Distinguishing Overconfidence from Rational Best-Response in Markets[1]

Shimon Kogan[2]
Tepper School of Business
Carnegie Mellon University

February 1, 2006

[1] This is an updated version of the paper titled: "Distinguishing Bounded Rationality from Overconfidence in Financial Markets - Theory and Experimental Results". The financial research support of The Center for Responsible Business, The Dean Witter Foundation and IBER is gratefully acknowledged. I would like to thank George Akerlof, Bob Anderson, Stefano DellaVigna, Teck Ho, Botond Koszegi, Lars Lochstoer, Rich Lyons, Ulrike Malmendier, Barbara Mellers, seminar participants at the Carnegie Mellon University, Federal Reserve Bank of Boston, Harvard Business School, London Business School, MIT Sloan School of Management, Tel-Aviv University, University of British Columbia, University of California at Berkeley, University of Texas at Austin, University of Utah, Western Finance Association 2005 Meetings, and Yale School of Management, for their valuable comments. I am particularly indebted to Jonathan Berk, Shachar Kariv, John Morgan, Terry Odean, Jacob Sagi and Nancy Wallace for their insightful suggestions and continuous support. All errors are of course mine alone.

[2] Tepper School of Business, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, tel: (412) 268-8501, fax: (412) 268-7357, email: kogan@andrew.cmu.edu

Abstract

This paper studies the causal effect of individuals' overconfidence and bounded rationality on asset markets. To do that, we combine a new market mechanism with an experimental design, where (1) players' interaction is centered on the inferences they make about each others' information, (2) overconfidence in private information is controlled by the experimenter (i.e., used as a treatment), and (3) natural analogs to prices, returns and volume exist.

We find that in sessions where subjects are induced to be overconfident, volume and price error analogs are higher than predicted by the fully-rational model. However, qualitatively similar results are obtained in sessions where there is no aggregate overconfidence. To explain this, we suggest an alternative possibility: participants strategically respond to the errors contained in others' actions by rationally discounting the informativeness of these actions. Estimating a structural model of individuals' decisions that allows for both overconfidence and errors, we are able to separate these two channels. We find that a substantial fraction of excess volume and price error analogs is attributable to strategic response to errors, while the remaining is attributable to overconfidence. Further, we show that price analog exhibit serial autocorrelation only in the overconfidence-induced sessions.

1 Introduction

Recent studies suggest that overconfidence on the part of traders can rationalize a set of long standing asset-pricing 'anomalies' such as excess trading volume, excess volatility and serially autocorrelated returns [see for example Kyle and Wang (1997), Odean (1998), Daniel et al. (1998)]. To understand the link between overconfidence and financial markets, consider a generic market populated by partially informed traders. Each trader's decision reflects a weighting of her private (yet imperfect) information and the information revealed by the actions of others. Overconfident traders perceive their signal to be more precise than it is, thus irrationally overweighting it. As a result (1) beliefs are more dispersed across traders, leading to greater volume, and (2) prices over-reflect overconfident traders' signals, leading to poorer price informativeness.[1]

We suggest an alternative reason that may cause traders to *rationally* 'overweight' their information relative to the information of others: they strategically respond to errors made by others. The idea that people make mistakes, in the sense that they do not *always* best respond when interacting with others, is well documented [see Camerer (2003)]. Generally, mean-zero mistakes in actions can have two effects: direct and strategic. If errors are added to players' Nash equilibrium strategies, the direct effect would wash out across many observations. However, if in addition players are aware of others' mistakes and react to them, this would lead to a strategic effect that as we show, does not necessarily average out. In this case, traders would discount the informativeness of fellow traders' actions, rationally overweighting their own information. Thus, overconfidence and response to errors both lead to *directionally* similar behavior and may therefore be observationally equivalent.

To separate these competing channels and quantify their relative magnitude on volume, prices and returns analogs, we study (theoretically and experimentally) a new game that explicitly links individual level behavior and asset markets. The experimental setting enables us to control and/or measure individuals' information, preferences and beliefs, which are key determinants of their decisions. While previous experimental studies have also looked at aggregation of information in financial markets [Plott and Sunder (1982), (1988) and Sunder (1995) for a survey], they focused on *market* outcomes and not on individual behavior. Other important studies have suggested rich descriptions of how *individuals* learn from each others' actions [see Bikhchandani, Hirshleifer and Welch (1998) for a survey] but used settings that are somewhat different from those found in asset markets. Our design is novel in that it outlines a way of bridging these two strands of literature.

In the game there are two players, each receiving (1) a private signal and (2) a private *signal-precision*. Players' task is to guess an unknown fundamental value, around which their signals are drawn. The game consists of multiple decision turns in which players first observe each others' *previously* submitted estimates and then simultaneously submit new ones. At the end of the game, one of the previous turns is chosen randomly and each subject is paid according to the accuracy of her estimate on that turn relative to the drawn value. Players' payoffs do not come from trading but rather depend on the accuracy of their *individual* estimates.[2] Thus trading intensity, which is related to confidence in valuation, is replaced by persistency: the more a player is confident in her private information, the less she adjusts her estimates across turns. Over time (under full-rationality), players are predicted

[1] We do not claim that these results would come out of all models of overconfidence; rather, we are trying to provide an intuition for how overconfidence *may* affect markets.

[2] Unlike most trading mechanisms, the one described here is not a fixed sum game. This allows us to simplify the inference problem by inducing subjects to act as if they were risk-neutral price takers.

to perfectly aggregate their private information, converging to the Rational Expectations Equilibrium level.

In this game, natural proxies for prices, returns and volume emerge. The idea that prices reflect a weighted average of traders' individual beliefs about the fundamental value is very common in information based asset pricing models [e.g. Diamond and Verrecchia (1981)]. These weights generally depend on the distribution of wealth and/or preferences in the economy. Since in this game we control for endowments and modulate the effect of risk attitudes, we define a price index as the (equally weighted) average of players' estimates. Also, speculative volume is generally a result of traders' differing valuation of the underlying asset [e.g., Wang (1998)]. We therefore create a volume index which is equal to the absolute difference of players' estimates (recall that the game induces players to submit estimates which are equal to their valuations). Using these definitions, we construct a return index.

We conducted experimental sessions in which subjects participated in this game and were rewarded in cash based on their decisions.[3] Private signal precision (high or low) was determined by the subjects' rank on a task that took place at the beginning of sessions. In some sessions, denoted as *baseline treatment (BLT)*, participants rolled a die (whose outcome was *privately* observed). In other sessions, denoted as *overconfidence treatment (OCT)*, participants answered a short SAT quiz [see Camerer and Lovallo (1999)].[4] While the die throw is a neutral treatment, the SAT is not; many previous studies document the tendency of individuals to perceive themselves as 'better than average' in variety of contexts [e.g., Svenson (1981)]. Therefore, subjects who mistakenly believe they are better than their median peer on the SAT quiz will also mistakenly believe their signal precision is better than it really is and are therefore going to be overconfident about their private signal – not by conjecture, but rather *by construction*.[5]

Analyzing the results from the OCT we find that volume and price error indexes are in excess of what is predicted by the fully-rational model, lending support to the hypothesized effect of overconfidence on markets. However, we find qualitatively similar results in the BLT, where subjects have no reason to be overconfident.[6] Specifically, we find that a substantial fraction of *excess* volume index is attributable to strategic response to errors, while the remaining is attributable to overconfidence. If one looks at the price error index, similar results are obtained.

To formally separate these competing channels we form a structural model, denoted as Noisy Actions Biased Beliefs ('NABB'), that maps exogenous information, endogenous information and beliefs into actions while allowing for *both* erroneous beliefs and erroneous actions. Applying it to the data enables us to back out participants' subjective confidence in their information, calculate best-responses and estimate the magnitude of errors to which they respond.[7] Fitting the model on both treatments confirms that subjects are on average

---

[3] Subjects were primarily UC Berkeley business and economics undergraduate students, earning \$5-\$15, depending on individual performance, for a one-hour long experiment.

[4] In both treatments subjects were not told what their rank was but were made aware of the way it was determined.

[5] Most other experimental studies used survey-based miscalibration results and relate them to trading activity either in experimental markets [see Biais et al (2002)] or naturally occurring markets [see Glaser and Weber (2003)]. There are two potential problems with this approach: (1) these surveys do not provide incentives for accuracy and (2) miscalibration tends to be domain specific; for example, the level of overconfidence tends to depend on the difficulty of the task [Fisschhoff et al. (1977), Lichtenstein et al. (1982)] and on the domain-expertise [Keren (1987)].

[6] As we show later on, subjects do not exhibit aggregate overconfidence in the BLT.

[7] This is in the spirit of Quantal Response Equilibrium (QRE) models previously studied in the context

overconfident in the OCT (around 15%) but not in the BLT and that subjects seem to be responding to mean-zero errors made by others.[8]

The remainder of the paper is organized as follows: section 2 sets up the theoretical model and derives the unique subgame-perfect Nash equilibrium and section 3 describes the experimental design mirroring this model. Section 4 discusses the model-independent empirical results. We proceed to specify a richer model of behavior in section 5 and discuss its estimation results in section 6. We summarize in section 7.

2 Literature Review

There are a number of voluminous strands of literature related to this paper which we do not cover in this short review including cognitive psychology studies of overconfidence and miscalibration, theoretical asset pricing models studying the effect of traders' overconfidence, and empirical and theoretical work on social learning. Rather, we focus on the most closely related experimental work on overconfidence and emphasize key differences between these studies and our paper.

Kirchler and Maciejovsky (2002) measure individuals' miscalibration and how it is affected by trading. They allow subjects to trade securities that pay stochastic dividends over a number of periods in double-auction markets. At the beginning of each period, subjects are asked to provide an assessment of the distribution of trading prices they expect to observe. The authors use these predictions to construct two different measures of overconfidence. They find substantial heterogeneity in subjects' levels of over/underconfidence but find no aggregate overconfidence.

Biais et. al. (2002) correlate individual measures of overconfidence and self-monitoring, collected through surveys at the beginning of the sessions, to their earnings from trading in an asset market similar to Plott and Sunder (1988). Absent private values, this setting is subject to winner's curse since participants hold imperfect private information. Thus, no-trade results apply. The authors find that subjects prone to overconfidence earn relatively low earnings and those that exhibit high self-monitoring abilities earn relatively high earnings. At the same time, they do not find that overconfidence leads to more intense trading.

Deaves, Luders and Luo (2003) study the link between miscalibration, gender and trading intensity. The authors design an experiment in which subjects' signal quality depends on the accuracy of their responses on a survey. They find that overconfidence leads to increased trading activity among subjects. In contrast with other studies, they find no difference in overconfidence and trading intensity between men and women.

Glaser and Weber (2003) conducts a survey among broker investors to assess their overconfidence as expressed in their miscalibration and better than average effect (as well as illusion of control / unrealistic optimism). Results from 215 individuals were then matched with their own trading volume. The authors find that the two measure yield distinctively different results: above average effect, but not miscalibration, is related to trading volume.

There are a number of important differences between these papers and our study. First, we measure participants' individual over/underconfidence that is *implicit* in their decisions

of normal and extensive form games [see McKelvey and Palfrey (1995), (1998)]. Notice that these games involved discrete actions while here we deal with continuous actions.

