On the trend recognition and forecasting ability

of professional traders

Markus Glaser, Thomas Langer, and Martin Weber*

April 15, 2003

Abstract

Empirical research documents that temporary trends in stock price movements exist. Moreover, riding a trend can be a profitable investment strategy. Thus, the ability to recognize trends in stock markets influences the quality of investment decisions. In this paper, we provide a thorough test of the trend recognition and forecasting ability of financial professionals who work in the trading room of a large bank and novices (MBA students). In an experimental study, we analyze two ways of trend prediction: probability estimates and confidence intervals. Subjects observe stock price charts, which are artificially generated by either a process with positive or negative trend and are asked to provide subjective probability estimates for the trend. In addition, the subjects were asked to state confidence intervals for the development of the chart in the future. We find that depending on the type of task *either* underconfidence (in probability estimates) *or* overconfidence (in confidence intervals) can be observed in the same trend prediction setting based on the same information. Underconfidence in probability estimates is more pronounced the longer the price history observed by subjects and the higher the discriminability of the price path generating processes. Furthermore, we find that the degree of overconfidence in both tasks is significantly positively correlated for all experimental subjects whereas performance measures are not. Our study has important implications for financial modelling. We argue that the question which psychological bias should be incorporated into a model does not depend on a specific informational setting but solely on the specific task considered. This paper demonstrates that a theorist has to be careful when deriving assumptions about the behavior of agents in financial markets from psychological findings.

Keywords: trend recognition, forecasting, conservatism, overconfidence, professionals, financial modelling

JEL Classification Code: C9, G1

On the trend recognition and forecasting ability

of professional traders

Abstract

Empirical research documents that temporary trends in stock price movements exist. Moreover, riding a trend can be a profitable investment strategy. Thus, the ability to recognize trends in stock markets influences the quality of investment decisions. In this paper, we provide a thorough test of the trend recognition and forecasting ability of financial professionals who work in the trading room of a large bank and novices (MBA students). In an experimental study, we analyze two ways of trend prediction: probability estimates and confidence intervals. Subjects observe stock price charts, which are artificially generated by either a process with positive or negative trend and are asked to provide subjective probability estimates for the trend. In addition, the subjects were asked to state confidence intervals for the development of the chart in the future. We find that depending on the type of task *either* underconfidence (in probability estimates) *or* overconfidence (in confidence intervals) can be observed in the same trend prediction setting based on the same information. Underconfidence in probability estimates is more pronounced the longer the price history observed by subjects and the higher the discriminability of the price path generating processes. Furthermore, we find that the degree of overconfidence in both tasks is significantly positively correlated for all experimental subjects whereas performance measures are not. Our study has important implications for financial modelling. We argue that the question which psychological bias should be incorporated into a model does not depend on a specific informational setting but solely on the specific task considered. This paper demonstrates that a theorist has to be careful when deriving assumptions about the behavior of agents in financial markets from psychological findings.

Keywords: trend recognition, forecasting, conservatism, overconfidence, professionals, financial modelling

JEL Classification Code: C9, G1

1 Introduction

Empirical research documents that temporary trends in stock price movements exist.[1] Moreover, riding a trend can be a profitable investment strategy.[2] Thus, the ability to recognize trends in stock markets influences the quality of investment decisions. This argument is in line with Shiller (2001) who states that "ultimately, people who choose asset allocations must use their subjective judgment about the probability that stock trends will continue".[3] The same idea can also be found in many finance models. These models assume that investors' expectations of future prices or their asset demand functions are influenced by past price movements. Some of these models are discussed in Section 2.

In this paper, we provide a thorough test of the trend recognition and forecasting ability of lay people as well as professional traders.[4] Gaining insights about how professional traders form beliefs is particularly important as financial market outcomes are mainly driven by transactions of institutional investors.[5] By choosing an experimental approach for this investigation we are able to isolate the effects of past prices on predictions. Such an analysis would be more difficult in a real market setting, where other factors like trading volume or macroeconomic conditions might play a role as well.

In our experiment, traders observe artificially generated stock price charts which have in

---

[1]See, for example, Jones and Litzenberger (1970) or Bernard and Thomas (1989).

[2]One example is the profitability of momentum strategies in which stocks with high (low) returns over the last three to 12 months are bought (sold short). Jegadeesh and Titman (1993, 2001) showed for US stocks that this strategy results in significant positive profits. This strategy has been successful in other stock markets as well (see Rouwenhorst (1998), Rouwenhorst (1999), and Glaser and Weber (2003) for international evidence on the profitability of momentum strategies).

[3]Shiller (2001), p. 3.

[4]In the following we use the term trend prediction to refer to the two concepts trend recognition and forecasting.

[5]See Nofsinger and Sias (1999), Lakonishok, Shleifer, and Vishny (1992), and Gompers and Metrick (2001).

the long run either a positive or a negative trend. This two state scenario is similar to the situation professionals face when judging whether the market is currently bullish or bearish. We use two natural ways of trend prediction: estimates of the probability for either trend and forecasts of the future price development via confidence intervals. How individuals perform in such tasks has been extensively studied in the psychological literature. In probability estimate tasks, conservatism or underconfidence is often observed, i.e. the stated probabilities are less extreme than they should be (see Edwards (1968)). In the confidence interval tasks, overconfidence is usually observed, i.e. confidence intervals are too tight (see the survey of the calibration literature by Lichtenstein, Fischhoff, and Phillips (1982)). However, most psychological studies use only students as experimental subjects. Our study allows to test whether professional traders are as prone to these biases as students or whether expertise moderates deviations from rationality. In contrast to the above mentioned psychological studies, we can furthermore analyze intrapersonal differences in the strength of possible biases as we embed the two different tasks in the same informational setting.

Our main results can be summarized as follows. Depending on the type of task *either* underconfidence (in probability estimates) *or* overconfidence (in the calibration tasks) can be observed in the same trend prediction setting based on the same information. Underconfidence in the trend recognition task is more pronounced the longer the price history observed by subjects and the higher the discriminability of the price path generating processes. Furthermore, we find that the degree of overconfidence in both tasks is significantly positively correlated for all experimental subjects whereas performance measures are not.[6] This result suggests robust individual differences in the degree of overconfidence.

---

[6]Underconfidence is defined to be a negative degree of overconfidence.

