Hindsight bias, risk perception and investment performance

Bruno Biais (Toulouse University) & Martin Weber (Mannheim University)[1]

June 2008

Abstract

Once they have observed information, hindsight biased agents fail to remember how ignorant they were initially, "they knew it all along." We formulate a theoretical model of this bias, providing a foundation for empirical measures, and implying that hindsight biased agents learning about volatility will underestimate it. In an experiment involving 67 students from Mannheim University, we find that hindsight bias reduces volatility estimates. In another experiment, involving 85 investment bankers in London and Frankfurt, we find that more biased agents have lower performance. These findings are robust to differences in location, information, overconfidence and experience.

[1] Many thanks for thoughtful comments to participants in the Toulouse and Amsterdam Universities seminars, the Warwick Conference on Behavioral Finance (2006), the Munich Summer School on Economics and Psychology (2006), and the UCL conference on biased beliefs, in particular Denis Hilton, Luis Garicano, Robin Hogarth, Florencio Lopez de Silanes, Thomas Mariotti, Paola Sapienza, Jean Tirole and Luigi Zingales. We like to thank Alen Nosic for excellent research assistance.

1 Introduction

Psychologists have extensively documented the prevalence of cognitive biases and deviations from rational choice. Are these results relevant for actual financial markets? Professionals, operating in their natural habitat and with a lot at stake, could be expected to behave rationally. Biases could be weeded out by learning and experience. Market forces and economic environments could induce optimizing. The present paper combines experimental evidence from the lab and from the field to shed light on these issues.

Decision making in financial markets relies crucially on information processing and learning. Efficient learning requires comparing new information to previous expectations. For example, after earnings announcements, investors must compare the news to their prior expectations. They must take into account the information content of the difference between the former and the latter. The hindsight bias, which is the inability to correctly remember one's prior expectations after observing new information, hinders such information processing. Biased agents are not surprised by new information, as "they knew it all along." Starting with Fischhoff (1975) and Fischhoff and Beyth (1975), this bias has been amply documented by the psychology literature.[2] This literature has shown that the hindsight bias arises in a large variety of contexts and that even the knowledge of this bias does not eliminate it.

The goal of this paper is to study the consequences of the hindsight bias for an important type of economic activity: investment and trading. We propose a simple theoretical model providing

[2] For an extended review of the literature and a discussion of psychological explanations of the bias see Christensen-Szalanski and Wilham (1991), Hawkins and Hastie (1990) and Roese (2004).

a foundation for the measurement of the bias and showing that hindsight biased agents fail to assess variances correctly. When volatility is stochastic, traders need to update their assessment of risk, based on return realizations. On observing unexpectedly positive or negative returns, rational agents should raise their volatility estimates. Hindsight biased agents, who "knew it all along," fail to understand that such returns were unexpected, and thus underestimate variances. More generally, hindsight biased agents will form inaccurate beliefs about asset returns, leading to suboptimal trades and inferior financial performance. For example, hindsight biased agents underestimating volatility will underprice options and fail to hedge appropriately. Hindsight biased agents will also fail to call in questions economic analyses at odds with the facts. And they will fail to estimate accurately the difference between their information and that of other traders. This will undermine their ability to earn trading profits based on superior information.

To test the hypothesis that the hindsight bias hinders learning about risk, we designed a lab experiment involving 67 students from Mannheim University. We gave the participants financial data and asked them to estimate variances. Then, we gave them new data and asked them to again estimate variances. The idea was to study how participants process this new data to update their volatility estimates. The experiment involved two treatments. In the first treatment, the experimenter reminded the participants of their initial estimates, thus muting the bias. In the second one, the experimenter asked the participants to remember their initial estimates, so that the bias could manifest itself. Consistently with our theoretical prediction, in second first treatment participants came up with lower variance estimates than in the first one.

To test the hypothesis that the hindsight bias reduces performance, we collected data from 85 investment bankers working in Frankfurt or London. We found these bankers to be significantly hindsight biased. And yet the questions we asked them related to their own field of expertise, such as stock market returns, macro variables or characteristics of the investment banking industry. We also found that more experienced bankers or bankers with more precise information were not less biased. Using data provided by the bank, we proxied bankers' performance by their compensation (including bonuses). Consistent with the hypothesis that the hindsight bias reduces performance, we found that the bankers in the highest earnings category had the lowest bias on average. We first obtained this finding in the Frankfurt subsample. Then, to test its robustness, we collected additional data for London. In this new sample we found the same result. We also checked that our results were not driven by other variables, such as the experience, overconfidence or information of the bankers.

Our paper is in line with Camerer et al (1989). We use similar methods to measure the hindsight bias. But our focus and findings differ from theirs. Camerer et al (1989) showed that markets reduce but don't eliminate the hindsight bias. We complement this finding by offering new results on the consequences of the bias for information processing and financial performance. Another distinguishing feature of our paper is that our findings on the link between the hindsight bias and performance are based on data collected in the field, from highly paid and experienced investment bankers. This is in the line of several recent field experiments on financial decision and psychology.[3] Our analysis is also in the line of several recent papers empirically relating psychological constructs to economic variables (see Camerer (1987), Fenton O'Creevy (2003), Biais et al (2005), and Glaser and Weber (2007)).

In the next section we present our theoretical model. In Section 3, we present our lab experiment. In Section 4, we present our field experiment. Section 5 concludes.

[3]For example, Haigh and List (2005) experimentally find that myopic loss aversion is even stronger for CBOT traders than students, while Sarin and Weber (1993) find that ambiguity aversion reflected in experimental market prices is also prevalent among investment bankers.

2 Theory

2.1 Defining the hindsight bias

The research on hindsight bias started with Fischhoff (1975) and Fischhoff and Beyth (1975). After events have been resolved, rational agents correctly remember their ex–ante expectations. Hindsight biased agents don't. By a process of memory reconstruction they incorporate the new knowledge into their remembrance of the prior expectation (see e.g., Carli, 1999). In the words of Wasserman et al (1991, p 30) this is "a projection of new knowledge into the past accompanied by a denial that the outcome information has influenced judgements." Correspondingly, for hindsight biased agents the ex–post recollection of the initial belief will be closer to the realization than the true ex–ante expectation. Numerous papers in experimental psychology have shown the prevalence of this bias and discussed its measurement (see e.g., Hawkins and Hastie (1990)). Musch (2003) finds empirically that individual differences in hindsight bias exist. This suggests that this bias acts as a trait, consistently influencing individual behavior in various environments. The notion of hindsight bias was initially developed in the context of binary variables. In that case, the bias arises if the ex-post recollection of the ex-ante probability is greater when the event actually occurred. The bias arising in the general case is often referred to as the "curse of knowledge." In the present paper we will not differentiate curse of knowledge and hindsight bias.