[8] Average overconfidence of 15% means that subjects perceive the probability of being perfectly informed to be on average 65% while in fact it is 50% (since half of the subjects are perfectly informed by design).

by estimating a structural model of behavior. That is, instead of using direct elicitation, we estimate revealed miscalibration as relevant for the context studied. This is important as overconfidence lacks a universal operational measure and may very well differ for the same individual across different domains. As a result, defining a measure of overconfidence elicited through a survey that covers a particular area may not necessarily capture overconfident behavior in a market setting (see Glaser and Weber (2003)). Related to that, our approach is more robust to survey methodology; it is quite possible that while individuals are not able to *communicate* probabilistic assessments well, they are able to incorporate them into their decisions.

Second, the market mechanism we utilize differs substantially from the canonical double-auction markets, utilized by virtually all experimental asset markets. The tractability of our mechanism allows us to study a rich set of aggregate measures, such as price informativeness, and to generate clear predictions under the fully-rational model about the levels and changes in price and volume indexes. In contrast, most previous work almost exclusively dealt with *comparative static* tests of trading intensity or volume only. No study that we know of looks for example at the effect of overconfidence on the quality of prices, while it is clearly of central interest to economists.

Third, most previous work in this area followed the approach of correlating individuals' psychological attributes and their behavior in markets. As such, it was centered on *individual level results.* Our focus is on understanding the aggregation process and the market level effects of individuals' biases, while also discussing a number of individual level findings.

3 Theory

3.1 General

In this game there are two players, both trying to estimate the realization of a random variable v, referred to as 'fundamental value', where $v \sim U[L, H]$. Each player is assigned a type: $t_i \in \{h, l\}$ such that one player is of type h and the other is of the complementary type, l. A player of type h receives a perfect signal while player of type l receives an imperfect signal:[9]

- Perfect signal: $s_i^h = v$
- Imperfect signal: $s_i^l = v + e_i$, where $e_i \overset{iid}{\sim} U[-Y, Y]$

Subjects do not know whether their type is h or l; instead, they observe a draw, q_i, representing the objective probability that they are of type h, where q_i is drawn IID from a known continuous distribution with a support $F \subseteq [0, 1]$. Since there are only two types of signals, q_i fully characterizes the precision of player i's private signal. For now, assume that subjects' beliefs about the likelihood that her own signal is perfect, denoted by $\widetilde{q}_i$, are correct (i.e., $\widetilde{q}_i = q_i$). Thus, the realization $\{s_i, q_i\}$ makes up subject i's private information.

At $t = 0$, the realizations of v and $\{s_i, q_i\}_{i=1}^2$ are drawn. The collection of $\{q_i\}_{i=1}^2$ is used to determine players' types: the player with the highest draw of q_i is assigned type h, while the other player is assigned type l.[10] The game consists of 3 turns: at the beginning of each turn, t, both players simultaneously submit an action, $a_{i,t}$, which comes in the form

[9] Subscripts $i.j$ index the players.

[10] Note that type realization is not part of players' information set.

of a numerical estimate of the realized fundamental value.[11] At the end of each turn, both players' estimates are announced.[12] As we show later, 3 turns are needed for players to arrive at the fully-revealing equilibrium. The intuition is straightforward. There are two dimensions of uncertainty for each player – other's signal and signal precision. Since each turn can allow for at most one new dimension to be observed, players need to observe each others' estimates for two turns, arriving at full-revelation in turn 3.

At the end of the game one turn is randomly chosen (with equal probability) and players receive a payoff $\pi_i(a_{i,t}, v)$ ensuring that expected utility is maximized at the expected value of v: $E(v|I_{i,t}) \in \arg\max E[u_i(\pi_i(a_{i,t}, v))|I_{i,t}]$, where $I_{i,t}$ represents player i's information set (both private and public) in turn t. Put differently, payoff scheme ensures that if players act myopically, they minimize the forecasting error at each turn of the game. For example, if players are risk neutral then $\pi_i = -(v - a_i)^2$; if players have log utility then $\pi_i = \exp(-(v - a_i)^2)$, etc. Note that since each player is paid according to the accuracy of her actions, irrespective of the actions of the other player, this is not a fixed sum game (unlike most trading institutions). This feature is important in neutralizing payoff externalities typically arising in market settings and removing strategic incentives.

To understand the dynamics of this game, notice that players' confidence in their information is not conveyed through trading intensity. Rather, it is communicated through the extent to which they revise their estimates. A player revising her estimate sharply, in response to observing a fellow player's estimate, is reflecting low confidence in her previously held information.

3.1.1 Optimal Actions

Now we turn to characterize the fully-rational solution of this game by calculating the optimal actions of players i, j (denoted by a_i^*, a_j^*). Recall that the game starts with subjects receiving their private information, $\{s_i, q_i\}$, followed by three decision turns. Since the exogenous information is fixed across the turns, subjects revise their submissions due to endogenous information only, obtained by observing others' actions. Also, since exactly one player is perfectly informed but the identity of that player is uncertain, optimal actions are a convex combination of players' signals. How far one's estimate is from her signal depends generally on the confidence she has in her signal.

Proposition 1 *There exist a Perfect Bayesian Equilibrium (PBE) where players optimal actions are:* $a_{i,t}^* = E(v|I_{i,t})\ \forall i, t$.

Proof In turn 1, optimal actions are:[13]

$$a_{i,1}^* = s_i$$
$$a_{j,1}^* = s_j$$

[11] We fixed the number of turns to be 3 because as we show later, this is the number of turns needed for full information revelation. That is, any additional turns are redundant.

[12] Each player is privy to the actions of the other player with whom they were paired in that round and not all players. Also, since there are 2 players per market, observing the average of actions is sufficient statistic for the action of the other player.

[13] Since we set the support from which v is drawn to be much larger than Y, we ignore the boundary cases where $s_i, s_j \in [H - Y, H] \cup [L, L + Y]$.

At the end of turn 1, $a_{i,1}$ and $a_{j,1}$ are announced.

In turn 2, $a^*_{i,2} = E(v|I_{i,2}) = E(v|\{s_i, q_i, a^*_{j,1}\}) = E(v|\{s_i, q_i, s_j\})$

Since player i cannot extract any information about the other players' realized signal precision, q_j, we obtain that:

$$a^*_{i,2} = q_i s_i + (1 - q_i)s_j \tag{1}$$
$$a^*_{j,2} = (1 - q_j)s_i + q_j s_j \tag{2}$$

Once again, at the end of turn 2, $a_{i,2}$ and $a_{j,2}$ are announced.

In turn 3, since both s_j and q_j are known[14], $a^*_{i,3} = E(v|\{s_i, q_i, a^*_{j,1}, a^*_{j,2}\}) = E(v|\{s_i, q_i, s_j, q_j\}) = a^*_{j,3}$

$$a^*_{i,3} = a^*_{j,3} = Ind_{(q_i>q_j)}s_i + Ind_{(q_i<q_j)}s_j + Ind_{(q_i=q_j)}(\frac{s_i + s_j}{2}) \tag{3}$$

Where Ind represents the indicator function. Since all information is now common knowledge full information revelation is obtained.

Proposition 2 *The Perfect Bayesian Equilibrium (PBE) characterized above is unique.*

Proof. We will use backward induction for this proof:
Since the myopic best-response equilibrium maximized expected payoffs at each turn of the game separately, player would deviate from it only if they can increase their future expected payoffs. Therefore, In the last stage of the game, both players follow $a^*_{i,3} = E(v|I_{i,3})$ since no future benefits can arise from deviation.
In turn 2, assume $a_{i,2} \neq a^*_{i,2} \Rightarrow E(u_{i,2}(a_{i,2})) < E(u_{i,2}(a^*_{i,2}))$ so it must be the case that $E(u_{i,3}(a^*_{i,3}(I(a_{j,2}(a_{i,2}))))) > E(u_{i,3}(a^*_{i,3}(I(a^*_{j,2}(a_{i,2})))))$ but since actions are submitted simultaneously, this can not hold. Thus, in turn 2, both players' actions are $a^*_{i,2} = E(v|I_{i,2})$
In turn 1, assume that $a_{i,1} \neq a^*_{i,1} \Rightarrow E(u_{i,2}(a_{i,1})) < E(u_{i,2}(a^*_{i,1}))$ so it must be the case that $E(u_{i,3}(a^*_{i,3}(I(a_{j,2}(a_{i,1}))))) > E(u_{i,3}(a^*_{i,3}(I(a_{j,2}(a^*_{i,1})))))$ but since turn 3 actions arrive at full information revelation (a.s.), this can not hold. ■

Before proceeding, a few features of this game should be emphasized. First, information is aggregated sequentially. In turn one, optimal action depends on ones' own signal. In turn two, optimal action depends on own signal as well as other's observed turn one action and own subjective confidence. Second, using the mapping outlined in the introduction between this game and financial markets, we denote level of disagreement by volume index ($Vol = |a_{i,t} - a_{j,t}|$), average estimate as price index ($P = \frac{a_{i,t}+a_{j,t}}{2}$) and the distance between price index and v as price error index ($PE = |v - P_t|$). Using these definitions, we claim that volume index and price error index strictly decrease across turns. This is an intuitive result – the effect of gradual information aggregation is that the average estimate participants hold gets closer to the underlying value, and the effect of gradual information dissemination is that participants' estimates get closer together.

Definition 3 *volume index:* $Vol_t = |a_{i,t} - a_{j,t}|$

Proposition 4 *volume index strictly decreases from turn* 1 *to* 3 *a.s.*
Proof. $Vol_1 = |s_i - s_j| = |e_j| > 0$ *a.s.*

[14] Since $a^*_{j,1} = s_j$ and $a^*_{j,2} = (1 - q_j)s_i + q_j s_j$ we obtain that $q_j = \frac{a^*_{j,2} - s_i}{a^*_{j,1} - s_i}$

$Vol_2 = |q_i s_i + (1-q_i)s_j - (1-q_j)s_i - q_j s_j| = |(1-q_i-q_j)s_j - (1-q_i-q_j)s_i|$
$= |(1-q_i-q_j)(s_j - s_i)|$

Since $-1 \leq (1-q_i-q_j) \leq 1$ and $-e_j \leq (s_j - s_i) \leq e_j$, we get that $|(1-q_i-q_j)e_j| < |e_j|$ a.s.(notice that $q_i + q_j$ need not equal 1)

Also, since $a^*_{i,3} = a^*_{j,3}$, $Vol_3 = 0$

Thus, $Vol_1 > Vol_2 > Vol_3 = 0$ ■

Notice that volume index (in round 2) is increasing in (1) the realized signal error of the imperfectly informed trader and (2) the *sum* of subjective beliefs.