In general, professional traders are more overconfident than students in trend prediction tasks. Our study has important implications for financial modelling. We argue that the question which psychological bias should be incorporated into a model does not depend on a specific informational setting (in our experiment both trend prediction tasks were embedded in the same informational setting) but solely on the specific task considered. It demonstrates that a theorist has to be careful when deriving assumptions about the behavior of agents in financial markets from psychological findings.

The rest of the paper is organized as follows. Section 2 surveys related finance models. Section 3 describes the design of the experiment. Sections 4 and 5 present our overconfidence measures, hypotheses, and results. Section 6 discusses the results and concludes.

2 Related Finance Models

In this section, we briefly review how judgment biases are incorporated into theoretical models of financial markets.

Overconfidence is most often modelled as underestimation of the variance of information signals and, as a consequence, the uncertain value of a risky asset. Assume there is a risky asset with liquidation value v which is a realization of $\tilde{v} \sim N(0, \sigma_{\tilde{v}}^2)$. Investors receive signals $\tilde{s} = \tilde{v} + c \cdot \tilde{e}$ (assuming that $\tilde{v}$ and $\tilde{e}$ are independent) with $\tilde{e} \sim N(0, \sigma_{\tilde{e}}^2) \Rightarrow \tilde{s} \sim N(0, \sigma_{\tilde{v}}^2 + c^2 \cdot \sigma_{\tilde{e}}^2)$. If $c = 1$, investors are rational, if $0 \leq c < 1$, investors are overconfident. Conditional expectation and conditional variance of $\tilde{v}$, given the realization s are

$$E[\tilde{v} \mid \tilde{s} = s] = E[\tilde{v}] + \frac{Cov[\tilde{v}, \tilde{s}]}{Var[\tilde{s}]}(s - E[\tilde{s}]) = \frac{\sigma_{\tilde{v}}^2}{\sigma_{\tilde{v}}^2 + c^2 \cdot \sigma_{\tilde{e}}^2} \cdot s \tag{1}$$

$$Var[\tilde{v} \mid \tilde{s} = s] = Var(\tilde{v}) - \frac{(Cov[\tilde{v}, \tilde{s}])^2}{Var[\tilde{s}]} = \sigma_{\tilde{v}}^2 - \frac{\sigma_{\tilde{v}}^4}{\sigma_{\tilde{v}}^2 + c^2 \cdot \sigma_{\tilde{e}}^2} \tag{2}$$

These equations show that overconfident investors underestimate the variance of the risky asset. In the extreme case $c = 0$ the conditional variance is zero. Benos (1998), Caballé and Sákovics (2003), Kyle and Wang (1997), Odean (1998), and Wang (1998) incorporate overconfidence into different types of models such as those of Diamond and Verrecchia (1981), Hellwig (1980), Grossman and Stiglitz (1980), Kyle (1985), and Kyle (1989). These models are useful to explain high levels of trading volume in financial markets. There are other overconfidence models that address questions like the dynamics of overconfidence, the survival of overconfident investors in markets, and the cross-section of expected returns. Examples are Daniel, Hirshleifer, and Subrahmanyam (1998, 2001), Hirshleifer and Luo (2001), Gervais and Odean (2001), and Wang (2001).

One of the most important facets of overconfidence is overestimating trends in random sequences or time series, in other words: betting on trends.[7] Shleifer and Summers (1990) state in their survey of the noise trader approach in finance that "one of the strongest investor tendencies documented in both experimental and survey evidence is the tendency to extrapolate or to chase the trend".[8] Betting on trends is closely related to noise trading, positive feedback trading and extrapolative expectations.[9] One explanation for positive feedback trading is the way investors form expectations. They naively extrapolate trends.[10] Furthermore, noise trading is related to overconfidence. Hirshleifer (2001) argues that pure

[7]See De Bondt (1993).

[8]Shleifer and Summers (1990), p. 28.

[9]See, for example, Hirshleifer (2001), p. 1545.

[10]See, for example, Hirshleifer (2001), p. 1567, or Daniel, Hirshleifer, and Teoh (2002), p. 145.

noise trading is a limiting case of overconfidence.[11]

Other models incorporate the second behavioral bias mentioned in the introduction, conservatism or underconfidence, into models of financial markets. One example is Barberis, Shleifer, and Vishny (1998) who model investors who try to infer information from earnings series.[12] Earnings follow a random walk. Because the firm pays out all earnings as dividends the true value of the firm follows a random walk, too. Investors do not understand that dividends and firm value follow a random walk and thus do not use the random walk model to forecast future earnings. They believe that the earnings process stochastically fluctuates between either a mean-reverting or a trending regime. Using past realizations of the earnings series investors try to find out which of the two regimes is currently generating earnings. This way of modeling beliefs or updating of probabilities is able to capture the conservatism bias (underconfidence, underreaction) and the representativeness heuristic. An investor using the mean-reverting model to forecast earnings reacts too little to an earnings announcement which leads to underreaction or conservatism. However, when the trending model is used to forecast earnings, strings of good or bad news are extrapolated too far into the future. This captures the idea of representativeness. Investors see patterns in random series and they thus extrapolate trends and overreact. One implication of this way of modeling beliefs is that investors erroneously use the frequency of recent earnings reversals to predict the likelihood that the current earnings change will be reversed in the future.[13]

This overview shows that there are many different ways of modelling judgment biases:

---

[11]Hirshleifer (2001), p. 1568.

[12]To be more precise, Barberis, Shleifer, and Vishny (1998) model a representative investor.

[13]See Bloomfield and Hales (2002) for an experimental investigation of this assumption.

Some models assume an underestimation of the variance of signals (or overestimation of their precision), i.e. too tight confidence intervals. Other theories model how investors form beliefs and predict future outcomes when they observe past realizations of a time series. Furthermore, there is the mechanistic modelling approach in positive feedback trading models in which investors form expectations of future prices by extrapolating trends or in which differences of past prices are directly incorporated into asset demand functions. Our study helps to evaluate these various modelling approaches and competing assumptions.