Let $\tilde{v}$ be a random variable with ex–ante expectation μ. Suppose the agent is told the realization v and asked to report his a priori expectation. If rational, he reports μ. If hindsight biased, he fails to remember how ignorant he was initially and forms a biased recollection of the prior mean, tilted towards the realization of the random variable. This can be modelled by specifying the reconstructed prior mean of the biased agent as a weighted average of the true prior mean and of the realization of the random variable:

$$\hat{E}_v(\tilde{v}) = \omega v + (1-\omega)\mu \tag{1}$$

where the constant ω measures the magnitude of the bias and $\hat{E}_v$ denotes the biased expectation. The subscript denotes that the biased expectation is reported after observing the realization v of $\tilde{v}$. This formulation is similar to equation (1) in Camerer et al (1989). When asked to report the a priori mean of the distribution, rational agents are not influenced by the realization of the random variable so that $\omega = 0$. In contrast, biased agents partially forget the ex–ante expectation. If they are completely biased, they totally forget the rational prior mean and place all the weight on the realization of the random variable ($\omega = 1$).

2.2 Adverse consequences of the hindsight bias

Several previous papers have discussed adverse consequences of the hindsight bias. For example, Mangelsdorff and Weber (1998) and Madarasz (2008) show that, in a principal agent relation, the hindsight bias will prevent the principal from correctly evaluating the performance of the agent. Indeed, biased principals fail to remember what was known when the agent's decision was taken, which is in line with the finding of Baron and Hershey (1988). Hlzl et al (2002) find that, when asked questions about economics before and after the introduction of the Euro, Euro supporters showed stronger hindsight bias for positive developments, while opponents showed stronger bias for negative developments.

One of the main disadvantages of the hindsight bias is that it prevents rational processing of information and learning from past data. Thus, Fischhoff (1982) notes that the hindsight bias will prevent one from rejecting one's hypotheses about the world:

> "When we attempt to understand past events, we implicitly test the hypotheses or rules we use both to interpret and anticipate the world around us. If, in hindsight,

> we systematically underestimate the surprises that the past held and holds for us, we are subjecting those hypotheses to inordinately weak tests and, presumably, finding little reason to change them." (Fischhoff, 1982, page 343.)

Similarly, observing students analyzing business cases, Buksar and Conolly (1988) found that the hindsight bias hindered learning from past experience.

In financial markets the inability to be surprised, to learn from the past and to reject hypotheses can be very damaging. Hindsight biased traders will fail to recognize that their view of the market was wrong. Hence they will fail to cut their losses when it is optimal to do so. Hindsight biased investors will inaccurately take into account the informational content of new signals, such as earnings announcement or macro–news. This will lead them to form suboptimal portfolios.[4] In the remainder of this subsection we focus on the adverse consequences of the hindsight bias for risk assessment.

Risk is one of the most important elements of the environment of financial decisions. Correctly factoring risk in these decisions is made difficult by random fluctuations in volatility. Because of these fluctuations, it is important for agents to conduct efficient learning about volatility. The hindsight bias prevents such efficient learning. To see this consider the following simple model. With probability λ the random variable $\tilde{v}$ is normally distributed with variance $\underline{\sigma}^2$ and with probability $1-\lambda$ it is normal with variance $\bar{\sigma}^2$, where $\bar{\sigma} > \underline{\sigma}$. After observing one realization of $\tilde{v}$, the agent has to update her beliefs about the true variance. Applying Bayes law, rational agents would update the probability that the variance is low to:

$$P(\underline{\sigma}|v) = \frac{f(v|\underline{\sigma})\lambda}{f(v|\underline{\sigma})\lambda + f(v|\bar{\sigma})(1-\lambda)}.$$

That is:

[4]The previous version of this paper included a formal analysis of such inefficient portfolio formation (see Biais and Weber (2007)).

$$P(\underline{\sigma}|v) = \frac{1}{1 + \frac{1-\lambda}{\lambda}\frac{\underline{\sigma}}{\bar{\sigma}} e^{-\frac{1}{2}\left(\frac{1}{\bar{\sigma}^2} - \frac{1}{\underline{\sigma}^2}\right)(v-\mu)^2}}. \tag{2}$$

Hindsight biased agents will proceed to a similar updating, except that, instead of using $E(\tilde{v}) = \mu$, they will rely on the biased expectation: $\hat{E}_v(\tilde{v}) = \omega v + (1-\omega)\mu$. This leads to the following biased probability:

$$\hat{P}_v(\underline{\sigma}|v) = \frac{1}{1 + \frac{1-\lambda}{\lambda}\frac{\underline{\sigma}}{\bar{\sigma}} e^{-\frac{1}{2}\left(\frac{1}{\bar{\sigma}^2} - \frac{1}{\underline{\sigma}^2}\right)(1-\omega)^2(v-\mu)^2}}. \tag{3}$$

Comparing (2) and (3), we obtain the following proposition:

Proposition 1: *In our framework, when agents must conduct learning about volatility, those with hindsight bias underestimate volatility.*

Agents increase the probability that variance is large when they observe realizations that differ a lot from their prior expectations, i.e., when they are surprised. Hindsight biased agents tend to not be surprised: "They knew it all along." Hence, they underestimate volatility.

Underestimation of risk, induced by the hindsight bias, will adversely affect decision making. Biased traders will underestimate the value of options and hedging strategies. They will inaccurately assess risk return tradeoffs. They will incorrectly estimate risk premia. Hence, the hindsight bias reduces performance.

2.3 Measuring the hindsight bias

Three main different empirical designs have been used to measure the hindsight bias.

Fischhoff and Beyth (1975) use a within person design where subjects are first asked to report their ex-ante expectations. Then, they learn the realization of the variable and are asked to report their ex-post recollection of their ex–ante expectations.