Definition 5 *Price index: $P_t = \frac{a_{i,t}+a_{j,t}}{2}$*

Definition 6 *Price error index: $PE_t = |v - P_t|$*

Proposition 7 *The price error index strictly decreases a.s. from turn 1 to 3*

Proof. $PE_1 = \left|\frac{s_i+s_j}{2} - v\right| = \left|\frac{v+v+e}{2} - v\right| = \left|\frac{e_j}{2}\right| > 0$ *a.s.*

$PE_2 = \left|\frac{q_i s_i+(1-q_i)s_j+(1-q_j)s_i+q_j s_j}{2} - v\right| = \left|\frac{(1-q_j+q_i)s_i+(1-q_i+q_j)s_j}{2} - v\right|$

Assume WLOG that player i is of type h. Then

$\left|\frac{(1-q_j+q_i)s_i+(1-q_i+q_j)s_j}{2} - v\right| = \left|\frac{(1-q_j+q_i)v+(1-q_i+q_j)(v+e_j)}{2} - v\right| = \left|\frac{(1-q_j+q_i+1-q_i+q_j)v+(1-q_i+q_j)e_j}{2} - v\right| = \left|\frac{2v+(1-q_i+q_j)e_j}{2} - v\right| = \left|\frac{(1-q_i+q_j)e_j}{2}\right|$

Since $q_i > q_j$, $0 < (1-q_i+q_j) < 1$ we get that $PE_2 = \left|\frac{(1-q_i+q_j)e_j}{2}\right| < \left|\frac{e_j}{2}\right| = PE_1$

Recall that since by turn 3, the price index is perfectly revealing (a.s.) and since one of the players is perfectly informed, $PE_3 = 0$. Thus, $PE_1 > PE_2 > PE_3 = 0$. ■

Thus, price error index increases (in turn 2) in realized signal error and in *difference* in subjective beliefs.

3.1.2 Miscalibration

Recall that we are interested in understanding the effects of two forms of bounded rationality: errors in actions and errors in beliefs. In this section, we provide some intuition for the latter. We will discuss the former in the context of our econometric model.

Consider the possibility that players hold erroneous beliefs about their probability of being perfectly informed. That is, individual subjective probability equals the objective probability plus miscalibration: $\widetilde{q}_i = q_i + MC_i$ where MC_i denotes subject i's miscalibration. Positive miscalibration represents overconfidence while negative miscalibration represents underconfidence. We allow for arbitrary subjective beliefs as long as they are admissible, that is $0 \leq \widetilde{q}_i \leq 1$ $\forall i$.[15]

To simplify matters, let us assume that subjects are naive in the sense that they are not aware of other players' potential miscalibration (this assumption will be maintained throughout this paper).[16] Now, we can rewrite volume and price error indexes in the case of miscalibration, denoting them with superscript MC.

Volume index**:**

[15] This definition corresponds to the way miscalibration is defined in the cognitive psychology literature [see Alpert and Raiffa (1977) for a review].

[16] We naturally assume that players are not aware of their own miscalibration [see for example Odean (1998)].

- Turn 1: $Vol_1^{MC} = |s_i - s_j| = |e| > 0 \ a.s$
- Turn 2: $Vol_2^{MC} = |1 - \widetilde{q}_i - \widetilde{q}_j|\,|e| = |1 - (q_i + q_j) - (MC_i + MC_j)|\,|e|$
- Turn 3: $Vol_3^{MC} = 0$

Notice that turn 3 volume index would be zero, even if players are miscalibrated, since both parties regard subjective beliefs to be equal to the objective beliefs and thus converge on the signal held by the player with the larger subjective probability of the two. Thus, unlike the fully-rational case, convergence will happen but it may be to the wrong signal.

Price error index**:**

- Turn 1: $PE_1^{MC} = \left|\frac{e}{2}\right|$
- Turn 2: $PE_2^{MC} = |1 - \widetilde{q}_i + \widetilde{q}_j|\left|\frac{e}{2}\right| = |1 - (q_i - q_j) - (MC_i - MC_j)|\left|\frac{e}{2}\right|$
- Turn 3: $PE_3^{MC} = \left|\frac{2\left(1_{\widetilde{s}_i > \widetilde{s}_j}\widetilde{s}_i + 1_{\widetilde{s}_j > \widetilde{s}_i}\widetilde{s}_j\right) + 1_{\widetilde{s}_j = \widetilde{s}_j}(\widetilde{q}_i s_i + (1-\widetilde{q}_i)s_j + (1-\widetilde{q}_j)s_i + \widetilde{q}_j s_j)}{2} - v\right| =$
  $\left|1_{\widetilde{s}_i > \widetilde{s}_j}\widetilde{s}_i + 1_{\widetilde{s}_j > \widetilde{s}_i}\widetilde{s}_j + \frac{1_{\widetilde{s}_j = \widetilde{s}_j}(\widetilde{q}_i s_i + (1-\widetilde{q}_i)s_j + (1-\widetilde{q}_j)s_i + \widetilde{q}_j s_j)}{2} - v\right|$

In the discussion below we will compare volume and price error indexes in the absence of miscalibration and in the presence of miscalibration to obtain comparative static predictions.

Proposition 8 *The expected volume index is (weakly) greater in the presence of average overconfidence for all turns (a.s.).*

Proof. For turns 1 and 3 the proof is trivial.

For turn 2, recall that $Vol_2^{OC} = |1 - (q_i + q_j) - (MC_i + MC_j)|\,|e| \propto |1 - (q_i + q_j) - (MC_i + MC_j)|$, since $|e| \geq 0$.

Denoting $(q_i + q_j) \equiv q_{ij}$ and $(MC_i + MC_j) \equiv MC_{ij}$, and squaring both sides of the expression we get:

$\left(Vol_2^{OC}\right)^2 = (1 - q_{ij} - MC_{ij})^2(s_j - s_i)^2 = (1 - q_{ij} - MC_{ij})^2(e_j)^2$

To find the parameter value ranges for which index volume is increasing, take a derivative with respect to MC_{ij}:

$\frac{d\left(Vol_2^{OC}\right)^2}{dMC_{ij}} = -2(e_j)^2(1 - q_{ij} - MC_{ij})$, which is increasing if

$q_{ij} + MC_{ij} - 1 > 0$. Since in expectations $q_{ij} = 1$, $E(Vol_2^{OC})$ is increasing in MC_{ij} if players are on average overconfident. ■

Proposition 9 *The price error index is (weakly) increasing in the dispersion of overconfidence for all turns.*

Proof For turn 1, the proof is obvious.

For turn 2, recall that $PE_2^{MC} = |1 - (q_i - q_j) - (MC_i - MC_j)|\left|\frac{e}{2}\right| \propto |1 - (q_i - q_j) - (MC_i - MC_j)|$

WLOG we have assumed that $q_i \geq q_j$ and since $|q_i - q_j| \leq 1$ we have that $1-(q_i-q_j) > 0$. Therefore:

- if the better informed player, i, is *less* overconfident than the worst informed player, j, price error index is *higher* than when players are well-calibrated.

- if the better informed player, i, is *more* overconfident than the worst informed player, j, price error index is *lower* than when players are well-calibrated.

In turn 3, price error index would a.s. be either equal to zero or to $|e|$, exceeding the price error index in the case when players are well-calibrated.

Compared with the case where subjects are well calibrated, price error index is greater if the worst informed is more overconfident than the better informed and price error index is smaller if the better informed is more overconfident than the worst informed.

4 Experimental Design

4.1 General

The experiment was run at the Haas School of Business: a total of 12 sessions were conducted in which 72 subjects participated; 5 were Baseline Treatment (BLT) and 7 were Overconfidence Treatment (OCT).[17] Subjects were recruited from undergraduate classes at the University of California, Berkeley and had no previous experience with similar experiments. They received a show-up payment of $5 and an additional performance-based pay of $0-$10, which was paid in private and in cash at the end of the session. Sessions were 1 hour long and included 6 participants each.

At the beginning of each session an administrator read the instructions aloud and answered questions in private.[18] Each subject entered their decision using a computerized interface, which was built for the purpose of this experiment [see figure 1], thus maintaining both isolation and anonymity. Particular emphasis was put on limiting interaction to that facilitated by the computerized system.[19]

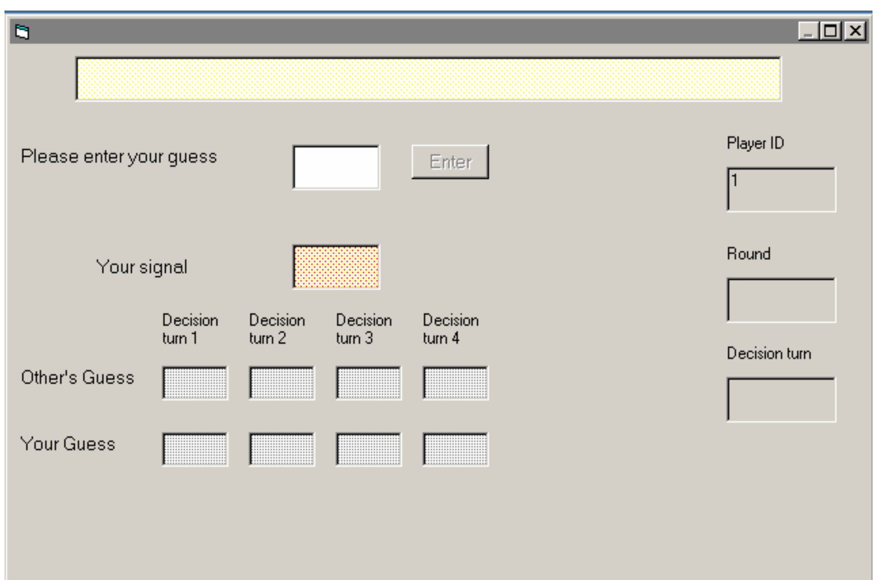


Figure 1: Interface: screen shot

[17] The order of treatments was determined randomly.
[18] Instructions are available upon request.
[19] The application developed by the author for this experiment is available upon request.

4.2 Structure

Each session started with an initial phase, followed by 10 independent and identical rounds.[20] At the beginning of each round subjects were randomly assigned into markets consisting of two players each and were presented with their private signal. Each round was composed of 4 decision turns and in each subjects were asked to enter their decision.[21] Throughout the turns, subjects' pairing and their private information remained the same. Transition from one turn to the next occurred only after all subjects submitted their action and no time restriction was imposed.

The experiment was carried out along a single treatment: base-line treatment (BLT) or overconfidence treatment (OCT), which differed *only* in their initial phase. In the BLT, the initial phase consisted of subjects privately throwing a die and observing its outcome.[22] Draws were recorded by the experiment administrator and fed into the computer which then determined the rank of the draws; three of the subjects, with the highest draws, were classified as perfectly informed while the other three, with the lowest draws were classified as imperfectly informed (ties were resolved randomly). Subjects observed their own draw but did not observe the draws obtained by other participants and were not told their rank.

In the OCT, subjects were asked to answer 20 multiple-choice SAT questions (taken from sample tests that were posted on the CollegeBoard website, see appendix). Scores were recorded by the computer which then ranked subjects according to the number of correct answers, as a primary key, and by the length of time required to complete the quiz, as a secondary key; three of the subjects, ranked top, were classified as perfectly informed while the other three were classified as imperfectly informed. Again, subjects were not told what their rank was.

The choice of using SAT questions was deliberate and intended to bias the results in favor of the null, stating that treatment would have no effect on market outcomes, by facing subjects with a task with which they are familiar - one that they have performed before and on which their ranking, with respect to the relevant peer group, is known.[23]

4.3 Information

The information structure was the following: at the beginning of each round a quantity v was drawn by the computer, where $v \sim U[50, 950]$. Then, subject i received an independent signal $s_i = v + e_i$ such that $e_i = 0$ for subjects that were classified as perfectly informed and $e_i \sim U[-30, 30]$ for subjects that were classified as imperfectly informed.