3 Experimental Design

The experiment we present in this paper was part of a larger project that consisted of four phases and covered a time span from February to April 2002.[14] In the first phase a questionnaire was presented that contained knowledge related overconfidence tasks that required a controlled environment. In this phase, we further assigned secret IDs to subjects to guarantee concealment and collected demographic data. The three other phases of the project were internet based. In prespecified time windows subjects had to access a web page and log in to the experimental software. Each phase took about 30 minutes to complete. Overall, 31 professionals of a large German bank participated in the project, 29 of them completed all parts of the study, the remaining two subjects dropped out or missed a phase due to vacation or other reasons. Based on a self report in the first phase, 11 professionals assigned their job to the area 'Derivatives', 10 to the area 'Proprietary Trading', 12 to the area 'Market Making' and 6 to 'Other Area'.[15] The age of the profes-

---

[14]In addition, there was a preexperimental meeting in which we interviewed the professional subjects to better understand their decision scope and goals.

[15]Subjects could assign themselves to more than one area.

sionals ranged from 23 to 55 with a median age of 33 years. 14 subjects had an university degree and the median subject worked five years for the bank (range 0.5 to 37 years). In addition to the professionals, we had a control group of 64 advanced students, all specializing in Banking and Finance at the University of Mannheim. Their median age was 24 and ranged from 22 to 30. While there were no women in the group of professionals, the control group consisted of 6 female and 58 male students. The control group was faced with exactly the same procedure as the professionals with the one exception that for organizational reasons the questionnaire of phase 1 was filled out at the end not the beginning of the project.

All data relevant for this paper was electronically collected in the phases 2 and 3.[16] It consists of two related tasks which we call 'trend prediction by probability estimates' or trend recognition and 'trend prediction by confidence intervals' or forecasting. In the trend prediction by probability estimates part, subjects were confronted with two simple distributions of price changes. Both processes could generate price movements of size -2, -1, 0, $+1$, and $+2$ with different probabilities. They were constructed, such that one process had a positive trend, i.e. a positive expected value, and the other process had a negative trend. Both processes were graphically and numerically displayed and their meaning explained in detail. Figure 1 shows a screenshot of the experiment. Subjects were further informed that one of the two processes was randomly picked to generate a price path. After observing this path it would be their task to state a subjective probability that this path is driven by a positive or a negative trend. Great care was taken to make sure subjects understood that this was a synthetically generated path and no relation to real market phenomena and price patterns must be assumed. Subjects then saw a developing

---

[16]There were further tasks in phases 2 and 3 that we do not address in this paper.

price path. The path started in t=0 at price 100 and moved on to t=5, where subjects had to make their first judgment. The probability elicitation was implemented as a two stage procedure. First, subjects had to state if they were at least 90% sure that a specific process would underly the observed path. Then, according to their answer they could choose an explicit percentage number in the restricted domain. This procedure was perfectly symmetric. When moving the slider the numerical percentage values simultaneously increased and decreased for both processes (see Figure 1). Thus, the fine tuning offered probabilities between 90% and 100% for the negative trend if subjects were at least 90% sure about the negative trend in the first step. The scale was restricted to a respective percentage range [0%, 10%], if they were 90% sure about a positive trend and to [10%, 90%] otherwise. It was possible in stage 2 to go back and change the judgment of the first stage. Subjects made such judgments at dates t=5, 6, 7, 8, 9, 10 and 20.[17] Three different pairs of price processes were used within the experiment. In all cases the distribution with positive trend was a symmetric copy of the distribution with negative trend. Thus by stating the negative trend probabilities for the outcomes −2, −1, 0, +1, and +2 the complete pair is defined. The following distributions were used:

$$D_{1.5} \quad = \quad (18\%, \quad 24\%, \quad 34\%, \quad 16\%, \quad 8\%)$$

$$D_{1.75} \quad = \quad (9.8\%, \quad 28\%, \quad 43\%, \quad 16\%, \quad 3.2\%)$$

$$D_2 \quad = \quad (12\%, \quad 32\%, \quad 37\%, \quad 16\%, \quad 3\%)$$

The index of D_k is chosen as the quotient of the probabilities for the outcome -1 and the outcome +1. These odds k will later turn out to be central for deriving the rational trend probabilities given some price change. In addition, k can be interpreted as a measure of

---

[17]On entering the second stage the slider was always set to the stated percentage of the previous round or the closest available percentage value (i.e. 90% or 10%). In t=5 it was set to 50% or the closest available number.

the discriminability of the two trends. In phase 2 of the experiment, all subjects saw each process pair exactly once. So, overall they had to estimate 21 probabilities (for each of the three price paths probability judgments for seven dates). Phase 3 was identical to phase 2 with different price paths presented to subjects.[18] Hence, overall we have 42 probability judgments for each subject in this 'trend prediction by probability estimates' task.

After stating the subjective probability at time t=20 subjects were further asked to predict how the price path will develop until t=40. They had to provide a price interval that they expect to contain the price at t=40 with 90% probability. The price at t=40 should fall short of the lower boundary and pass over the higher boundary with only 5% probability each. Such an interval judgments was collected for each process pair in phases 2 and 3. Thus, overall we have 6 intervals for each subject in this 'trend prediction by confidence intervals' part.[19]

4   General Hypotheses and Results

4.1   Trend Prediction by Probability Estimates

In this subsection we compare the subjective probability estimate with the mathematically correct probability that a given chart was generated by the process with the negative trend (which, of course, determines the probability that this chart was generated by the process with the positive trend). It turns out that $P(\text{neg. trend}|\text{level} = x)$ neither depends on date t nor on details of the price path but only on the absolute price change from the

---

[18]To to be more explicit, the same overall set of price paths was used in both phases but with a different assignment across subjects.