Fischhoff (1975) complements this design by adding a control group. At the stage where subjects are asked to report their recollection of their initial estimate two groups are formed. In one group, they receive no information. In the other group they are told the true outcome of the event.

Camerer et al (1989) use a third design, also with two groups. In group two, subjects are told the true outcome of the event and asked to estimate the average expectation of group one (knowing that that group had no information).

To pave the way for our empirical analysis in the next section, we now analyze in our simple model the measures corresponding to designs 1 and 3. Before being told the answer to a question, people don't know it exactly. For them it is a random variable. In line with the notation of our simple model, we denote it by $\tilde{v}$. For simplicity we assume information is homogeneous.[5]

In this context, first consider the "within person design" of Fischhoff and Beyth (1975). In that design, the hindsight bias can be measured by the ratio of i) the agent's remembered estimate ($\hat{E}_v(\tilde{v})$) minus his initial estimate (μ) to ii) the difference between the realization of the variable (v) and the agent's initial estimate (μ). Substituting $\hat{E}_v(\tilde{v})$ from equation (1), this ratio simplifies to : ω.

Next, turn to the "between subjects design" of Camerer et al (1989). They form two groups (A and B) and ask a question to the participants in group B. Denote by W_B the average answer of group B participants. The experimenter gives the answer v to participants in group A. Then the experimenter asks these participants to assess the average answer of group B. Denote by $\hat{E}_v^a(W_B)$ the assessment of W_B by participant a in group A. The index of hindsight bias used by Camerer et al (1989) is:

$$\frac{\hat{E}_v^a(W_B) - W_B}{v - W_B}, \tag{4}$$

[5]The previous version of this paper included an extension of this analysis to the case where agents observe private signals (see Biais and Weber (2007)).

i.e., the difference between a's assessment of the average group B's answer minus the actual average, divided by the difference between the true answer and group B's average answer. As noted by Camerer et al (1989) rational agents should on average correctly estimate W_B, hence the average hindsight bias index is 0. In contrast, biased agents will be influenced by the revelation of v. If completely biased, they will believe the others gave the exact right answer, and the index will be 1. If partially biased they will believe:

$$\hat{E}^a_v(W_B) = \omega^a v + (1 - \omega^a)\mu,$$

Noting that all participants in group B should give the same answer μ so that $W_B = \mu$, and substituting W_B and $\hat{E}^a_v(W_B)$ in equation *(4)*, we obtain our next Proposition.

Proposition 2: *In our framework, the Camerer et al (1989) index for each participant is equal to his (her) hindsight bias parameter (ω^a).*

3 Study 1: hindsight bias and risk perception

The goal of this experiment is to test whether hindsight biased agents fail to remember ex–post how much uncertainty there was ex–ante, and thus come up with lower risk estimates than unbiased agents. To do so we compare two treatments. In the first treatment, the hindsight bias is muted because, ex–post, the experimenter reminds the participants of their ex–ante estimates. In the second treatment, the hindsight bias can arise because the participants are not told their initial forecasts, rather they are asked to remember these forecasts. We test the hypothesis that participants give higher volatility estimates in the first treatment than in the second one.

3.1 Experimental design

The experiment was run in the context of the behavioural finance class, taught by one of the authors (Martin Weber) to fourth year students in Mannheim University. 67 students participated in the experiment, which involved two tasks.

The first task was on 29 November 2007. We gave the participants the market price of 5 German stocks (BASF, IKB, EON, Postbank, Premiere) and 2 commodities (oil and gold) as well as the euro dollar exchange rate, for November 21 and November 28. We asked them to predict the price one week later (on December 5). We also asked them to give an upper bound and a lower bound such that there would be only one chance out of ten that the December 5 price would be outside the bounds. We then transformed this confidence interval into a standard deviation using the method introduced by Keefer and Bodily (1983). Thus we directly elicited their estimate of the expected price and indirectly estimated their estimate of the standard deviation. We resorted to such an indirect elicitation because it is more natural and intuitive for participants to give a confidence interval than a standard deviation.

The second task was a week later, on December 6. At that point we gave the participants the realized price for December 5 (and also reminded them of the November 28 price). We asked them again to give us one week ahead (for December 12) expectations and confidence intervals for the 5 stocks, the 2 commodities and the exchange rate.

Students were randomly assigned to one of two treatments. In the first treatment, on December 6, we gave the participants their estimates (prediction and bounds) from last week (formed on November 29 for December 5). In the second treatment we did not remind the participants of their estimates from the previous week. Rather we asked them to remember and write down these estimates (prediction and bounds). Thus, while in the first treatment, we muted the hindsight bias by explicitly reminding the participants of their forecasts, in the second treatment the hindsight bias could manifest itself. 35 students participated in the first treatment, 32 participated in the second treatment. At each round, we randomly drew 10 participants who each received 10 euros for participation.[6]

[6]It's likely that stronger financial incentives would not have led to significantly different results. As stated by

3.2 Results

Our empirical results are summarized in Table 1. First consider the hindsight bias. As mentioned above, in this "within person design"the hindsight bias (ω) can be measured by the ratio of i) the difference between the agent's remembered estimate and his initial estimate to ii) the difference between the realization of the variable and the agent's initial estimate. For each participant in the second treatment, and for each asset, we compute this ratio. The first row of Table 1 reports (for each asset) the median across agents for this score (as well as, in parentheses, the p–value for the hypothesis that the score is equal to 0).For all eight assets the median score is positive (and in 4 cases out of 8 it is significantly different from 0 at the 10% level). Thus participants are hindsight biased on average.

As mentioned above, the hindsight bias stems from the inability of the agents to remember ex–post how much uncertainty there was ex–ante. Our experiment offers an opportunity to document the extent to which people are able to remember previous uncertainty. In the second treatment, on December 6, we asked the participants to remember and state the confidence interval they had previously indicated (on November 29). The initial confidence interval (given on November 29) and the corresponding implied volatility, reflected the participants' ex–ante perception of uncertainty. The remembered confidence interval (given on December 6) and the corresponding implied volatility, reflected how much the agents remembered they were uncertain ex–ante. Thus, for each of the 8 assets, we compute the median across participants in the second treatment of the ratio of i) the volatility estimate implied by their initial confidence interval on November 29 to ii) the volatility estimate implied by their remembrance on December 6. The

Camerer and Hogarth (1999, page 8), drawing the lessons of the experimental literature: "In the kind of tasks economists are most interested in, like trading in markets, bargaining in games and choosing among risky gambles, the overwhelming finding is that increased incentives do not change average behaviour substantially (although the variance of responses often decreases)."

medians are presented in the second row of Table 1 (along with, in parentheses, the p–values for the hypothesis that the ratio is equal to 1). For 5 assets out of 8 the median ratio is significantly above 1 at the 10% level. The median ratio is strictly below one for only one asset (and in that case the difference to 1 is not statistically significant). This also suggests that participants are hindsight biased on average. As far as we can tell, this is a new way to document this bias.