All information was continuously displayed on subjects' interfaces for them to observe. Note that aside from the information specified above, no additional feedback was given. In particular, the realization of the unknown quantity, v, was not revealed at any stage of the experiment (not even at the end of the round) and subjects did not see their earnings until the end of the *session*. This may be likened to an environment where traders never get to observe the liquidating value; subjects can only learn from their interaction with other

[20] We report here the results from the first 10 rounds while a few sessions were conducted with more rounds.

[21] While the 4th turn is redundant under the fully rational model it need not be redundant in practice. We also run 2 sessions (not reported here) with 6 decision turns but behavior during the last two turns seemed very close to the one exhibited in turn 4.

[22] At the beginning of each experiment, one subject was publicly asked to examine the die and confirm that it appeared normal.

[23] Note that SAT scores already reflect ranking as they are curved.

players, which is an endogenously generated information, not from exogenous cues. [24]

4.4 Assignment

Pairing into markets is randomly determined while ensuring that exactly one subject is perfectly informed and the other one is imperfectly informed. This (1) makes ex-ante distribution of information equal across all market instances (2) disables subjects from easily unveiling their type and (3) allows posterior probability updating to take on a particularly simple and intuitive form.[25,26]

4.5 Actions and payoff

At the beginning of each turn t subjects simultaneously submit their estimates $a_{i,t}$ by entering a number on their screen. No restrictions are imposed on the value the report can take. Upon receiving submissions from both subjects, the turn comes to an end and no changes are accepted. At that point subjects are informed of each others' estimate and are given a short transition time into the next turn.

At the end of the session, one turn from each round is randomly drawn and earnings (for subject i in round r) are calculated as follows:

$$\pi_i = \sum_r c_1 * \exp\left(-\frac{(v_r - a_{i,r})^2}{c_2}\right) \tag{4}$$

Where we parametrized $c_1 = 100, c_2 = 50$. At the end of the experiment, the total number of points earned was converted into dollars using an exchange rate of 100 to 1 and subjects were paid in private and in cash. Average earnings were \$12 with standard deviation of \$3.5.

This payoff function was chosen for a number of reasons. First, its convexity ensures that payoffs are non-negative everywhere. This is desirable because of the bankruptcy possibility arising from subjects submitting estimates that are distant from the fundamental value (due to errors).[27] Generally, bankruptcy is nonenforceable in the lab and once encountered may influence subjects' decisions in a substantial manner and may result in loss of experimental control [see Friedman and Sunder (1994)]. Second, the symmetry of payoffs around the fundamental value suggest to subjects that they should submit estimates that minimize estimation error. Indeed, the instructions reinforce this idea by stating that "the more precise your guesses are the more money you will earn at the end of the experiment" (see appendix). While formally this payoff function induces truth-telling if players maximized expected log utility, we do not find evidence to suggest that subjects' risk attitudes, deviating systematically from log utility, influence our results.[28]

[24] In a few sessions, we have extended the number of rounds to include full feedback round: subjects' payoff and the realization of v was revealed at the end of the round. Sure enough, subjects discovered whether they were the perfectly or imperfectly informed type almost immediately.

[25] If two subjects submit the same estimate in turn 1, most chances are that they are both perfectly informed and thus from the next round on both players know their type with certainty.

[26] We have conducted a few sessions (not reported in this paper) with different rules of market assignment. The problem discussed here does appear: when two perfectly informed subjects are paired together, they tend to find out their type. Nonetheless, the qualitative features of the experiment and the results are similar.

[27] In a number of sessions (not reported here) we have used a quadratic payoff function. In each of those sessions, about a third of the participants ended up with negative payoffs after the first few rounds.

[28] As an aside, assuming that subjets are risk-averse, rather than risk neutral, is desirable, as suggested

5 Model independent results

Our main findings can be divided into

- individual level
  - subjects to incorporate both private information dimensions into their estimates: signal realization and signal precision
  - subjects seem to exhibit overconfidence and/or react strategically to others' errors
- aggregate level
  - information is aggregated and disseminated under both treatments but not with the same degree of efficiency
  - the volume and price error index levels are in excess of the fully rational model prediction in *both* treatments; however, these indexes are higher in the OCT compared with the BLT
  - the price index exhibits negative serial autocorrelation in the OCT but *not* in the BLT

5.1 Individual level

Given the central role that the second moment of information plays in this game, we seek to characterize individual level behavior by focusing on a measure that captures subjects' weighting of their own information with that of their fellow player. Recall that confidence in this game is expressed by the rate at which estimates are adjusted across turns. That is measured by "adjustment rate", which we define as the change in estimate, from turn 1 to turn 2, divided by the difference between players' turn 1 estimates. This quantity is represented in figure 2 as the fraction B/A.

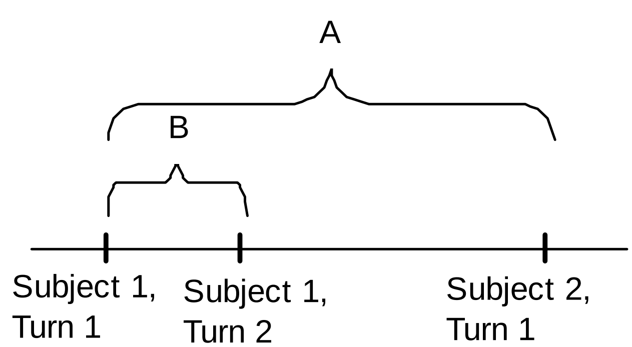


Figure 2: Adjustment rate

To better interpret this measure, recall that the *incremental* information obtainable at each turn is the following:

by experimental studies of first-price sealed bid auctions [see Kagel and Roth (1995) and Davis and Holt (1993) for a review].

- Turn 1: own signal realization
- Turn 2: other's signal realization and own signal precision
- Turn 3: other's signal precision
- Turn 4: none

As was shown in the theory section (in the absence of errors) subject i's turn 2 optimal estimate is $a_{i,2} = \widetilde{q}_i s_i + (1+\widetilde{q}_i) a_{j,1}$, which can be rearranged as $\widetilde{q}_i = \frac{a_{i,2}-a_{j,1}}{s_i - a_{j,1}}$. Since $s_i = a_{i,1}$, we obtain that $\widetilde{q}_i = \frac{a_{i,2}-a_{j,1}}{a_{i,1}-a_{j,1}}$. This quantity matches the definition of the adjustment rate. We show in a later section that in the case where subjects react to errors in others' estimates, $a_{i,2} = \widetilde{q}_i s_i + (1+\widetilde{q}_i)(a_{j,1} + c_{i,2})$ where $c_{i,2} \leq 0$ represents the discount player i applies to the turn 1 estimate of player j. In that case, $\widetilde{q}_i = \frac{a_{i,2}-a_{j,1}}{a_{i,1}-a_{j,1}-c_{i,2}}$.

Thus, the adjustment rate is a unit-free measure of the weight assigned to the private signal realization and as such depends on:

- Subjective probability of being perfectly informed, which in turn can be broken into
  - objective probability of being perfectly informed
  - miscalibration (over/under confidence)
- Reaction to possible errors made by other players

Figure 4 depicts observed ("Data") and fully-rational, no-overconfidence theoretical ("RE") adjustment rates for the BLT and the OCT, sorted by subjects' *objective* probability of being perfectly informed.[29] The main findings suggest that subjects act in a coherent manner that is also consistent with our intuition of the game:

- Objective confidence is monotonically related to adjustment rates *in both treatments.* Higher objective probability corresponds to lower adjustment rates. This suggests that subjects incorporate their signal precision into their estimates and that objective and subjective probabilities are related.
- In the OCT, subjects adjust *less* than they do in the BLT, which is consistent the presence of more overconfidence in the former treatment.
- Subjects adjust *uniformly less* (in both treatments) than predicted by the RE model, suggesting that players discount others' actions. This is consistent with strategic response to others' errors.

5.2 Aggregate-level results

5.2.1 Measures

The empirical analysis discussed in this section focuses on the following market-level measures, each representing a dependent variable of interest. These are:

[29] For the purpose of this plot only we exclude observation where B and A do not have the same sign as these cases clearly represent erroneous behavior; by doing that we have taken out less than 5% of the observations, across both treatments.

- **Volume index** represents the extent to which players diverge in their estimates and is defined as $Vol_{r,t} = |a_{i,r,t} - a_{j,r,t}|$.[30]
- **Price index** represents the average estimates and is defined as $P_{r,t} = \frac{a_{i,r,t}+a_{j,r,t}}{2}$.
- **Price error index** represents the distance between the average estimate and the fundamental value and is defined as $PE_{r,t} = |v_r - P_{r,t}|$.

5.2.2 Aggregation and dissemination of information

The markets we study here have the potential to aggregate and disseminate private information held by individual players. The extent to which they succeed in performing these functions can be measured through the level and change in volume and price error indexes. The volume index is indicative of the degree to which information held by players is close together, thus proxying for the information disseminated. The price error index is indicative of the degree to which the aggregate information-set is informative, thus proxying for information aggregation. Thus, we look at the changes and levels of volume and price error indexes across turns.

Table 1 provides average levels (across market instances and rounds) of volume and price indexed for both treatments while table 2 summarizes the non parametric Wilcoxon rank-sum test results of the null that median indexes are constant across turns.[31] In the BLT both volume and price error indexes decrease from turn 1 to turn 3 but not after that, in line with the rational model (recall that turn 4 is in principle redundant). Quite remarkably, median price error and volume indexes in the last turn are very close to zero (0.5 and 1.5, respectively). In the OCT, we find similar pattern, delayed by a turn: there is a significant drop going from turn 1 to 2 and from 3 to 4 but not in between. Notice that while the *volume index* decreases toward the end of the round, the *price error index* does not. That is, subjects seem to converge but to the *wrong* value, which can be explained by subjects' ignorance of their and others' overconfidence. We provide further evidence on that in our discussion of return autocorrelation.

5.2.3 Excess volume and price error indexes

A central question of interest is: are excess volume and price error indexes linked to overconfidence? To answer that, we compare the OCT results to the predicted RE levels. We find that consistent with the predictions of many asset pricing overconfidence models, observed levels are higher than predicted by the fully-rational model (see table 1).[32] To gauge how much of these deviations from the rational model prediction are due to overconfidence, we perform the same comparison on the BLT results, in which subjects are not induced to overconfidence. First, we find *qualitatively* similar patterns to those observed in the OCT, while the magnitude is lower (see tables 3 for a formal cross-treatment significance test). A

[30] The subscript notation consists of player identification, round number and turn number, in that order. We may drop the round or turn subscript when appropriate.

[31] To obtain a feel for the results recall that in each market instance there is exactly one subject who receives an imperfect signal. This signal is uniformly distributed around the liquidating value with bounds of +/-30. Therefore, if these markets did not aggregate or disseminate information at all, expected volume index would have been 15 and expected price error index would have been 7.5.