[19]Two professionals participated in only one of the two phases and thus provided only three intervals.

starting price level 100. This can be seen as follows: If we observe a price step of $+1$ then the odds are given as

$$\frac{P(+1|\text{pos. })}{P(+1|\text{neg. })} = k. \tag{3}$$

Due to the specific construction of the risk profiles it is easily verified that it holds more generally

$$\frac{P(z|\text{pos. })}{P(z|\text{neg. })} = k^z \quad \text{for} \quad z \in \{-2, -1, 0, +1, +2\}. \tag{4}$$

From $P(\text{level} = 100 \mid \text{pos. })/P(\text{level} = 100 \mid \text{neg. }) = 1$ it then follows by induction

$$\frac{P(\text{level} = x|\text{pos.})}{P(\text{level} = x|\text{neg.})} = k^{x-100} \quad \text{for all } x. \tag{5}$$

Bayes' Law yields

$$P(\text{neg. } | \text{ level} = x) = \frac{P(\text{level} = x|\text{neg.}) \cdot P(\text{neg. })}{P(\text{level} = x|\text{neg.}) \cdot P(\text{neg.}) + P(\text{level} = x|\text{pos.}) \cdot P(\text{pos. })} \tag{6}$$

and with $P(\text{neg.}) = P(\text{pos.})$ it follows from equation (5),

$$P(\text{neg.}|\text{level} = x) = \frac{1}{1 + k^{x-100}}. \tag{7}$$

Assume, for example, $k = 2$, i.e. distribution $D_2 = (12\%, 32\%, 37\%, 16\%, 3\%)$ and a price level of 102 after 5 periods ($t = 5$). Then, the probability $P(\text{neg. trend}|\text{level} = 102)$

is $1/(1 + 2^{102-100}) = 1/5$. Our design is similar to the study of Griffin and Tversky (1992). We discuss this study more deeply as this similarity to their study yields some of our hypotheses. In the study of Griffin and Tversky (1992), there was a coin with either a $3:2$ bias in favor of heads or tails. Subjects received outcomes of coin spinning, i.e. the number of heads and the number of tails.[20] Then, subjects were asked to state their subjective confidence that the respective outcome was generated by the coin with the $3:2$ bias in favor of heads. As in our experiment, the probability distributions were symmetric. It turned out that the only relevant determinant of the mathematically correct probability was the difference between the numbers of heads and the number of tails whereas sample size and sequence of heads and tails were irrelevant. This shows the analogy to our study if one interprets the biases in favor of heads or tails as positive or negative trend. Counterparts of sample size and sequence are time period and price path in our study.

Griffin and Tversky (1992) report in this scenario an underestimation of the mathematically correct probability. This leads to our first hypothesis:

> *Hypothesis 1:*
>
> *Subjects are, on average, underconfident, i.e. they underestimate the mathematically correct probability.*

Furthermore, Griffin and Tversky (1992) find that underconfidence is stronger when the weight of evidence (sample size) and the discriminability of the coin biases are high. This finding motivates our second and third hypotheses:

---

[20]There was always a majority of the number of heads.

Hypothesis 2:

Underconfidence is stronger the higher the sample size, i.e. the longer the time period.

Hypothesis 3:

Underconfidence is stronger the higher the discriminability of the two symmetric processes, i.e. the higher k.

Equation (7) shows that as long as the price level is higher than 100 the probability that the chart was generated by the process with the negative trend should be lower than 0.5. If the price level is lower than 100 the probability that the chart was generated by the process with the negative trend should be higher than 0.5. If subjects predict the rational direction, we are able to compare the subjective probability with the mathematically correct probability. Assume, for example, the rational probability for the negative trend is 0.7. If a person states a subjective probability of 0.9 this person is classified as overconfident. If the person states a probability of 0.6 the subject is regarded as underconfident. This classification is far less clear if a subject states a probability of, say, 0.2 for the negative trend. Is this person extremely overconfident? Or is she underconfident because her subjective certainty is lower than the objective certainty? We argue that these subjects cannot be reasonably classified as either overconfident or underconfident.[21] We thus omit

---

[21]There are other studies that measure overconfidence in a different way. According to these studies we should classify subjects that predict the wrong trend to be true as extremely overconfident. However, it can be shown that this methodology could lead to serious biases. Furthermore, other studies restrict the probability range for the subjects' answers. Griffin and Tversky (1992), for example, only analyze situations with a majority of heads. Other studies present knowledge questions. Subjects have to provide an answer and have to state their subjective confidence from 50 % to 100 %. In both cases, the probability range for the subjects' answers is restricted so that apparently illogical answers (e.g. the answer yes with a subjective certainty of 40 %) are impossible.

these subjects from our analysis. In period $t = 5$, 24 of 552 probability (4.3 %) estimates have to be eliminated (37 in period $t = 10$ and 28 in period $t = 20$).

In the following, we focus on the three time periods $t = 5$, $t = 10$, and $t = 20$.[22] Table 1 presents the number of probability estimates above (overconfident), equal to (well calibrated), and below the rational probability (underconfident).[23] The majority of probability estimates is below the rational probability thus indicating underconfidence. For example, in time period $t = 5$, 175 probability estimates are above the correct probability (33.14 %) whereas 293 probability estimates are below the correct probability (55.50 %). This preliminary analysis suggests that hypothesis 1 is confirmed. Furthermore, Table 1 shows that underconfidence increases with time. This observation is consistent with our second hypothesis.

To analyze these preliminary observations more deeply we define an overconfidence measure doc_1 which is simply the difference between the subjective and the mathematically correct probability.[24] This measure indicates overconfidence if $doc_1 > 0$ and underconfidence if $doc_1 < 0$. Table 1 presents the medians of doc_1 for dates $t = 5$, $t = 10$, and $t = 20$. The medians of doc_1 are always negative indicating underconfidence. For example, doc_1 in $t = 5$ is -4.12, i.e. the median probability estimate underestimates the mathematically correct probability by -4.12 percentage points. Again, underconfidence increases with time. In $t = 10$ and $t = 20$, doc_1 is even lower. A Wilcoxon signed rank test shows that all

[22]Analysis of the remaining time periods yields similar results.

[23]Almost 90% of the correct probability estimates ("well calibrated") correspond to price levels of 100 with a rational probability of 0.5.

[24]Note, that we only analyze subjects that predict the rational direction. Thus, all probabilities are equal to or above 0.5. The degree of overconfidence doc_1 is the rational probability for the positive (negative) trend minus the subjective probability for the positive (negative) trend.

three doc_1-values are significantly different from 0. When we analyze the medians of doc_1 separately for each risk profile we find results consistent with hypothesis 3. The higher the discriminability between the two risk profiles (the larger k), the larger underconfidence (lower doc_1-values).