As discussed in the previous section, hindsight biased agents are "never surprised" . In our experiment, in line with the model presented in the previous section, we measure surprise as the square of the difference between the realization and the expectation, divided by the standard deviation. When this surprise is measured using the initial expectation and volatility estimates of the agents, we refer to it as the "true surprise."When it is measured using the remembered expectation and volatility estimates, we refer to it as the "biased surprise." Of course, while the true surprise can be computed in both treatments, the biased surprise can be computed in the second treatment only. The biased surprise is a way to jointly assess the effect of the hindsight bias on expectations and on variances. For each asset and each agent we computed the true surprise and, in the second treatment, the biased surprise. We then took the medians across agents. These medians are depicted in Table 1, in the third, fourth and sixth rows.[7] As expected, there is no systematic ranking between the true surprises between the two treatments, and the differences are not significantly different from 0 (the p–levels are above 10% for all 8 assets). In contrast, for the second treatment, the biased surprise is systematically smaller than the true surprise. However, the difference is significant in only two cases (Gold and dollar).

The main objective of this study was to test the implication from our theoretical analysis that the inability to be surprised leads biased agents to come up with lower volatility estimates than unbiased agents (see Proposition 1). To document this point we computed, for each asset and each participant, the ratio of the initial one–period ahead volatility estimate (formed on

[7] The figures have been rounded to integers.

November 29) to the subsequent one–period ahead volatility estimate (formed on December 6).We then took the medians across agents. These medians are depicted in the last two rows of Table 1. For all assets except Gold, the median ratio is lower for the first treatment, and in 4 cases the difference is statistically significant. Thus the upward revision of volatility forecasts is stronger for participants who are shown their initial estimates than for participants who are asked to remember initial estimates. This is consistent with the hypothesis that hindsight biased agents end up with lower volatility updates, because they are never surprised.

Finally note that our design offers the opportunity to observe two variables related to the hindsight bias: On one hand, we measure ω as the ratio of i) the difference between the agent's remembered estimate and his initial estimate to ii) the difference between the realization of the variable and the agent's initial estimate. On the other hand, people with stronger hindsight bias are predicted to have a greater ratio of initial implied volatility to remembered implied volatility. For each of the 8 assets in our study, we computed the correlation across participants between these two variables. The Spearman correlation coefficients (and corresponding p–values) are: .134 (.463) for BASF, -.006 (.974) for IKB, -.018 (.920) for EON, .288 (.109) for Postbank, .029 (.877) for Premiere, .465 (.007) for oil, .132 (.471) for gold, and .132 (.473) for dollar. Consistent with our theory, the vast majority of the coefficients is positive. But the result is not very strong, since the correlation is significantly different from 0 only in one case (oil).

4 Study 2: hindsight bias and performance

While the goal of Study 1 was to analyze the effects of the hindsight bias on the way in which agents process information, the goal of Study 2 is to test its effects on performance.

4.1 Data Collection

To assess the effect of the hindsight bias on performance in the field we collected psychometric and financial data for bankers working in a large investment bank, in Frankfurt and London. The psychometric data used in the present study was collected as part of a larger seven page questionnaire where other psychological constructs were also measured.[8] In the present study we use data relative to the hindsight bias, and also, to control for other possible effects, the miscalibration bias and the better than average effect.

The participants in the study are from the trading department of the bank. While they have various tasks (sales, research and trading), they all participate in the elaboration of portfolio allocation, trading and investment decisions. The data was collected by inviting groups of participants into one of two conference rooms in the bank. There, they filled out the questionnaire under the supervision of an experimenter. After finishing the questionnaire, which took about 30 minutes, participants were asked not to talk to their colleagues until the end of the data collection. The overall data collection effort took about two hours. The data was collected anonymously with each questionnaire being labelled with a number. This number was used to sort the questionnaires into earnings categories. For each number the personnel department of the bank informed us whether the respondent was in the top, middle or low earnings category. Earnings meant overall compensation, including bonus. Sorting subjects into three categories was a compromise between our wishes (as much performance data as possible) and the privacy requirements of the participants and institution. It should be noted that the distribution of earnings is skewed (as suggested by common wisdom and an interview with the human resource department). Therefore earnings in the low earnings and medium classes are less variable than in the high earnings class.

[8]These variables, together with similar ones collected from another bank are analysed in Glaser et al (2005).

We first collected the data in the Frankfurt branch of the bank, on December 5, 2001. We had 41 respondents (each member of the department who was present that day). Then, to assess the robustness of our results, we collected similar data, in the London branch of the bank, on October 9, 2002. There we had 49 respondents, corresponding to the vast majority of the department.[9] Bankers in both locations had similar jobs. In Frankfurt, the questionnaire had to be approved by the management and by the worker's council. We made it very explicit that the study was done by two universities and that the management of the bank would not see any specific questionnaire. The only data made available to participants and management were aggregate data. No payments were given.

4.2 Descriptive statistics

Table 2 offers some descriptive statistics on the population of bankers in our sample. It shows that low and mid earnings bankers have approximately the same age in Frankfurt and London. High earnings bankers are considerably older in London. In Frankfurt the bankers are relatively equally distributed across the different departments (sales, trading, research and other). Unfortunately, the London traders did not fill in our questionnaire. Hence, the bankers in our London subsample only belong to sales and research.

To estimate the hindsight biases of the investment bankers, we used the "between person" experimental design of Camerer et al (1989). We feared that highly competitive bankers would find it difficult to admit they could not perfectly forecast the future. Such considerations could have contaminated hindsight bias measures obtained in a "within person" design (see e.g., Campbell and Tesser (1983)). Also, as Rabin (1998) points out, as economists, we care about a

[9] In London we received 49 questionnaires. However, only 44 participants were grouped into earnings categories. Hence, as regards the London subsample, the test of hypotheses is based on 44 subjects whereas the pure hindsight bias results are based on 49 subjects.

person's belief about others and our between person design captures this aspect.