[32] Notice that observed levels of volume and price-errors indexes are not significantly different from those predicted by the RE model in turn 1. This is to be expected since the precision of information should not affect turn 1 estimates.

large fraction of *excess* volume and price error indexes is attributable to strategic response to errors, while the remaining is attributable to overconfidence.[33] In a later section we show that a model, which allows for erroneous beliefs and actions, can replicate the BLT levels of excess volume and price errors indexes and establish that subjects are not overconfident.

5.2.4 Return autocorrelation

In contrast to the spirit of the findings above, we proceed to show that return autocorrelation – another phenomenon attributed to overconfidence – is found *only* in the OCT. To explore this, we sort within-round price index changes by size (-30 to -20, -20 to -10 etc.) and plot the average change from turn 1 to 4 relative to the average change that would have occurred at the close of the round had the fundamental value been announced (see figure 5). If returns are uncorrelated, these two should not be related. As we can see, in the BLT the average close-to-fundamental value is virtually zero irrespective of the sort. In the OCT, the price changes seem to exhibit reversals.

To provide an econometric test, we estimate (using a robust regression technique) the following relation:

$$Ln(V_i/P_{4,i}) = a + bLn(P_{4,i}/P_{1,i}) + e_i$$

Table 4 summarizes the results. First, there is no indication of unconditional return predictability in either of the treatments. Second, we find negative serial autocorrelation of returns in the OCT (at a very high significance level) but not in the BLT; i.e., we find price index reversals. Third, change in the volume index (within the round) helps predict returns; when the volume index *decreases* (i.e., subjects' valuation get closer to each other) price reversals are *more* pronounced. The intuition is the following: in the course of the game, average estimates move away from a naive average of signals and closer to the signal held by the *poorly* informed player. This is due, as we show later, to the heterogeneity of overconfidence; since players that are very unlikely to receive perfect signals tend to be very overconfident while players that are very likely to receive perfect signals tend to be somewhat underconfident. Each of these effects increases the market weight on the poorly informed player. To the extent that subjects are naive about the existence of overconfidence, they would fail to properly offset the "negative externality" brought about by the poorly informed yet overconfident players.

6 Econometric Model

To substantiate the effect of the two forms of deviations from rationality discussed here, we need to obtain subjects beliefs (which will allow us to estimate their individual miscalibration) and separate them from strategic response to errors. In the next section we describe a structural model that allows us to do just that. It maps exogenous information (signal realization), endogenous information (players' publicly submitted history of estimates) and beliefs into best-response actions. Since all the model's inputs are known except for subjects' beliefs, we fit the data to back-out subjective probabilities implicit in subjects' decisions at each stage of the session.

The econometric model we suggest, termed Noisy Actions Biased Beliefs ('NABB'), is designed to separate errors in action from errors in beliefs by nesting them. Recall that

[33] We do not include turn 1 results since they are not influenced by overconfidence.

these two channels of deviation from full rationality have potentially competing effects: too high subjective probability of being perfectly informed as well as strategic reaction to fellow players' errors both result in increase of the relative weight players' assign to their own private information. We use a probabilistic choice model [see Goeree and Holt (1999), McKelvey and Palfrey (1995, 1998)] in which players' actions are *distributed around* their best-responses; best responses are formed while taking into account that others' actions include errors. Additionally, we let each subject hold arbitrary beliefs about the precision of their private information.

We build on the following principles:

- Subjects' actions include errors
- The magnitude of errors is inversely related to their cost
- Subjects have some expectations about the distribution of errors; the distribution of errors is common knowledge and matches the observed distribution
- Subjects react strategically to errors of others when forming their best responses
- Subjects' confidence in the precision of their private signal need not match its true precision, i.e., they may be miscalibrated. However they are naive about the possibility that either they or others may be miscalibrated, i.e. we are fixing subjects' higher order beliefs about overconfidence

There are two notable differences between the model we propose and those discussed in the literature. First, we estimate *individual* beliefs and miscalibration without imposing any parametric assumptions about them and without using direct elicitation. Thus we end up with estimates of each participant's implicit confidence over the course of the session.[34] One reason we can achieve this is because the game involves repeated interaction – each unit of observation includes three pairs of simultaneous decisions (corresponding to turns 1-3) over which beliefs are constant. Second, we do not assume that participants are perfect econometricians; in particular, we allow for the possibility that they misestimate the variance of errors. Third, we allow actions to be continuous while most previous applications involved mostly discrete choice games [see Celen and Kariv (2003) for an example].

We formulate the problem recursively [see Anderson and Holt (1997)], so that actions in turn t best respond to the distribution of errors in turns $\{t-1, t-2, ...\}$. To allow for continuous actions, we suggest that observed actions are composed of best-response action (conditional on private and public information) and white noise-error term, the following specification:

$$a_{i,t} = a^*_{i,t}\left(I_{i,t}, \{c\sigma_j\}_{j=1}^{t-1}\right) + e_{i,t} \tag{5}$$

where $a^*_{i,t}$ is player i's best-response in turn t, $\{\sigma_j\}_{j=1}^{t-1}$ is set of the error disturbance parameters for previous turns (1 through $(t-1)$), $e_{i,t} \sim N(0, \sigma_t^2)$ is turn t realized error, and c represents player's beliefs about the size of σ.

[34] In contrast, previous studies took one of two approaches: they either used direct elicitation, thus ending up with *static* individual measure, or estimated the *distribution* of belief parameters, thus ending up with a collective measure.

This model assumes that subjects' observed actions are normally distributed around the optimal action in that turn.[35] Specifying a normal distribution for errors automatically satisfies the condition that the probability of observing deviations from best response is inversely related to their cost since the payoff function is of the same functional form as the normal density (both are negative quadratic exponential). Further, the distribution of errors is the same across players. In particular, subject i's error in a particular turn does not change the likelihood of observing a given size of error in the previous or subsequent turns.[36] Optimal actions are a function of all information (private and public) as well as knowledge of previous turns' error distributions.

6.1 Optimal Actions

In this part we derive expressions for best-responses and actions for a generic player i during turns 1-2.[37] We retain a key feature discussed in the context of the fully-rational model: in each turn, players' estimates correspond to their expected value of the underlying asset, given their private information.

In turn 1 (suppressing the round index and using subscript 1/2 for the first/second player), we obtain that:

$$a_{1,1} = s_1 + e_{1,1} \tag{6}$$

$$a_{2,1} = s_2 + e_{2,1} \tag{7}$$

where $e_{i,1} \sim N(0, \sigma_1^2)$.

In turn 2, $a^*_{1,2} = s_1 q_1 + E(s_2|a_{21}, s_1)(1 - q_1)$, but,

$E(s_2|a_{21}, s_1) = \int_{s_1-Y}^{s_1+Y} s_2 \Pr(s_2|a_{2,1}, s_1) ds_2$.

After some calculations, we obtain (see page 26 for details) that:

$$a_{1,2} = q_1 s_1 + (1 - q_1) \left(a_{2,1} + \frac{2c^2\sigma_1^2 \left(\phi(a_{2,1}; s_1 - Y, c\sigma_1) - \phi(a_{2,1}; s_1 + Y, c\sigma_1)\right)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2c\sigma_1}\left(Y - s_1 + a_{2,1}\right)\right) - \operatorname{erf}\left(\frac{\sqrt{2}}{2c\sigma_1}\left(a_{2,1} - s_1 - Y\right)\right)} \right) + e_{1,2} \tag{8}$$

[35] We impose mean zero error distribution for all turns.

[36] Some one-shot games assume that subjects come from a pool that includes some random players and other best-responding players. Using this specification in a context of a repeated interaction game, like the one studied here, would introduce further complexity [see for example Celen and Kariv (2003)].

[37] We focus on these turns only for tractability reasons (recall that in this setting the updating rule is not stationary). In a previous version of the paper, we derived the solution to optimal actions in turn 3. However, including turn 3 in the estimation requires the use of a lenghty numerical optimization and relies on assumptions about higher order expectations.

$$a_{2,2} = q_2 s_2 + (1 - q_2) \left(a_{1,1} + \frac{2c^2\sigma_1^2 \left(\phi(a_{1,1}; s_2 - Y, c\sigma_1) - \phi(a_{1,1}; s_2 + Y, c\sigma_1)\right)}{\mathrm{erf}\left(\frac{\sqrt{2}}{2c\sigma_1}(Y - s_2 + a_{1,1})\right) - \mathrm{erf}\left(\frac{\sqrt{2}}{2c\sigma_1}(a_{1,1} - s_2 - Y)\right)} \right) + e_{2,2} \tag{9}$$

To interpret these expressions, we contrast them with the solution obtained in the fully-rational model:

$$a_{1,2} = q_1 s_1 + (1 - q_1) a_{2,1}$$
$$a_{2,2} = q_2 s_2 + (1 - q_2) a_{1,1}$$

Notice that subject 1's reaction to the action of player 2 in the previous turn is now adjusted, relative to the fully rational case. This adjustment decreases the marginal weight put on the other's action the more extreme it is. To see that, figure 3 plots the expected value of the other player's signal, conditional on their estimate, $E(s_2|a_{2,1})$, relative to their estimate, $a_{2,1}$ (recall that in the fully-rational model $E(s_2|a_{2,1}) = a_{2,1}$), for the following parameter values: $s_1 = 500, Y = 30, \sigma_1 = 5, c = 1$.

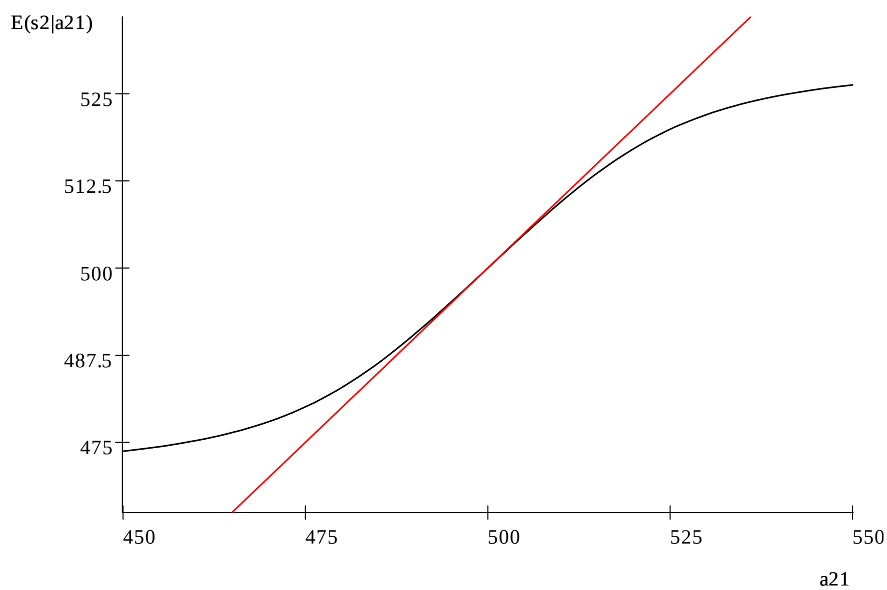


Figure 3: Turn 2 best response adjustment term effect - numerical example

One can see that for admissible values $(470 - 530)$, the adjustment is small. Going outside that range results in a steep adjustment in the direction of the signal possessed by the receiver.