Our doc_1 measure reflects the deviation of the subject's answer from the rational benchmark on the probability scale. Alternatively, the bias could be measured on the price scale as each probability uniquely corresponds to some price level. The choice of scale is not irrelevant, since the transformation from probabilities to prices is not linear. For example, a doc_1 value of -5 corresponds to a much larger price difference if it is derived from a 95 % probability estimate and a 90 % rational probability than if it results from the values 65 % and 60%, respectively. To make sure that our results are robust with respect to the chosen scale, we consider a second measure doc_2 on the price scale that is defined as the difference between the price level implied by the probability estimate and the observed price level.[25] If doc_2 is positive, subjects are overconfident, if doc_2 is negative, people are underconfident.

Table 1 presents the medians of doc_2 for dates $t = 5$, $t = 10$, and $t = 20$. Again, all doc_2 are negative indicating underconfidence.

So far, we have analyzed medians for all probability estimates although these observations are not independent because several observations come from the same subject. We now calculate doc_1 and doc_2 separately for each individual as simply the median of doc_1 and doc_2, respectively.

---

[25]We replace stated probabilities of 100 and 0 with 99.75 and 0.25, respectively. This takes into account that a subject with a subjective probability estimates in the range [99.5, 100] has to choose the answer 100 due to the discrete scale of possible probability estimates.

Table 1 shows medians of the medians doc_1 and doc_2 values per subject. All values indicate underconfidence. All doc_1 and doc_2 values are significantly different from zero at the 1 % level except for doc_2 in $t = 5$ which is only significant at the 5 % level.[26]

4.2   Trend Prediction by Confidence Intervals

Intertwined with the trend probability estimation, a second task was used to further examine possible judgment biases related to trend prediction. Following each final probability judgment, based on a price path from $t = 0$ to $t = 20$, subjects were asked to make a prediction about the price level at time $t = 40$. This judgment was elicited via a confidence interval, consisting of an upper and lower limit. Subjects were instructed that the price at $t = 40$ should be above the upper limit and below the lower limit with 5 % probability each. Thus the stated confidence interval was supposed to contain the price at $t = 40$ with 90 % probability. The experimental subjects had to provide such confidence intervals for three price paths in phase 2 as well as in phase 3. Thus overall we have six [low, high]-intervals, two for each of the price process pairs $D_{1.5}$, $D_{1.75}$, and D_2. The correct distribution of prices at $t = 40$ can be computed from the price at $t = 20$ and the process distributions. Thus the stated confidence intervals can be translated into probability intervals. For a perfectly calibrated subject these induced probability intervals $[p_{low}, p_{high}]$ should be [5%, 95%].[27] We measure the degree of under-/overconfidence via the length of the induced probability interval. If a subject is too sure about the price at t=40 and provides a too narrow interval, it is classified as overconfident. More explicit, our third measure of overconfidence doc_3 is defined as: $doc_3 = 90\% - (p_{high} - p_{low})$. Posi-

---

[26]When we analyze the medians of the means per subject we find similar results.

[27]Note, however, that due to the discreteness of the distribution it was not possible in general to exactly hit this target.

tive doc_3-values correspond to overconfidence, negative doc_3-values to underconfidence. In previous literature it was observed that individuals usually provide too narrow intervals when making similar judgments.[28] This leads to our fourth hypothesis:

Hypothesis 4:

In trend prediction by confidence intervals, subjects are overconfident on average,

i.e. they provide too narrow intervals ($doc_3 > 0$).

In the first column of Table 2, the main results regarding hypothesis 4 are summarized. From the 552 observations, 335 are classified as overconfident ($doc_3 > 0$) and 217 as underconfident ($doc_3 < 0$) with a median doc_3-value of $+10.03$.[29] A Wilcoxon signed rank test shows that the doc_3-values are significantly different from 0. In Table 2, we also present doc_3-values on a more aggregated level. From the 93 subjects in the data set 61 have a positive median doc_3-value and are thus classified as overconfident. The 32 other subjects are underconfident. The median subject has a doc_3-value of 8.01. Again, these doc_3-values are significantly different from 0. To address a potential concern, we will next look at a slightly different overconfidence measure doc_4. In the computation of the doc_3-values we had to determine the rational price distribution at $t = 40$. This t_{40}-distribution depends on the trend process pair as well as on the current ($t = 20$) price level. The current price level enters the computation in two ways. It obviously determines the absolute location of the t_{40}-distribution, but it also influences the shape of the distribution via the rational trend probabilities that follow from the price level. By an individual misjudgment of the trend probabilities in $t = 20$, as documented in

---

[28]See, for example, Lichtenstein, Fischhoff, and Phillips (1982).

[29]We do not report mean values because of the obvious asymmetry in possible doc_3-values (ranging from -10% to +90%).

the previous section, a different "rational" t_{40}-distribution would thus be derived. In the computation of our fourth overconfidence measure doc_4 we replace the rational trend probability derived from the price level at $t = 20$ by the trend probability stated by the individual in $t = 20$. As shown in Table 2, this alternation of the rational benchmark does not have a strong impact on the results, however. From the 524 observations, 341 are classified as overconfident and 183 as underconfident with a median doc_4-value of 9.11.[30] These doc_4-values as well as the aggregated numbers (67 subjects are classified as overconfident, 26 as underconfident) are significantly different from 0.

5 Traders versus Students: The Correlation of Overconfidence Measures

5.1 Traders versus Students

Table 3 analyzes the various overconfidence measures defined in the previous sections separately for the two subject groups: professional traders and lay people (students). The first part of Table 3 shows the medians of all observations. Accordingly, the second column shows figures that are also part of Table 1 and Table 2. For example, Table 3 shows again that the median doc_1-value among all observations in $t = 5$ is -4.12. The next two columns present the median doc_1-value in $t = 5$ as well as all the other doc_i-values, $i \in \{1, 2, 3, 4\}$, for the two groups of subjects separately. The main observation is that financial professionals have higher doc_i-values than students. Thus, the degree of overconfidence is always higher for traders. Note, again, that underconfidence is considered

---

[30]We restrict the doc_4-analysis to the same 524 observations that we considered in the analysis of the measures doc_1 and doc_2 in t=20.

to be a negative degree of overconfidence. Focusing on doc_1- and doc_2-values shows that all medians are negative (except for doc_1 and doc_2 in $t = 5$ with a median value of 0) implying that professionals perform better in this task: They are closer to the rational probability. In other words: They are less underconfident than students in this task. Analysis of doc_3- and doc_4-values yields that professionals are not generally better than students but, again, more overconfident. The last column of Table 3 presents the p-value of a Kruskal-Wallis test. Null hypothesis is equality of populations. The results show that we can reject the null hypothesis of the equality of populations in most cases at least at the 5 % level.