Our experimental design involved two types of tasks for each participant. In the first task, they were asked to estimate the true value of some unknown variables (own estimate). In the second task, they had to guess the average answers of the other group to some questions (others estimate). In that second task, the participants were given the true answer to the question asked to the other group. Each subject had to answer ten questions, five in the own estimate condition and five in the others estimate condition. We formed two groups in each location, hereafter referred to as Group A and B.

The items were taken from the natural habitat of the bankers. Nevertheless, we used questions for which we thought they did not know exactly the correct answer. The items were similar for Frankfurt and London. Whenever it was reasonable, the questions were exactly the same for the two locations. The second sample however, was collected 10 months later than the first and in a different country. Hence, we had to change some questions to ensure comparability. The questions are listed in Table 3.

Table 3 reports the average hindsight bias, measured using the Camerer et al (1989) index. Participants in our study are hindsight biased on average, although the questions are from the natural habitat of the bankers and that the participants are experienced professionals. The degree of hindsight bias is different for different questions but between 0 and 1 in the majority of the cases. And the magnitude of the hindsight bias score observed in this between person design is similar to that observed in the within person design of Study 1. To summarize the data we computed the average bias across participants and across the 5 questions. The mean hindsight bias is very similar in London (.547) and in Frankfurt (.579). It is significantly lower than 1 and above 0 in both locations. The t–statistics are: 3.60 and 4.35 respectively for London, and 3.43 and 4.72 respectively for Frankfurt.

4.3 Psychometric issues

4.3.1 Precision of information

Does our measure of the hindsight bias really capture this bias, or spuriously reflects other variables? One possibility is that it simply mirrors the precision of the knowledge of the participants regarding the questions we asked them. To check this we estimated a measure of this precision. In the own estimate treatment, for each participant and each question, we computed the absolute value of the difference between the participant's answer and the true answer. Then we summed these mistakes across questions. Finally, we normalized the total mistake of each participant by the average in his group. The greater this ratio, which we hereafter refer to as the mistake ratio, the lower the precision of the information of the participant. To examine if the hindsight bias is related to the mistake ratio, we computed the correlation between the two. The Spearman correlation coefficient is -.0837. The probability to observe such a value under the null that the two variables are uncorrelated is 44.6%. Hence, the mistake ratio and the hindsight bias are weakly and insignificantly negatively correlated.

4.3.2 Underconfidence and overconfidence

Another possibility is that our measure of the hindsight bias actually reflects underconfidence. The agent would believe the others to be better informed than him and than they really are. In this context, while the agent would be surprised by the realization of the random variable, he/she would believe that the others had predicted it very accurately and were not surprised. While *in theory* underconfidence could generate positive values for the Camerer et al (1989) index, *in practice* this is unlikely to be driving the results in our experiment. As mentioned above, in our questionnaire we collected the answers to "better than average" questions. Participants were asked to answer what percentage of their colleagues they expected to perform better than themselves along four dimensions (trading skills, communication skills, market vision, technical skills.) Averaging across the four dimensions, and across participants, the bankers in our sample answered that 24.61% of the others had better skills than themselves. This suggests overconfidence rather than underconfidence. In addition, the correlation across bankers of the better than average score with the hindsight bias index is not significantly different from 0. For the Frankfurt bankers it is 0.055 (with a p–value of 0.733). For the London bankers, the correlation is 0.0975 (p–value = 0.51.)

4.3.3 Miscalibration

Another form of overconfidence which has received attention in finance is miscalibration (see Biais et al (2005)). In line with Russo and Schoemaker (1992) and Klayman et al (1999), we measure miscalibration by eliciting confidence intervals. We show participants ten questions. For each question we ask them to give an upper bound and a lower bound such that the true answer should lie outside the bounds with only one chance in ten. If participants were correctly calibrated, the frequency of answers lying outside the bounds should be 10%. In our sample, the frequency of such answers was 68%. Such strong miscalibration, is consistent with the findings of Klayman et al (1999) and Russo and Schoemaker (1992). In line with this literature we thus use as miscalibration score the percentage of questions for which the true answer fell outside the indicated bounds. To assess if miscalibration is independent from the hindsight bias, we computed the correlation between the miscalibration score and the hindsight bias score. For the London sample this correlation is equal to -.1 while for the Frankfurt sample it is equal to .16. In both samples, the correlation is not significantly different from zero.

4.4 Empirical results

Up to that point we have demonstrated that investment bankers exhibit hindsight bias. We have found that our measure of hindsight is not significantly correlated with other variables such as information precision and overconfidence. But does the hindsight bias affect performance? We now turn to that question.

4.4.1 Main result

As can be seen from Table 4, the hindsight bias (median as well as mean) is lower for the high earnings category than for the two other categories. The average hindsight bias is somewhat lower for the low earnings category than for the middle earnings category. Note however that when one focuses on medians the difference is not large. Table 4 also shows that, for the middle earnings category, the median bias is much lower than the mean, while for the high earnings category, the mean bias is much lower than the median, and for the low earnings category, there is no clear pattern. This suggests that the mean in the high earnings category is driven down by a few participants with very low bias, while the mean in the middle earnings category is driven up by a few participants with very high bias.

In addition, Table 4 shows that the relation between the hindsight bias of bankers and their performance is prevalent in both locations. Both in London and in Frankfurt the bankers with the highest performance are also those with the lowest bias. Such robustness speaks to the out–of–sample validity of the result. It is particularly striking since, as was mentioned above, we first designed and implemented the experiment with Frankfurt participants, and then, to evaluate the robustness of our results, we replicated the same experiment with London participants.

Table 4 also suggests that the result is not driven by other variables, such as age or experience. Indeed, the age and experience structure differs in the two locations. While in London the high earnings bankers have longer experience than the middle earnings bankers (18 versus 9.9), in

Frankfurt this is not the case (11.8 versus 12.5. See Table 2).