6.2 Maximum Likelihood Estimation

Given the highly non-linear nature of the model specified above, we utilize maximum likelihood estimation. The objective is to provide point estimates of the following quantities:

subjects' individual confidence, subjects' degree of response to others' errors, their optimal actions at each point of the game, and the variance of those errors around the optimal actions. In estimating the model, we make the following specifications:

1. All errors are conditionally independent and identical of all other errors.
2. Subjective beliefs are constant within subject and across rounds. In order to allow for heterogeneous overconfidence while keeping the number of estimated parameters to a minimum, we further assume that subjective beliefs take one of three values (determined by the estimation).[38] Specifically, we assume that subjects in the BLT with objective probabilities of $[0, \frac{1}{3}]$, $[\frac{1}{3}, \frac{2}{3}]$,$[\frac{2}{3}, 1]$ hold beliefs equal to $\widetilde{q}_{BLT}^{[0,\frac{1}{3}]}$, $\widetilde{q}_{BLT}^{[\frac{1}{3},\frac{2}{3}]}$, and $\widetilde{q}_{BLT}^{[\frac{2}{3},1]}$ respectively, and subjects in the OCT hold beliefs equal to $\widetilde{q}_{OCT}^{[0,\frac{1}{3}]}$, $\widetilde{q}_{OCT}^{[\frac{1}{3},\frac{2}{3}]}$, and $\widetilde{q}_{OCT}^{[\frac{2}{3},1]}$.
3. Subjects need not have rational expectations about others' mistakes. That is, in turn 2 subjects may under or overestimate the magnitude of errors embedded in others' observed actions.

Assumption (2) assumes that subjects can be categorized into three groups based on how objectively informed they are. At the same time, it does not specify the relation between objective and subjective beliefs (see for example Tversky and Kahneman (1992)). This strikes a balance between the desire to limit the restriction about admissible forms of overconfidence while keeping the number of estimated parameters to a minimum. After all, if we wanted to allow subjects to have individually time varying beliefs would could end up having upward of 140 parameters. Assumption (3) allows us to test (rather than assume) whether subjects rationally respond to others' errors in actions.

This leads to the following likelihood function:

$$L = \prod_{v}^{\{BLT,OCT\}} \prod_{r=1}^{10}\prod_{i=1}^{72} \phi\left(a_{v,r,1,i} - a^*_{v,r,1,i}\left(I_{v,r,i,1}\right); 0, \sigma\right) \phi(a_{v,r,2,i} - a^*_{v,r,2,i}(c\sigma, I_{v,r,i,2}, \widetilde{q}_{v,i}); 0, \sigma) \tag{10}$$

Where the vectors $I_{v,r,i,1}$ and $I_{v,r,i,2}$ denote the set of observables (e.g., own signal, the signal of the other subject, etc.). The indexes v, i, r denote the treatment variant (BLT or OCT), the subject number, and the round number, respectively; the numbers $\{1, 2\}$ denote the turn number. For example, $a_{BLT,3,2,10}$ refers to action submitted in the BLT , round 3, turn 2, by player 10. We fit this model to the data collected in the first two turns under both treatments. Nonetheless, we allow for different levels of miscalibration across the two variants.[39]

In summary, the model estimated a total of eight parameters:

- Magnitude of observed error (difference between best-response and observed decisions), denoted by σ.

[38] In a previous version of the paper we estimate a model in the number of subjective-belief parameters is equal to the number of subjects, while the model estimated here makes due with only three parameters. We obtain qualitatively similar results to the ones presented here.

[39] We implicitly assume that magnitude of errors and average response to them is equal across treatments. Estimation results not reported here suggest that this assumption is not consequential.

- Subjective confidence, denoted by $\widetilde{q}_{BLT}^{[0,\frac{1}{3}]}$, $\widetilde{q}_{BLT}^{[\frac{1}{3},\frac{2}{3}]}$, $\widetilde{q}_{BLT}^{[\frac{2}{3},1]}$, and $\widetilde{q}_{OCT}^{[0,\frac{1}{3}]}$,$\widetilde{q}_{OCT}^{[\frac{1}{3},\frac{2}{3}]}$, and $\widetilde{q}_{OCT}^{[\frac{2}{3},1]}$.
- Degree of response to others' errors, denoted by c.

Since subjects are randomly paired into markets, and since we assume beliefs are fixed across subsequent rounds, the likelihood function captures a complex set of interactions. For example, subject #1 may be paired with subject #2 in round 1 and so their subjective probabilities will enter into both of their optimal actions. In round 2, subject #1 may be paired with subject #3 and subject #2 may be paired with subject #4. Since we are estimating $\widetilde{q}_i$ assuming it is constant across rounds 1 and 2, all four will be linked through the likelihood function. Thus, no analytical solution can be derived. Instead, we set-up and solve the maximization problem using a Sequential Quadratic Programming (SQP) method.[40] Since these procedures find local minima, results may be sensitive to starting points. To avoid biasing our estimations, we specify starting points for subjective probabilities equal to objective probabilities and repeat a subset of estimations by specifying different starting points. The results do not seem to be sensitive to variations in the starting points.

7 Model estimation results

Our main results are as follows:

1. **Model fit** - test of the NABB model and the two alternative models that it nests – Rational Expectations (RE) or Private Information (PI) – suggests that the alternative models can be rejected while the NABB cannot.
2. **Miscalibration** - subjects in the OCT exhibit moderate levels of overconfidence while subjects in the BLT are on average well calibrated.
3. **Rational response to mistakes**: subjects appear to be strategically responding to errors of others, while underestimating the magnitude of those errors.

7.1 Testing the NABB Model

We start by testing the null hypothesis that the NABB model fits the data well. To do that, we use a standard log-likelihood model test statistic:

$$G^2 = -2\Sigma X_i \ln\left(\frac{X_i^{prd}\left(\widehat{\theta}\right)}{X_i}\right) \sim \chi^2\left(n-k-1\right)$$

Where X_i denotes observation i and $X_i^{prd}\left(\widehat{\theta}\right)$ denotes model prediction for observation i given the vector of estimated parameters $\widehat{\theta}$. Since the calculated test statistic for the model is equal to 3.6076, we cannot reject the null that the NABB describes the data well at conventional statistical levels (corresponding p value is 0.5000).

Next, we test two alternative models: Rational Expectations (RE) and Private Information (PI), which are nested within the specifications of the NABB model.[41] If players

[40] Matlab code is available upon request.

[41] We chose these two alternatives as they often serve in examinations of information aggregation markets [see Plott and Sunder (1998)].

were perfectly calibrated and made no errors, we would obtain the RE results; if players were all completely confident in their private information, thus disregarding the actions of others, we would obtain the PI results.

That is, the RE model implies that subjects are well calibrated *and* that the error magnitude to which they are responding is equal to zero. The model, discussed in the theory section, assumes that:

- Players' subjective probabilities of being perfectly informed are equal to the objective ones. We bootstrap from the distribution of outcomes (die throw or quiz score) to calculate the probability of each draw/score being ranked above average among a group of 6, which determined the objective probability of being perfectly informed.
- Players' actions in *all* turns are Bayesian. That is, subject i reacts in a particular turn to the perfectly rational, and not to the actual estimate submitted by player j in the previous turn.

In the context of the NABB model, these implications can be translated into a set of non-linear restrictions on the parameters: $\{c\sigma = 0, q - \widetilde{q} = 0\}$. Using Wald test we find that the null can be strongly rejected ($W = 6.0102e + 010$ in the BLT and $4.5185e + 010$ in the OCT).

The PI model suggests that agents attend only to their private signal, thereby completely disregarding the observed actions of others. In our setting, this translates into players holding subjective beliefs of being perfectly informed equal to one; that is, $\left\{\widetilde{q}_{BLT}^{[0,\frac{1}{3}]} = 1, \widetilde{q}_{BLT}^{[\frac{1}{3},\frac{2}{3}]} = 1, \widetilde{q}_{BLT}^{[\frac{2}{3},1]} = 1\right\}$ and $\left\{\widetilde{q}_{OCT}^{[0,\frac{1}{3}]} = 1, \widetilde{q}_{OCT}^{[\frac{1}{3},\frac{2}{3}]} = 1, \widetilde{q}_{OCT}^{[\frac{2}{3},1]} = 1\right\}$. We examine the prediction by constructing a Wald test corresponding to those restrictions. The results suggest that we can reject the hypothesis for each of the treatments (test statistic equal to 371.2 in the BLT and 219.1 in the OCT).

7.2 Miscalibration

We define miscalibration, denoted by $MC_{i,r}$ as the difference between subjective (e.g., $\widetilde{q}_{BLT}^{[0,\frac{1}{3}]}$) and objective probability (e.g., $q_{BLT}^{[0,\frac{1}{3}]}$) of being perfectly informed. This measure captures the degree to which an individual is over/underconfident, adjusting for their objective confidence. As a convention, we interpret positive miscalibration as representing overconfidence and negative miscalibration as representing underconfidence. Note that an individual that scored high on the initial task (die roll or quiz) would be assigned high confidence level (q_i) but would not necessarily be assigned high overconfidence level ($\widetilde{q}_i - q_i$).

We first seek to establish that the conjectured treatment effect induces (overall) overconfidence in the OCT but *not* the BLT. Table 5 reports mean and standard errors of estimated miscalibration levels grouped by treatment (BLT and OCT) and by objective confidence ('poorly informed', 'averagely informed', 'well informed', and 'all').

First, we find that overall miscalibration in the BLT is statistically indistinguishable from zero (mean estimate of -0.0427 with standard errors of 0.0300) while overall miscalibration in the OCT is positive (mean estimate of 0.1477 with standard error of 0.0250). Subjects appear to be overconfident in the OCT but well-calibrated in the BLT. The level of overconfidence in the OCT is not only statistically but also economically significant; it suggests that on average subjects (in the OCT) believe that their probability of being perfectly informed is about 15% greater than it actually is.

Second, we test the null of no difference in miscalibration level for all subgroups.[42] The Wald test statistic ($W = 37.59$) allow us to reject the null at the 1% level.

Third, overconfidence seem to be driven primarily by the poorly informed subjects in the OCT. The estimated level of miscalibration for the averagely informed and well informed are similar across treatments (for the averagely informed it is 0.0035 (BLT) and 0.010 4 (OCT), and for the well informed it is $-0.104\,3$ (BLT) and 0.0372 (OCT)). Comparing the levels for the poorly informed group suggests major differences: miscalibration is -0.03 50 in the BLT and 0.432 7 in the OCT. This finding suggests that overconfidence may be particularly damaging for price efficiency as it is concentrated among the poorly informed.[43] It can also serve to explain why price index was found to be serial autocorrelation in the OCT.

7.3 Rational best-response

The NABB model allows for the *possibility* that players best respond to errors made by fellow players. Rather than impose this as an assumption, we estimate the model in a way that allows subjects to respond to some fraction, denoted by c, of the empirically observed magnitude of errors. Recall that subjects generally cannot directly observe errors made by others, let alone measure their magnitude. Thus, the parameter c estimates subjects' implicit response to others errors such that parameter value of zero means that subjects completely ignore others' errors and parameter value of one suggest that subjects respond to others' errors and have rational expectations about their magnitude.