The second part of Table 3 shows the median doc_i-value of the medians per subject. In contrast to Table 1 we calculate the median of doc_1 and doc_2 over the three dates $t = 5$, $t = 10$, and $t = 20$ for ease of exposition. Again, all doc_i-values are higher for traders although a Kruskal-Wallis test suggests that we cannot reject the null hypothesis of equality of populations for doc_3 and doc_4.

The main message of this subsection can be summarized as follows. Traders always have a higher degree of overconfidence. As in the 'trend prediction by probability estimates' task underconfidence is observed, professionals perform better in this task. But traders are not better in general. The 'trend prediction by confidence intervals' task shows that professionals are again more overconfident.

5.2 Correlation of Overconfidence and Performance Measures

Table 4 presents correlations as well as significance levels of correlations of doc_i-measures, $i \in \{1, 2, 3, 4\}$. doc_1 and doc_2 are the medians over the three dates $t = 5$, $t = 10$, and $t = 20$ per subject. doc_3 and doc_4 are the medians per subject. Table 4 shows that

all overconfidence measures are positively correlated. These correlations are significant at the 1 % level except for $corr(doc_1, doc_4)$ with a p-value of 0.0166. Accordingly, we find interpersonal differences. A higher degree of overconfidence in the 'trend prediction by probability estimates' task implies a higher degree of overconfidence in the 'trend prediction by confidence intervals' task and vice versa. This result is remarkable as these are completely different tasks.

Table 5 shows an according analysis for the subjects' performance instead of their overconfidence. The variables $perf_i$ are calculated as absolute values of doc_i, $i \in \{1, 2, 3, 4\}$, and are aggregated to median values as described above. The lower the $perf_i$ values, the closer the subjects approached the rational benchmark, i.e. the better they performed in the task. In contrast to the overconfidence measures, we do not find a significantly positive correlation for individual performance in the different task types, i.e. between $perf_1$, $perf_2$ and $perf_3$, $perf_4$. As shown in Table 5, the lowest p-value is 0.1826 for the correlation between $perf_2$ and $perf_3$. Combining the results of Table 4 and Table 5, we can conclude that overconfidence rather than performance is a robust individual characteristic: Individuals are not generally better or worse in such judgment tasks, but generally more or less overconfident.

6   Discussion and Conclusion

Our main findings can be summarized as follows.

1. Depending on the type of task *either* underconfidence *or* overconfidence can be observed in the same trend prediction setting.

2. The degrees of overconfidence derived from different trend prediction tasks are significantly positively correlated whereas performance measures are not.

3. In trend prediction tasks professional traders are generally more overconfident than students.

The first finding demonstrates that a theorist has to be careful when deriving assumptions that are based on psychological evidence. This is in line with Hirshleifer's argument that "it is often not obvious how to translate preexisting evidence from psychological experiments into assumptions about investors in real financial settings. Routine experimental testing of the assumptions and conclusions of asset-pricing theories is needed to guide modeling."[31] Our finding is particularly interesting as we consider two types of trend prediction that have obvious practical relevance and are both near at hand facets of the general phenomenon that people are biased when detecting trends. In addition, they both are realistic forms of trend prediction tasks that are separately modelled in theoretical analyses of financial markets.

But if a model is built in which investors observe asset prices and draw conclusions about future price developments, how should possible biases in trend prediction be incorporated? Should a theorist rely on the result of our first experimental task, which indicates that subjects underestimate the mathematically correct probability indicating underconfidence? Or should a theorist model overconfidence in this setting by assuming confidence intervals that are too narrow or, in other words, underestimation of the variance? It is important to stress that the answer to this question does not depend on a specific informational setting (in our experiment both trend prediction tasks were embedded in the

---

[31]Hirshleifer (2001), p. 1577.

same informational setting) but solely on the specific task considered.

The second result is of interest as the documented positive correlation of the degrees of overconfidence suggest robust individual differences in reasoning. In this context, it should be emphasized that trend prediction by probability estimates and trend prediction by confidence intervals are completely different tasks and not only a different way of measuring the same aspects.[32] Moreover, this result is of interest as it shows that individuals are not generally better or worse in different judgment tasks, but just more or less overconfident. Overconfidence rather than performance seems to be the robust individual characteristic.

The third result is important because it shows that professional experience does not eliminate biases in job-related judgment tasks. This contradicts the argument that overconfidence might not play a significant role in financial markets since professionals that dominate the markets are less susceptible to such biases than lay people.

---

[32]This is best illustrated by pointing out that in the trend probability estimation task the mathematically correct answer does neither depend on the path length nor directly on the variance of the price processes, but only on the odds k, measuring the relative strengths of both trends. For the rational confidence intervals, however, the variance of the processes and the prediction time horizons are most relevant.

References

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor senti-ment, *Journal of Financial Economics* 49, 307–343.

Benos, Alexandros V., 1998, Aggressiveness and survival of overconfident traders, *Journal of Financial Markets* 1, 353–383.

Bernard, Victor L., and Jacob K. Thomas, 1989, Post-earnings-announcement drift: De-layed price response or risk premium, *Journal of Accounting Research* 27, 1–36 Supple-ment 1989.

Bloomfield, Robert, and Jeffrey Hales, 2002, Predicting the next step of a random walk: experimental evidence of regime-shifting beliefs, *Journal of Financial Economics* 65, 397–414.

Caballé, Jordi, and József Sákovics, 2003, Speculating against an overconfident market, *Journal of Financial Markets* 6, 199–225.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

——— , 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Fi-nance* 56, 921–965.

Daniel, Kent, David Hirshleifer, and Siew Hong Teoh, 2002, Investor psychology in capital markets: evidence and policy implications, *Journal of Monetary Economics* 49, 139–209.

DeBondt, Werner F.M., 1993, Betting on trends: Intuitive forecasts of financial risk and return, *International Journal of Forecasting* pp. 355–371.

Diamond, Douglas W., and Robert E. Verrecchia, 1981, Information aggregation in a noisy rational expectations economy, *Journal of Financial Economics* 9, 221–235.