To evaluate the significance of our results we rely on a non-parametric Wilcoxon rank test (z-stat). The difference between the bias of the high earnings bankers and that of the middle earnings bankers is significant, with a z–statistics of 3.33 and a p–value of .09%. The difference between high earners and low earners is also significant, with a z–statistics of 2.27 and a p–value of 2.29%. The difference between low and mid earners is not significant, with a z–statistics of 1.02 and a p–value of 30.6%. This lack of significance is consistent with the skewness of the earning distribution mentioned above. Some stars receive very high compensation, rewarding exceptional performance, while the bulk of the bankers have relatively similar compensation. Consequently, there is only a small difference between the monetary compensation of the low earnings categories and that of the middle earnings categories.

Table 5 reports the estimated probabilities of the performance of the bankers conditionally on their biases. We classified participants as having low (resp. high bias) bias if they were among the 33% least (resp. most) biased of their occupational category. Conditional probabilities were estimated as empirical frequencies. Panel A reports the results obtained for the whole bankers' sample. Consistently with the hypothesis that the hindsight bias reduces performance, the largest number in each column is on the first diagonal: bankers with the lowest bias are most likely to be high–earners, bankers with medium bias are most likely to be mid–earners, and bankers with the highest bias are more likely to be low–earners. To test the null hypothesis that performance is independent from bias, we performed a Chi square test. Since there are three rows and three columns in Table 5, the relevant statistic is a Chi Square with 4 degrees of freedom, for which the critical value at the 5% level is 9.49. The distance between the empirical contingency table and the theoretical one (computed under the null) was found to be 10.05. The null hypothesis that performance and bias are independent can thus be rejected at the 5% level.

4.4.2 Robustness

Experience Experience could influence the bias of the bankers and its impact on their performance. For example, List (2003) provides evidence on the consequences of market experience on market anomalies. To control for experience we split the sample of bankers in two groups. The first group includes the bankers with less than 10 years of experience ("juniors") and the second group includes the bankers with 10 years of experience or more ("seniors"). For juniors, the median bias is .61 for bankers with low–earnings, .85 for bankers with medium earnings and .06 for bankers with high earnings. For seniors, the median bias is .77 for bankers with low–earnings, .64 for bankers with medium earnings and .43 for bankers with high earnings. Thus, the result that bankers with high performance are least biased obtains in both age groups. Our findings also suggest that experience does not reduce the bias.

Jobs Some participants in the experiment were employed in research, others in sales, and yet others in trading (or sales and trading). As an additional robustness check, we analyzed if our results obtained for each of these three different occupational categories. Panel B,C and D of Table 5 split the estimation of conditional probabilities according to occupational categories. The results suggest the link between bias and performance is reasonably robust across occupational categories in the bank. Consistently with the hypothesis that the hindsight bias reduces performance, bankers with the lowest bias are most likely to be high–earners for research and for sales. Bankers with medium bias are most likely to be mid–earners, for sales and for trading. And bankers with the highest bias are more likely to be low–earners for research and for trading. In addition, for research and in trading, no banker with high bias has high earnings.

Mistakes Yet another robustness check is to examine whether our results still hold after one controls for the precision of the information of the participants, relative to the questions asked. As discussed above, for each participant, we computed a mistake ratio, decreasing with the precision of the answers of the agent in the own answer treatment. Interestingly, the mistake ratio varies somewhat with performance. Participants with high performance have an average mistake ratio of .890, while participants with medium performance have a mistake ratio of 1.004, and participants with low performance have a mistake ratio of 1.051. Yet, the correlation between the mistake ratio and the hindsight bias is not significant. The Wilcoxon rank sum test comparing the mistake ratio of high and medium earnings is $z = .549$, with a $p-$value of .583. The Wilcoxon rank sum test comparing the mistake ratio of high and low earnings is $z = 1.206$, with a $p-$value of .2278. And the Wilcoxon rank sum test comparing the mistake ratio of medium and low earnings is $z = .595$, with a $p-$value of .5516. Furthermore, as mentioned above, the mistake ratio is not significantly correlated with the hindsight bias.

To control for mistake ratios, we divided the population of bankers in three categories. The first category includes the 28 bankers with the lowest mistake ratios (their average mistake ratio is .425). The second category includes the 28 bankers with intermediate mistake ratios (their average mistake ratio is .765). The third category includes the 29 bankers with the highest mistake ratios (their average mistake ratio is 1.745). We also divided bankers in three bias categories, similarly to Table 5. We then estimated, within each "mistakes category", the probability to be in the high earnings category conditional on their bias. We found that, in the three "mistakes categories," bankers with the lowest bias were the most likely to be high–earners: For the 33% least biased bankers, the probability of being in the high earnings category is 55% for low mistake ratios group, .45% for the medium "mistakes category," and .34% for the high "mistakes category". For the medium and high bias bankers, the probability of being in the high earnings category is between 20% and 33% across all "mistakes categories."

Overconfidence Other psychological biases than the hindsight bias could affect performance.[10] To control for such possible effects, we use the two other psychometric variables we measured: the better than average score and the miscalibration score. Table 6 presents their mean and median for the three earnings group. As can be seen in the table, there is no strong variation in biases from one earnings group to the other. The median miscalibration score is .726 for low earnings bankers and .650 for high earnings bankers. That is, for low earnings bankers the true answer fell outside the bounds given by the participants 72.6% of the time, while for the high earnings bankers it did so 65% of the time. The median better than average score is .207 for the low earnings bankers and .261 for the high earnings bankers. That is, low earnings bankers answered that 20.7% of their colleagues were better than themselves, while high earnings bankers answered that 26.1% of their colleagues were better than themselves. Such small (and non significant) variations from one earnings group to the other contrast with the large variations observed for the median hindsight bias, which falls from .636 for the low earnings bankers to .375 for the high earnings bankers.

5 Conclusion

Do cognitive biases affect information processing and performance in financial markets? In this paper we addressed that question, focusing on the hindsight bias. Agents who exhibit this bias fail to remember how ignorant they were before observing outcomes and answers. We show that this hinders learning, and, in particular, lead agents to underestimate volatility. This results in inefficient portfolio choice, loss–making trades and poor risk management. We rely on two experimental studies to test these claims.

In the first experimental study, we focus on the consequences of the hindsight bias on learning

[10] It would also have been interesting to control for other psychological variables, such as risk–aversion. Unfortunately, we did not collect data on this variable.

about volatility. We compare two treatments: one in which the bias is muted, and the other where it can manifest itself. Agents give lower volatility updates in the latter treatment than in the former, as implied by our model.