The estimate of the parameter c is 0.7839 (with standard error of 0.0631), implying that subjects' decisions take into account about 80% of observed error size. The main hypothesis we test is that subjects do not appear to be best responding to errors made by others. This null can be strongly rejected with test statistic of $W = 5e + 009$. That is, subjects' decisions seem to respond to others' errors. At the same time, they do not appear to hold fully rational expectations about others' errors as they underestimate the magnitude of errors; we can reject the null that $c = 1$ at conventional significance levels with test statistics of $W = 239.2$.

8 Summary

In this paper we suggest a game through which we study - theoretically and empirically - how market participants aggregate multidimensional private information. In order to separate out two widespread behavioral biases, erroneous actions and mistaken beliefs, we combine an experimental design, which controls for the presence of overconfidence, and an econometric model that nests both biases. As a result, we are able to estimate subjects' overconfidence and quantify the importance of errors.

We show that subjects strategically respond to others' mistakes and that this feature generates a pattern of volume and price efficiency index levels similar to those predicted by models based on overconfidence. Nonetheless, most canonical predictions linking investors'

[42] Using notation introduced above we test the null that $\left\{ \widetilde{q}_{BLT}^{[0,\frac{1}{3}]} = \widetilde{q}_{OCT}^{[0,\frac{1}{3}]}, \widetilde{q}_{BLT}^{[\frac{1}{3},\frac{2}{3}]} = \widetilde{q}_{OCT}^{[\frac{1}{3},\frac{2}{3}]}, \widetilde{q}_{BLT}^{[\frac{2}{3},1]} = \widetilde{q}_{OCT}^{[\frac{2}{3},1]} \right\}$.

[43] To see the intuition, imagine two subjects - one poorly informed and the other well informed. If both of them are equally overconfident, the price index would not overweight either of their private signals. However, if the poorly informed subject is more overconfident than the well informed agent, the price index would overweight the signal held by the former, decreasing the price informativeness index.

overconfidence to markets is borne out: we find a higher volume and lower informational efficiency index levels in the overconfidence treatment. Serial correlation of price index is observed only in the presence of overconfidence. We suggest that this stems from the fact that the worst informed are found to be the most overconfident, imposing negative externalities of their overconfidence.

We believe that the setup and the results discussed here open the door to promising future research. We have started exploring the lead-lag interaction between the volume index, the change in volume index, the price error index, and the return index. Preliminary results suggest intriguing dynamics, in the spirit of Llorente, Michaely, Saar and Wang (2002). While this paper made the simplifying assumption that overconfidence is static, it would we interesting to further explore the role market interaction plays in the process of overconfidence updating. Also, in this experiment we have made information acquisition exogenous but it would be interesting to endogenize it, allowing one to study how miscalibration feeds into investment in information. Related to that, we study a setting in which only private signals are present. This framework could be easily extended to include public information, thereby allowing one to ask how overconfidence effects the response to that type of news.[44] Last, since our game does not depend on a large number of participants, it provides a solution to situations characterized by either asynchronous trading or thin markets, a topic of interest to both experimentalists [see Plot (2000)] and practitioners [see Lange and Economides (2003)].

References

[1] Alpert M., and Raiffa H. (1982): "A progress report on the training of probability assessors", in Kahneman D., Slovic P., and Tversky A., Judgment Under Uncertainty: Heuristics and Biases, Cambridge University Press

[2] Anderson L. R., and Holt C.H. (1997): "Information cascades in the laboratory", American Economic Review, December, 847-862

[3] Arrow K.J. (1986): "Rationality of self and others in an economic-system", Journal of Business, 59 (4): S385-S399 Part 2

[4] Becker G.S. (1962): "Irrational behavior and economic theory", Journal of Political Economy, LXX(1) pp. 1-13

[5] Biais B. Hilton D., Mazurier K., and Pouget S. (2002): "Psychological dispositions and trading behavior", CEPR working paper

[6] Bikhchandani S., Hirshleifer D., Welch I. (1998): "Learning from the behavior of others: Conformity, fads, and informational cascades", Journal of Economic Perspectives 12 (3): 151-170

[7] Camerer F. C. (1987): "Do biases in probability judgment matter in markets? experimental evidence", The American Economic Review, 77(5) pp.981-997

[44] This topic has been of interested to a large body of empirical literature that documents under-reaction to public announcements, such as earning releases, and theoretical literature that suggest a roll for overconfidence in this.

[8] Camerer F. C. (2003): "Behavioral game theory: experiments in strategic interaction", Princeton University Press

[9] Camerer C, Lovallo D. (1999): "Overconfidence and excess entry: An experimental approach", American Economic Review 89 (1): 306-318

[10] Celen B. and Kariv S. (2003): "Distinguishing Informational Cascades from Herd Behavior in the Laboratory", Forthcoming, American Economic Review

[11] Daniel K., Hirshleifer D., Subrahmanyam A.(1998): "Investor psychology and security market under- and overreactions", Journal of Finance, 53 (6): 1839-1885

[12] Davis D.D and Holt C.A. (1993): "Experimental economics", Princeton University Press

[13] Deaves R, Luders E. and Luo G.Y. (2003): "An experimental test of the impact of overconfidence and gender on trading activity", working paper, McMaster University

[14] DeBondt, W. and R. Thaler (1985): "Does the stock market overreact", Journal of Finance, 40, 793-807

[15] DeBondt, W. and R. Thaler (1995): "Financial decision-making in markets and firms: a behavioral perspective", in R., Jarrow, V. Maksimovic and W. Ziemba, eds., Finance, Elsevier North Holland, Amsterdam, Handbook in Operation in Research and Management Science, Vol. 9 Chap 13. pp. 385-410

[16] Diamond D.W. and Verrecchia R.E (1981): "Information aggregation in a noisy rational expectations economy", Journal of Finance, 9, pp. 221-235

[17] Fiscoff, B., Slovic P., and Lichtenstein S. (1977): "Knowing with certainty: the appropriateness of extreme confidence", Journal of Experimental Psychology, 3, 552-564

[18] Forsythe R., Palfrey T.R., Plott C.R. (1982): "Asset valuation in an experimental market", Econometrica, 50 (3): 537-567

[19] Frank J. D. (1935): "Some psychological determinants of the level of aspiration", American Journal of Psychology, 47, 285-293

[20] Friedman D. and Rust J. (1991): "The double auction market institutions, theories, and evidence.:, Technical report, Santa Fe Institute Studies in the Sciences of Complexity

[21] Friedman D. and Sunder S. (1994): "Experimental methods: a primer for economists", Cambridge University Press

[22] Gale D. (1996): "What have we learned from social learning?", European Economic Review, 40 (3-5): 617-628

[23] Gervais S., Odean T. (2001): "Learning to be overconfident", Review of Financial Studies, 14 (1): 1-27

[24] Glaser M., and Weber M. (2003): "Overconfidence and trading volume", working paper

[25] Glosten L.R., Milgron P.R. (1985): "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders", Journal of Financial Economics, 14 (1): 71-100

[26] Gode D.K. and Sunder S. (1993): "Allocative efficiency of markets with zero-intelligence traders - market as a partial substitute for individual rationality", Journal of Political Economics, 101 (1): 119-137

[27] Goeree J.K and Holt C.A. (1999): "Stochastic game theory: for playing games, not just for doing theory", Proc. Natl. Acas. Sci. USA, Vol 96, pp. 10564-10567

[28] Griffin D.W. and Varey C.A. (1996): "Commentary: towards a consensus on overconfidence", Organizational Behavior and Human Decision Processes, Vol. 65(3), pp. 227-231

[29] Grossman S.J. and Stiglitz J.E. (1980): "On the impossibility of informationally efficient markets", American Economic Review, 70 (3): 393-408

[30] Hirshleifer D. and Teoh S.H. (2003): "Herd behavior and cascading in capital markets: a review and synthesis", European Financial Management, Vol. 9, No. 1, 25-66

[31] Jegadeesh N., and Titman (1993): "Returns to buying winners and selling losers: implications for stock market efficiency", Journal of Finance, 48, pp. 65-91

[32] Kagel J.H. and Roth A.E. (1995): "Handbook of experimental economics", Princeton University Press

[33] Kahneman D., Slovic P. and Tversky P. (1982): "Judgment under uncertainty: Heuristics and biases", Cambridge University Press

[34] Keren G. B. (1987): "Facing uncertainty in the game of bridge: a calibration study", Organizational Behavior and Human Decision Processes, 39, pp. 98-114

[35] Keren G. B. (1997): "On the calibration of probability judgments: Some critical comments and alternative perspectives", Journal of Behavioral Decision Making, 10 (3): 269-278

[36] Kirchler, E., and B. Maciejovsky (2002): "Simultaneous over- and underconfidence: Evidence from experimental asset markets", Journal of Risk and Uncertainty, 25: 65-85.

[37] Kyle A.S., Wang F.A. (1997): "Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?", Journal of Finance, 52 (5): 2073-2090

[38] Lange J. and Economides N. (2003): "A parimutuel market microstructure for contingent claims", Stern School of Business

[39] Llorente G., R. Michaely, G. Saar and J. Wang (2002). "Dynamic volume-return relation of individual stocks", Review of Financial Studies, 15(4): 1005-1047

[40] Palfrey T. R., and McKelvey R. (1995): "Quantal response equilibria in normal form games", Games and Economic Behavior (special issue on Experimental Game Theory), v. 10, 6-38

[41] Palfrey T. R., and McKelvey R. (1998): "Quantal response equilibria in extensive form games", Experimental Economics, v. 1, pp. 9-41

[42] Odean T. (1998): "Volume, volatility, price, and profit when all traders are above average", Journal of Finance, 53 (6): 1887-1934

[43] Plott C.R. (2000): "Markets as information gathering tools", Southern Economic Journal 67(1):1-15.

[44] Plott C.R. and Sunder S. (1982): "Efficiency of experimental security markets with insider information - an application of rational-expectations models", Journal of Political Economy, 90 (4): 663-698

[45] Plott C.R., Sunder S. (1988): "Rational-expectations and the aggregation of diverse information in laboratory security markets", Econometrica, 56 (5): 1085-1118

[46] Roll, R. (1984): "Orange juice and weather", American Economic Review, 74, 5, 861-880

[47] Roll R. (1986): "The hubris hypothesis of corporate takeovers", Journal of Business, 59 (2): 197-216 Part 1

[48] Shalen C.T (1993): "Volume, volatility, and the dispersion of beliefs", The Review of Financial Studies, 6(2), 405-434

[49] Shiller R.J. (1997): "Human behavior and the efficiency of the financial system", in J. Taylor J.B. and M. Woodford, Editors, Handbook of Macroeconomics

[50] Sunder S (1995): "Experimental asset markets: a survey", in Kagel J.H. and Roth A.E., Handbook if Experimental Economics, Princeton University Press

[51] Svenson O. (1981): "Are we all less risky and more skillful than our fellow drivers", Acta Psychologica, 47 (2): 143-148

[52] Taylor S., and Brown J. D. (1988): "Illusion of well-being: a social psychological perspective on mental health", Psychological Bulletin, 103, 193-210

[53] Tversky A., and Kahneman D. (1992): "Advances in prospect thoery: Cumulative representation of uncertainty", Journal of Risk and Uncertainty, Vol. 5, 297-323.