Edwards, W., 1968, Conservatism in human information processing, in B. Kleinmuntz, ed.: *Formal Representation of Human Judgment* . pp. 17–52 (Wiley, New York).

Gervais, Simon, and Terrance Odean, 2001, Learning to be overconfident, *Review of Financial Studies* 14, 1–27.

Glaser, Markus, and Martin Weber, 2003, Momentum and turnover: Evidence from the german stock market, *Schmalenbach Business Review* 55, 108–135.

Gompers, Paul A., and Andrew Metrick, 2001, Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229–259.

Griffin, Dale, and Amos Tversky, 1992, The weighing of evidence and the determinants of confidence, *Cognitive Psychology* 24, 411–435.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hellwig, Martin F., 1980, On the aggregation of information in competitive markets, *Journal of Economic Theory* 22, 477–498.

Hirshleifer, David, 2001, Investor psychology and asset pricing, *Journal of Finance* 56, 1533–1597.

———— , and Guo Ying Luo, 2001, On the survival of overconfident traders in a competitive securities market, *Journal of Financial Markets* 4, 73–84.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 1, 65–91.

———, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 2, 699–720.

Jones, Charles P., and Robert H. Litzenberger, 1970, Quarterly earnings reports and intermediate stock price trends, *Journal of Finance* 25, 143–148.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1336.

———, 1989, Informed speculation with imperfect competition, *Review of Economic Studies* 56, 317–356.

———, and F. Albert Wang, 1997, Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?, *Journal of Finance* 52, 2073–2090.

Lakonishok, Joseph, Andrei Shleifer, and Robert W. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–43.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D. Phillips, 1982, Calibration of probabilities: The state of the art to 1980, in Daniel Kahneman, Paul Slovic, and Amos Tversky, ed.: *Judgment under uncertainty: Heuristics and Biases* . pp. 306–334 (Cambridge University Press).

Nofsinger, John R., and Richard W. Sias, 1999, Herding and feedback trading by institutional and individual investors, *Journal of Finance* 54, 2263–2295.

Odean, Terrance, 1998, Volume, volatility, price, and profit when all traders are above average, *Journal of Finance* 53, 1887–1934.

Rouwenhorst, K. Geert, 1998, International momentum strategies, *Journal of Finance* 53, 267–284.

——— , 1999, Local return factors and turnover in emerging markets, *Journal of Finance* 54, 1439–1464.

Shiller, Robert J., 2001, Bubbles, human judgment, and expert opinion, Cowles Foundation Discussion Paper, Yale University.

Shleifer, Andrei, and Lawrence H. Summers, 1990, The noise trader approach in finance, *Journal of Economic Perspectives* 4, 19–33.

Wang, F. Albert, 1998, Strategic trading, asymmetric information and heterogeneous prior beliefs, *Journal of Financial Markets* 1, 321–352.

——— , 2001, Overconfidence, investor sentiment, and evolution, *Journal of Financial Intermediation* 10, 138–170.

Table 1: **Trend Prediction by Probability Estimates**

This table presents the number of probability estimates above (overconfident), equal to (well calibrated), and below the rational probability for the three time periods $t = 5$, $t = 10$, and $t = 20$. doc_1 is the difference between the subjective and the mathematically correct probability. doc_2 is the difference between the implied price change and the observed price change. Median values of these measures are also presented.

	t=5	t=10	t=20
Number of observations (all observations)	528	515	524
Overconfident ($doc_1 > 0$)	175	140	106
	(33.14 %)	(27.18 %)	(20.23 %)
Well calibrated ($doc_1 = 0$)	60	25	17
	(11.36 %)	(4.85 %)	(3.24 %)
Underconfident ($doc_1 < 0$)	293	350	401
	(55.50 %)	(67.96 %)	(76.53 %)
doc_1 (medians; all risk profiles)	-4.12	-5.26	-8.75
Wilcoxon signed-rank test (H_0 : $doc_1 = 0$)	$p = 0.0000$	$p = 0.0000$	$p = 0.0000$
doc_1 (medians; $k = 1.5$)	-3.00	-2.14	-5.23
doc_1 (medians; $k = 1.75$)	-3.64	-5.51	-9.35
doc_1 (medians; $k = 2$)	-5.67	-7.64	-9.81
doc_2 (medians; all risk profiles)	-0.42	-1.00	-2.43
Wilcoxon signed-rank test (H_0 : $doc_2 = 0$)	$p = 0.0000$	$p = 0.0000$	$p = 0.0000$
doc_2 (medians; $k = 1.5$)	-0.30	-0.47	-1.43
doc_2 (medians; $k = 1.75$)	-0.28	-1.07	-3.00
doc_2 (medians; $k = 2$)	-0.60	-1.00	-3.00
Number of observations (medians per subject)	93	93	93
Overconfident ($doc_1 > 0$)	28	19	13
	(30.11 %)	(20.43 %)	(13.98 %)
Well calibrated ($doc_1 = 0$)	7	3	0
	(7.53 %)	(3.23 %)	(0.00 %)
Underconfident ($doc_1 < 0$)	58	71	80
	(62.37 %)	(76.34 %)	(86.02 %)
doc_1 (medians of medians per subject; all risk profiles)	-5.45	-5.03	-7.46
Wilcoxon signed-rank test (H_0 : $doc_1 = 0$)	$p = 0.0011$	$p = 0.0000$	$p = 0.0000$
doc_1 (medians of medians per subject; $k = 1.5$)	-2.57	-2.54	-5.11
doc_1 (medians of medians per subject; $k = 1.75$)	-3.64	-6.60	-9.93
doc_1 (medians of medians per subject; $k = 2$)	-8.06	-9.17	-9.90
doc_2 (medians of medians per subject; all risk profiles)	-0.55	-0.81	-2.39
Wilcoxon signed-rank test (H_0 : $doc_2 = 0$)	$p = 0.0183$	$p = 0.0000$	$p = 0.0000$
doc_2 (medians of medians per subject; $k = 1.5$)	-0.40	-0.49	-1.40
doc_2 (medians of medians per subject; $k = 1.75$)	-0.46	-1.07	-2.95
doc_2 (medians of medians per subject; $k = 2$)	-0.60	-1.25	-3.48

Table 2: **Trend Prediction by Confidence Intervals**

This table presents the number of confidence intervals classified as overconfident or underconfident by the measures doc_3 and doc_4. doc_3 is defined as: $doc_3 = 90\% - (p_{high} - p_{low})$ with positive doc_3-values corresponding to overconfidence, negative doc_3-values to underconfidence. In the computation of doc_4 we replace the rational trend probability derived from the price level at $t = 20$ by the trend probability stated by the individual in $t = 20$. Median values of these measures are also presented.

