72	.74	.68
	(.12)	(.11)	(.15)	(.10)	(.07)	(.11)	(.13)	(.12)	(.14)	(.14)	(.13)
50-66	.59	.68	–	–	–	.78	.75	.62	.76	.69	.69
	(.11)	(.09)	–	–	–	(.09)	(.12)	(.16)	(.12)	(.12)	(.18)
51-55	.56	.71	–	–	–	.69	.73	.56	.72	.77	.64
	(.13)	(.11)	–	–	–	(.08)	(.12)	(.13)	(.13)	(.15)	(.14)
55-80	.53	–	–	–	–	.78	.81	.70	.83	.73	.79
	(.11)	–	–	–	–	(.09)	(.07)	(.14)	(.24)	(.09)	(.02)
60-51	–	–	–	–	–	.84	.57	.50	.77	.62	.49
	–	–	–	–	–	(.10)	(.13)	(.12)	(.14)	(.16)	(.15)
60-55	.51	–	–	–	–	.78	.62	.53	–	–	–
	(.13)	–	–	–	–	(.09)	(.13)	(.10)	–	–	–
60-60	.60	–	–	–	–	.69	.69	.59	.73	.63	.68
	(.10)	–	–	–	–	(.11)	(.11)	(.14)	(.16)	(.17)	(.23)

Note: Standard deviations in parentheses; own* denotes the fractions of subjects that would have followed their own signal if all subjects had behaved rationally; in all treatments except *BHW* and *BHW+AS* we have *ruck*=*own* and *own**=1 by definition.

Appendix D: Details of estimation of the error model

The estimation is done in the following way. It is assumed that subjects decide according to a logistic model with independent shocks to the expected payoff difference between assets A and B. For reasons of tractability, we only estimate the error model for treatment group *P-N*. Formally, the probability that subject i decides to buy asset A (which we denote by "$D_i = A$") is given by

$$P(D_i = A|H_t, s_i) = \frac{1}{1 + e^{-\beta_t(\pi_i^A - \pi_i^B)}}, \tag{8}$$

where π_i^S is the expected profit of buying asset $S \in \{A, B\}$. The parameter β_t characterizes the sensitivity to payoff differences. Subjects buy randomly if $\beta_t \to 0$ but play rational best replies if $\beta_t \to \infty$. Since the expected profits of the first player only depend on the realization of the signal, the estimation of β_1 is straightforward with a logit regression. The estimation of subsequent error parameters is more involved in that expected profits in a certain round depend on the error parameters of all previous rounds which implies path-dependency.

To estimate β_2, we first calculate the probability $P(D_1 = S|s)$ that in round 1 the subject chose asset S in case he received signal $s \in \{a, b\}$ taking the error parameter β_1 into account. In a second step, this information can be used to calculate

$$P(D_1 = A|A) = P(D_1|a) \cdot P(a|A) + P(D_1|b) \cdot P(b|A). \tag{9}$$

Hence, if $D_1 = A$, the posterior that asset A is successful is given by

$$P(A|D_1 = A) = \frac{P(D_1 = A|A) \cdot P(A)}{P(D_1 = A|A) \cdot P(A) + P(D_1 = A|B) \cdot P(B)}. \tag{10}$$

Combining this with private signals, one can calculate expected profits for second round players. With those, β_2 can be estimated yielding a new prior for player 3, and so on for all subsequent rounds.

Below we provide our estimates of the β_t's. As it is interesting to compare our results to those of Anderson and Holt (1997), note that the level of the estimated coefficients depends on the payoff resulting from a correct decision. To illustrate this, assume that p denotes the probability of asset A being successful. In Anderson and Holt (1997) a correct decision yielded a payoff of 2, and hence the expected payoff difference was given by $\pi^A - \pi^B = 2(p - (1-p))$. In our experiment, besides the payoff resulting from a correct decision (which was equal to 10), the expected payoff difference depended also on the current market price of A. Hence, the expected payoff difference was given by $\pi^A - \pi^B = 2[p(10 - p_A) - (1 - p)p_A]$. Given that the actual price in treatment *P-N* was on average equal to 5.44, our estimates have to be multiplied by (approximately) 5 to make them comparable.

Table 16: Coefficients of error model

	coefficient	standard deviation
β_1	.58**	.12
β_2	.08	.09
β_3	.52**	.13
β_4	.31**	.10
β_5	.30**	.09
β_6	.27**	.09
β_7	.55**	.13
β_8	.41**	.11
β_9	.23*	.09
β_{10}	.37**	.11
β_{11}	.32**	.10
β_{12}	.24*	.10
β_{13}	.35**	.10
β_{14}	.31**	.11
β_{15}	.43**	.10
β_{16}	.19*	.09
β_{17}	.27**	.09
β_{18}	.39**	.10
β_{19}	.37**	.10
β_{20}	.35**	.10

Note: ** indicates significance at the 1% level; * at the 5% level.