In the second experimental study, we test the hypothesis that the hindsight bias hurts financial performance. We collect psychometric and performance data about highly paid investment bankers. We find that they exhibit hindsight bias when asked questions about economics, banking and finance, and that experience does not reduce this bias. Most importantly, we find that bankers with low bias obtain significantly better performance.

Given that our results suggest the hindsight bias matters, it would be interesting to provide a deeper theoretical analysis of this bias. Could it emerge from cognitive constraints, such as limited memory? Can our model be extended to accommodate a more general "projection bias" phenomenon (see Loewenstein, O'Donoghue and Rabin, 2003) ? And how would biased agents strategically interact in a trading game? We leave these issues for further research.

References

Baron, J., and J. Hershey, 1988, "Outcome Bias in decision evaluation," *Journal of Personality and Social Psychology*, 54, 569-579.

Biais, B., D. Hilton, K. Mazurier, and S. Pouget, 2005, "Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market," *Review of Economic Studies*, 72, 287-312.

Biais, B., and M. Weber, 2007, "Hindsight bias and investment performance," Working Paper # 476, IDEI, Toulouse.

Bukszar, E. and T. Connolly, 1988, "Hindsight bias and strategic choice: Some problems in learning from experience," *Academy of Management Journal*, 31, 628-641.

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

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

Camerer, C., and R. Hogarth, 1999, "The effects of financial incentives in experiments: A review and capital-labor-production framework," *Journal of Risk and Uncertainty*, 19, 7-42.

Campbell, J. and A. Tesser, 1983, "Motivational Interpretations of hindsight bias: An individual difference analysis," Journal of Personality, 51, 605-620.

Carli, L., 1999, "Cognitive Reconstruction, Hindsight, and Reactions to Victims and Perpetrators," *Personality and Social Psychology Bulletin*, 25, 966–979.

Christensen-Szalanski, J. and C. Wilham, 1991, "The hindsight bias: A meta–analysis," *Organizational Behavior and Human Decision Processes*, 48, 147–168.

Fenton O'Creevy, M., N. Nicholson, E. Soane, and P. Willman, 2003, "Trading on illusions: Unrealistic perceptions of control and trading performance, " *Journal of Occupational and Or-*

ganizational Psychology, 76, 53-68.

Fischhoff, B., 1975, “Hindsight ≠ Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty,” *Journal of Experimental Psychology: Human Perception and Performance*, 1, 288-299.

Fischhoff, B., 1982, “For Those Condemned to Study the Past: Heuristics and Biases in Hindsight,” In D. Kahneman, P. Slovic, and A. Tversky (eds.), Judgment Under Uncertainty: Heuristics and Biases, NY: Cambridge University Press, 332-351.

Fischhoff, B., and R. Beyth, 1975, “I knew it would happen – Remembered probabilities of once-future things,” *Organizational Behavior and Human Performance*, 13, 1–16.

Glaser, M., T. Langer, and M. Weber, 2005, “Overconfidence of Professionals and Laymen: Individual Differences Within and Between Tasks?,” Working Paper, Mannheim University.

Glaser, M. and M. Weber, 2007, “Overconfidence and Trading Volume,” The Geneva Risk and Insurance Review, 32, 1-36.

Haigh, M., and J. List, 2005, “Do professional traders exhibit myopic loss aversion? An experimental analysis,” *The Journal of Finance*, 60, 523-534.

Hawkins, S., and R. Hastie, 1990, “Hindsight: Biased Judgments of Past Events After the Outcomes are Known,” *Psychological Bulletin*, 107, 311-327.

Hlzl, E., E. Kirchler, and C. Rodler, 2002, “Hindsight Bias in Economic Expectations: I Knew All Along What I Want to Hear,” *Journal of Applied Psychology*, 87, 437-443.

Klayman, J., J. B. Soll, C. Gonzales-Vallejo, and S. Barlas, 1999, “Overconfidence: it depends on how, what and whom you ask, ” *Organizational Behavior and Human Decision Processes*, 79, 216-247.

Keefer, D. L., and S. E. Bodily, 1983, “Three-Point Approximations for Continuous RandomVariables,” *Management Science*, 29, 595-609.

List, J., 2003, “Does market experience eliminate market anomalies?” *Quarterly Journal of*

Economics, 118, 41–71.

Loewenstein, G., O'Donoghue, T. and Rabin, M., 2003, "Projection bias in predicting future utility. "*Quarterly Journal of Economics*, 118, 1209-1248.

Madarasz, K, 2008, "Information projection: Model and applications," Working Paper, Berkeley.

Mangelsdorff, L. and M. Weber, 1998, "Hindsight Bias im Prinzipal Agent Kontext," in H. Glaser, E. Schröder, and A. v. Werder (eds.), Organisation im Wandel der Märkte, Gabler, 461-478.

Musch, J., 2003, "Personality differences in hindsight bias," *Memory*, 11, 473-489.

Rabin, M., 1998, "Psychology and Economics," *Journal of Economic Literature*, 35, 11-46.

Roese, N. J., 2004, "Twisted pair: Counterfactual thinking and the hindsight bias," in D. Koehler and N. Harvey (eds.), Blackwell handbook of judgment and decision making, Oxford: Blackwell, 258-273.

Russo J. and P. J. H. Schoemaker, 1992, "Managing overconfidence," *Sloan Management Review*, 33, 7-17.

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

Wasserman, D., R.O. Lempert, and R. Hastie, 1991, "Hindsight and causality," *Personality and Social Psychology Bulletin*, 17, 30–35.

Table 1: Hindsight bias and risk perception in Study 1

On 29 November, 2007, participants were asked to predict prices one week later and give confidence intervals, from which we backed out implied volatilities. On December 6 they performed the same task again. In the first treatment they were told their previous estimates. In the second one they were asked to remember their estimates. The table reports medians across agents. p–values are in parentheses.