All observations	doc_3	doc_4
Number of observations	552	524
$doc_i > 0, \quad i \in \{3,4\}$ (overconfident)	335	341
	(60.69 %)	(65.08 %)
$doc_i < 0, \quad i \in \{3,4\}$ (underconfident)	217	183
	(39.31 %)	(34.92 %)
$doc_i, \quad i \in \{3,4\}$ (median of all observations)	10.03	9.11
Wilcoxon signed-rank test ($H_0 : \quad doc_i = 0, \quad i \in \{3,4\}$)	$p = 0.0000$	$p = 0.0000$

Medians per subjects	doc_3	doc_4
Number of observations	93	93
$doc_i > 0, \quad i \in \{3,4\}$ (overconfident)	61	67
	(65.59 %)	(72.04 %)
$doc_i < 0, \quad i \in \{3,4\}$ (underconfident)	32	26
	(34.41 %)	(27.96 %)
$doc_i, \quad i \in \{3,4\}$ (median of median per subject)	8.01	13.22
Wilcoxon signed-rank test ($H_0 : \quad doc_i = 0, \quad i \in \{3,4\}$)	$p = 0.0000$	$p = 0.0000$

Table 3: **Trader versus Lay People**

This table presents the various overconfidence measures defined in Subsection 4.1 and Subsection 4.2 separately for the two subject groups: professional traders and lay people (students). The first part of Table 3 shows the medians of all observations. Accordingly, the second column shows figures that are also part of Table 1 and Table 2. The next two columns present the median doc_1-value in $t = 5$ as well as all the other doc_i-values, $i \in \{1, 2, 3, 4\}$, for the two groups of subjects separately. The last column of Table 3 presents the p-value of a Kruskal-Wallis test. Null hypothesis is equality of populations. The second part of Table 3 shows the median doc_i-value of the medians per subject. In contrast to Table 1 we calculate the median of doc_1 and doc_2 over the three dates $t = 5$, $t = 10$, and $t = 20$ for ease of exposition.

	all Subjects	Professionals	Students	p-value of Kruskal-Wallis test H_0 : Equality of populations
doc_1 (medians; all observations; $t = 5$)	-4.12	0.00	-5.00	0.0094
doc_1 (medians; all observations; $t = 10$)	-5.26	-1.93	-6.97	0.0001
doc_1 (medians; all observations; $t = 20$)	-8.75	-8.00	-9.22	0.0314
doc_2 (medians; all observations; $t = 5$)	-0.42	0.00	-0.58	0.0023
doc_2 (medians; all observations; $t = 10$)	-1.00	-.042	-1.05	0.0001
doc_2 (medians; all observations; $t = 20$)	-2.43	-2.04	-2.49	0.0772
doc_3 (medians; all observations)	10.03	16.68	5.88	0.0313
doc_4 (medians; all observations)	9.11	13.35	8.07	0.1616
doc_1 (medians of medians)	-5.00	-1.46	-6.86	0.0089
doc_2 (medians of medians)	-0.92	-0.36	-1.09	0.0062
doc_3 (medians of medians)	8.01	17.42	4.79	0.2277
doc_4 (medians of medians)	13.22	15.31	8.51	0.4388

Table 4: **Correlation of Overconfidence Measures**

This table presents correlations as well as significance levels of correlations of doc_i-measures, $i \in \{1, 2, 3, 4\}$. doc_1 and doc_2 are the medians over the three dates $t = 5$, $t = 10$, and $t = 20$ per subject. doc_3 and doc_4 are the medians per subject.

	doc_1	doc_2	doc_3	doc_4
doc_1	1			
doc_2	0.8756	1		
	(0.0000)			
doc_3	0.2924	0.3513	1	
	(0.0045)	(0.0006)		
doc_4	0.2479	0.3032	0.983	1
	(0.0166)	(0.0031)	(0.0000)	

Table 5: **Correlation of Performance Measures**

This table presents correlations as well as significance levels of correlations of $perf_i$-measures, $i \in \{1, 2, 3, 4\}$. $perf_1$ and $perf_2$ are the medians over the three dates $t = 5$, $t = 10$, and $t = 20$ per subject. $perf_3$ and $perf_4$ are the medians per subject.

	$perf_1$	$perf_2$	$perf_3$	$perf_4$
$perf_1$	1			
$perf_2$	0.6345	1		
	(0.0000)			
$perf_3$	0.0759	0.1394	1	
	(0.4694)	(0.1826)		
$perf_4$	0.0943	0.1104	0.9781	1
	(0.3686)	(0.2923)	(0.0000)	

Figure 1: **Screen Shot**

This figure shows a screen-shot of the 'trend prediction by probability estimates' task. In this part, subjects were confronted with two simple distributions of price changes. Both processes could generate price movements of size -2, -1, 0, $+1$, and $+2$ with different probabilities. They were constructed, such that the one process had a positive trend, i.e. a positive expected value, and the other process had a negative trend. Both processes were graphically and numerically displayed and their meaning explained in detail. Subjects were further informed that one of the two processes was randomly picked to generate a price path. After observing this path it would be their task to state a subjective probability that this path is driven by a positive or a negative trend. The path started in t=0 at price 100 and moved on to t=5, where subjects had to make their first judgment. The probability elicitation was implemented as a two stage procedure. First, subjects had to state if they were at least 90% sure that a specific process would underly the observed path. Then, according to their answer they could choose an explicit percentage number in the restricted domain. This procedure was perfectly symmetric. When moving the slider the numerical percentage values simultaneously increased and decreased for both processes. Thus, the fine tuning offered probabilities between 90% and 100%, if subjects were at least 90% sure about the negative trend in the first step. The scale was restricted to a percentage range [0%, 10%], if they were 90% sure about a positive trend and to [10%, 90%] otherwise. It was possible in stage 2 to go back and change the judgment of the first stage. Subjects made such judgments at times t=5, 6, 7, 8, 9, 10 and 20.