	BASF	IKB	EON	Postbank	Premiere	Oil	Gold	Dollar
Hindsight bias Second treatment	0.086 (0.344)	0.463 (0.009)	0.155 (0.166)	0.784 (0.000)	0.375 (0.626)	0.182 (0.074)	0.094 (0.052)	0.264 (0.108)
Initial implied σ/ remembered σ Second treatment	1.142 (0.082)	1.267 (0.002)	1.000 (0.822)	1.577 (0.000)	1.273 (0.000)	1.090 (0.087)	1.000 (0.517)	.944 (0.525)
True surprise. First treatment	2340	25	57004	737	12	3320	418167	1.5
True surprise. Second treatment	3026	22	34649	717	18	5867	226082	1.1
Rank sum test difference between Both treatments	(0.969)	(0.787)	(0.150)	(0.435)	(0.817)	(0.777)	(0.164)	(0.472)
Biased surprise Second treatment	2073	8	21007	215	38	5504	187809	.7
Rank sum test difference between true & biased surprise	(0.910)	(0.667)	(0.158)	(0.866)	(0.278)	(0.765)	(0.024)	(0.041)
Initial implied σ/ subsequent implied σ First treatment	.923	1.024	.630	1.001	1.041	.981	.996	.892
Initial implied σ/ subsequent implied σ Second treatment	.978	1.373	0.858	1.854	1.407	1.003	0.846	0.910
Rank sum test for difference between the 2 treatments	(0.172)	(0.021)	(0.099)	(0.002)	(0.001)	(0.349)	(0.520)	(0.369)

Table 2: Bankers population in Study 2

Location	Earnings	Number	Average Age	Average Experience
FRANKFURT	High	12	34	11.83
	Mid	12	34.50	12.50
	Low	17	29.76	7.06
LONDON	High	14	41.86	18
	Mid	14	34.07	9.89
	Low	16	32.38	6.36

Table 3, Hindsight bias questions and results in Study 2.

The figures in the cells are averages computed across participants. The first column reports the true answer to the question, the second column reports the average answer of the group which was asked that question, the third column reports the guess of the other group about that answer. The last column reports the hindsight bias, measured using Camerer et al (1989) index.

Panel A: Frankfurt sample.

Question	True answer	Other group answer	This group prediction (given true answer)	Bias
FRANKFURT A				
Consumer price change Germany 10/2000 to 10/2001	2.5	1.68	2.9	1.49
Drop of Swiss stock market from all time high to 10/2001	30	38.55	35.24	0.39
Change in price of gold 10/2000 to 10/2001	6.2	13.69	6.53	0.96
Number of bankers at Lazard 8/2001	200	3477.27	691.74	0.85
% of Merrill Lynch's earnings from asset management in 2001	10	31.32	20.29	0.52
FRANKFURT B				
Euro/Yen exchange rate 2/13/2001	109	135.85	103.41	1.21
Ratio of foreign debt to GDP in Brazil at the end of 2000	38.4	176.69	59.5	0.85
Net revenues drop Texas Instruments end of third quarter 2001	1285	1034.06	925.05	-0.43
Growth rate of real GDP Russia 2000	8.3	5.57	5.02	-0.20
Growth rate of real GDP OECD countries, 2000	4.1	3.5	3.66	0.27

Panel B: London sample.

Question	True answer	Other group answer	This group prediction (given true answer)	Bias
LONDON A				
Retail Price Index change 10/2000 to 10/2001	1.57	5.37	5.61	-0.06
Drop of Swiss stock market from all time high to 8/2002	39.17	32.06	37.4	0.75
Change in price of gold from 10/2000 to 10/2001	6.2	21.88	10.48	0.73
Bankers at Lazard in August 2001	200	1876.36	641.88	0.74
% Merrill Lynch's earnings from asset management, 2001	10	22.08	18.88	0.27
LONDON B				
Euro/Yen exchange rate 2/13/2002	116.1	135.5	122.22	0.68
Ratio of foreign debt to GDP Brazil, end of 2000	38.4	126.24	57.83	0.78
Net revenue increase TI 2^{nd} quarter 2002	1.06	-4.35	-3.47	0.16
Growth rate of real GDP Russia, 2001	5	4.25	4.92	0.89
Growth rate of real in GDP OECD countries 2001	1.2	3.75	2.12	0.64

Table 4: Hindsight bias for the three earnings categories in Study 2

The bias is measured using Camerer et al (1989) index). Medians, means and std; dev. are computed across subjects.

	Pool			Frankfurt			London		
Earnings	Low	Mid	High	Low	Mid	High	Low	Mid	High
Median	0.636	0.684	0.375	0.512	0.63	0.303	0.636	0.809	0.429
Mean	0.546	0.924	0.192	0.678	0.883	0.133	0.546	0.959	0.242
Std. Dev.	1.009	0.833	0.638	0.737	0.948	0.473	1.009	0.757	0.766
Nb obs	33	26	26	17	12	12	16	16	14

Table 5: Probability of performance, conditional on hindsight bias in Study 2

The bankers were split in three categories of equal sizes. The first one included the 33% bankers with the lowest bias, while the third one included the 33% bankers with the highest bias. Empirical frequencies are used to estimate conditional probabilities. The probabilities in each column add up to one.

Panel A: All bankers

	Low bias	Mid bias	High bias
Prob(High Earnings|bias)	.46	.32	.14
Prob(Mid Earnings|bias)	.14	.39	.38
Prob(Low Earnings|bias)	.39	.29	.48

Panel B: Research

	Low bias	Mid bias	High bias
Prob(High Earnings|bias)	.57	.14	.00
Prob(Mid Earnings|bias)	.0	.14	.50
Prob(Low Earnings|bias)	.43	.71	.50

Panel C: Sales

	Low bias	Mid bias	High bias
Prob(High Earnings|bias)	.57	.29	.21
Prob(Mid Earnings|bias)	.14	.43	.43
Prob(Low Earnings|bias)	.29	.29	.36

Panel D: Trading and Sales-trading

	Low bias	Mid bias	High bias
Prob(High Earnings|bias)	.20	.40	.00
Prob(Mid Earnings|bias)	.20	.40	.25
Prob(Low Earnings|bias)	.60	.20	.75

Table 6: Hindsight bias, miscalibration and better than average for the three earnings categories in Study 2

		Miscalibration	Better than average	Hindsight bias
Low earnings	Mean	0.701	0.207	0.546
	Median	0.726	0.207	0.636
Medium Earnings	Mean	0.681	0.296	0.924
	Median	0.663	0.294	0.684
High Earnings	Mean	0.666	0.245	0.192
	Median	0.650	0.261	0.375