[54] Wang F.A (1998): "Strategic trading, asymmetric information and heterogenous prior beliefs", Journal of Financial Markets, 1, 321-352

9 Appendix

9.1 Proofs

Derivation of turn two optimal actions:

Generally, it is easy to show that if x, y, z are r.v.:
$\Pr(x|y,z) = \frac{\Pr(y|x,z)\Pr(x|z)}{\Pr(y|z)}$ and therefore, $\Pr(s_2|a_{21}, s_1) = \frac{\Pr(a_{21}|s_2,s_1)\Pr(s_2|s_1)}{\Pr(a_{21}|s_1)}$ and calculating the elements of this expression we get:
- $\Pr(a_{2,1}|s_2, s_1) = \Pr(a_{2,1}|s_2) = \phi(a_{21} - s_2; 0, \sigma_1)$
- $\Pr(s_2|s_1) = \frac{1}{2Y}$

$\Pr(a_{21}|s_1) = \int\limits_{s_1-Y}^{s_1+Y} \frac{1}{2Y}\phi(a_{21} - s_2; 0, \sigma_1)ds_2 =$

$$\frac{\sqrt{4\pi}\,\mathrm{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}Y\sqrt{2}-\frac{1}{2}s_1\sqrt{2}+\frac{1}{2}a_{21}\sqrt{2}\right)\right)}{\frac{1}{8\sqrt{\pi Y}}}$$
$$-\frac{\sqrt{4\pi}\,\mathrm{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}a_{21}\sqrt{2}-\frac{1}{2}s_1\sqrt{2}-\frac{1}{2}Y\sqrt{2}\right)\right)}{\frac{1}{8\sqrt{\pi Y}}} =$$
$$\frac{1}{4Y}\left(\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}\left(Y-s_1+a_{21}\right)\right)-\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}\left(a_{21}-s_1-Y\right)\right)\right)$$

Collecting these terms we obtain that:

$$E(s_2|a_{21},s_1) = \int_{s_1-Y}^{s_1+Y} s_2 \Pr(s_2|a_{21},s_1)ds_2 =$$
$$\int_{s_1-Y}^{s_1+Y} s_2 \frac{\phi(a_{21}-s_2;0,\sigma_1)\frac{1}{2Y}}{\frac{1}{4Y}\left(\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)\right)}ds_2 =$$
$$-\frac{\sqrt{2}\left(\frac{1}{2}\sigma_1\sqrt{2}e^{-\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\left(-\sigma_1\sqrt{2}\exp\left(\frac{a_{21}}{\sigma_1^2}(Y+s_1)-\frac{1}{2\sigma_1^2}(Y+s_1)^2\right)-\sqrt{\pi}a_{21}e^{\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\mathrm{erf}\left(\frac{1}{2}\frac{a_{21}}{\sigma_1}\sqrt{2}-\frac{1}{2\sigma_1}(Y+s_1)\sqrt{2}\right)\right)-\right)}{\sqrt{\pi}\sigma_1\,\mathrm{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}a_{21}\sqrt{2}-\frac{1}{2}s_1\sqrt{2}-\frac{1}{2}Y\sqrt{2}\right)\right)-\sqrt{\pi}\sigma_1\,\mathrm{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}Y\sqrt{2}-\frac{1}{2}s_1\sqrt{2}+\frac{1}{2}a_{21}\sqrt{2}\right)\right)}+$$
$$\frac{\sqrt{2}\left(\frac{1}{2}\sigma_1\sqrt{2}e^{-\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\left(-\sigma_1\sqrt{2}\exp\left(\frac{a_{21}}{\sigma_1^2}(s_1-Y)-\frac{1}{2\sigma_1^2}(s_1-Y)^2\right)-\sqrt{\pi}a_{21}e^{\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\mathrm{erf}\left(\frac{1}{2}\frac{a_{21}}{\sigma_1}\sqrt{2}-\frac{1}{2\sigma_1}\sqrt{2}(s_1-Y)\right)\right)\right)}{\sqrt{\pi}\sigma_1\,\mathrm{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}a_{21}\sqrt{2}-\frac{1}{2}s_1\sqrt{2}-\frac{1}{2}Y\sqrt{2}\right)\right)-\sqrt{\pi}\sigma_1\,\mathrm{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}Y\sqrt{2}-\frac{1}{2}s_1\sqrt{2}+\frac{1}{2}a_{21}\sqrt{2}\right)\right)} =$$
$$a_{21} + \frac{e^{-\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\left(\sigma_1\sqrt{2}\left(\exp\left(\frac{a_{21}}{\sigma_1^2}(s_1-Y)-\frac{1}{2\sigma_1^2}(s_1-Y)^2\right)-\exp\left(\frac{a_{21}}{\sigma_1^2}(Y+s_1)-\frac{1}{2\sigma_1^2}(Y+s_1)^2\right)\right)\right)}{-\sqrt{\pi}\,\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)+\sqrt{\pi}\,\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)} =$$
$$a_{21} + \frac{2\sigma_1^2(\phi(a_{21};s_1-Y,\sigma_1)-\phi(a_{21};s_1+Y,\sigma_1))}{\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}$$

9.2 Tables

Table 1: Comparison between observed, fully-rational and boundedly rational models

Means (BLT/OCT)				
Turn	1	2	3	4
		Volume index		
Obs	17.57/16.97	7.51/11.95	5.62/13.19	5.00/9.74
RE	14.95/15.34	4.03/3.80	0.00/0.00	0.00/0.00
NABB	14.95/15.34	4.97/6.15	NA	NA
Obs-RE	2.62/1.63	3.48/8.15	5.62/13.19	5.00/9.74
		Price error index		
Obs	8.93/8.51	4.46/7.41	3.41/7.77	3.49/6.96
RE	7.47/7.67	3.14/3.15	0.68/0.17	0.68/0.17
NABB	7.47/7.67	4.69/6.88	NA	NA
Obs-RE	1.46/0.84	1.32/4.26	2.73/7.60	2.81/6.79

Notes: this table reports the average levels of volume and price error indexes as given by observed data (labeled "Obs"), predictions of the Rational Expectations model (labeled "RE") model, predictions of the Noisy Actions Biased Beliefs (labeled "NABB") model, and observed minus predicted rational expectations model (labeled "Obs-RE"). The left (right) figures represent the results for the BLT (OCT).

Table 2: Change in price-error and volume indexes

(p values)	Turn 1 vs 2	Turn 2 vs 3	Turn 3 vs 4	Turn 2 vs 4
		Price Error index		
BLT	0.0000	0.0021	0.9330	0.0216
OCT	0.0143	0.8068	0.2638	0.2414
		Volume index		
BLT	0.0000	0.0027	0.3950	0.0005
OCT	0.0000	0.4062	0.0030	0.0313

Notes: this table reports p values resulting from non-parametric (Wilcoxon) tests of the null that the price error index or the volume index remain constant across *turns* for each treatment separately. For example, in column one, labeled "Turn 1 vs 2", the lines labeled "BLT" report the probability that observed turn 1 price error index or volume index are on average the same as those observed in turn 2.

Table 3: Price error and volume index comparison across treatments

(p value reported)				
Turn	1	2	3	4
		Price Error index		
All	0.338	0.000	0.000	0.000
Rounds 1-3	0.840	0.046	0.004	0.000
Rounds 4-7	0.369	0.014	0.000	0.004
Rounds 8-10	0.300	0.000	0.000	0.009
		Volume index		
All	0.610	0.004	0.000	0.000
Rounds 1-3	0.122	0.580	0.008	0.001
Rounds 4-7	0.442	0.156	0.002	0.009
Rounds 8-10	0.095	0.002	0.000	0.104

Notes: this table reports non-parametric (Mann-Whitney) test results of the null that median observed price error and volume indexes are the same across *treatments* (BLT vs. OCT) for a given subset of rounds and turns. For example, the column labeled "3" and the row labeled "Rounds 1-3" under "Price error index" report the probability that the price error index level is the same across treatments in turn 3 during rounds 1-3.

Table 4: Return autocorrelation

	$\widehat{a}$	$\widehat{b}$	$\widehat{c}$
	$Ln(V_i/P_{4,i}) = a + bLn(P_{4,i}/P_{1,i}) + e_i$		
BLT	0.000	-0.000	
OCT	-0.002	-0.109^{***}	
	$Ln(V_i/P_{4,i}) = a + bLn(P_{4,i}/P_{1,i}) + c(Vol_{4,i} - Vol_{1,i}) + e_i$		
BLT	0.000	0.000	0.000
OCT	-0.003^{**}	-0.140^{***}	-0.0003^{***}

Notes: this table reports regression results of log price index change, from open to close of round (turns 1 to 4), $Ln(P_{4,i}/P_{1,i})$, and the volume index changes from open to close of round (turns 1 to 4), $Vol_{4,i} - Vol_{1,i}$, on log ratio of fundamental value to closing round price index, $Ln(V_i/P_{4,i})$. Superscript **(***) denote significance level of 5%(1%).

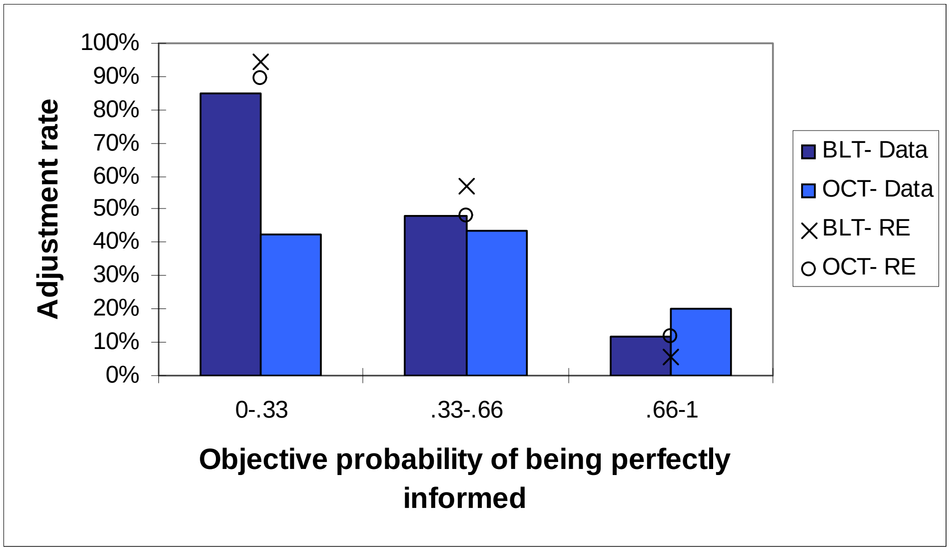
Table 5: Mean miscalibration

Group	$[0, 1/3]$	$[1/3, 2/3]$	$[2/3, 1]$	All
		BLT		
Miscalibration	-0.0350	0.0035	-0.1043^{**}	-0.0427
S.E.	0.0631	0.0500	0.0372	0.0300
		OCT		
Miscalibration	0.4327	0.0104	-0.0870^{*}	0.1477^{**}
S.E.	0.0414	0.0806	0.0498	0.0250

Notes: this table reports the average miscalibration (across subjects) along with test results of the null that the average level is equal to zero. Superscript ** (*) represent significance levels greater than 5% (10%).

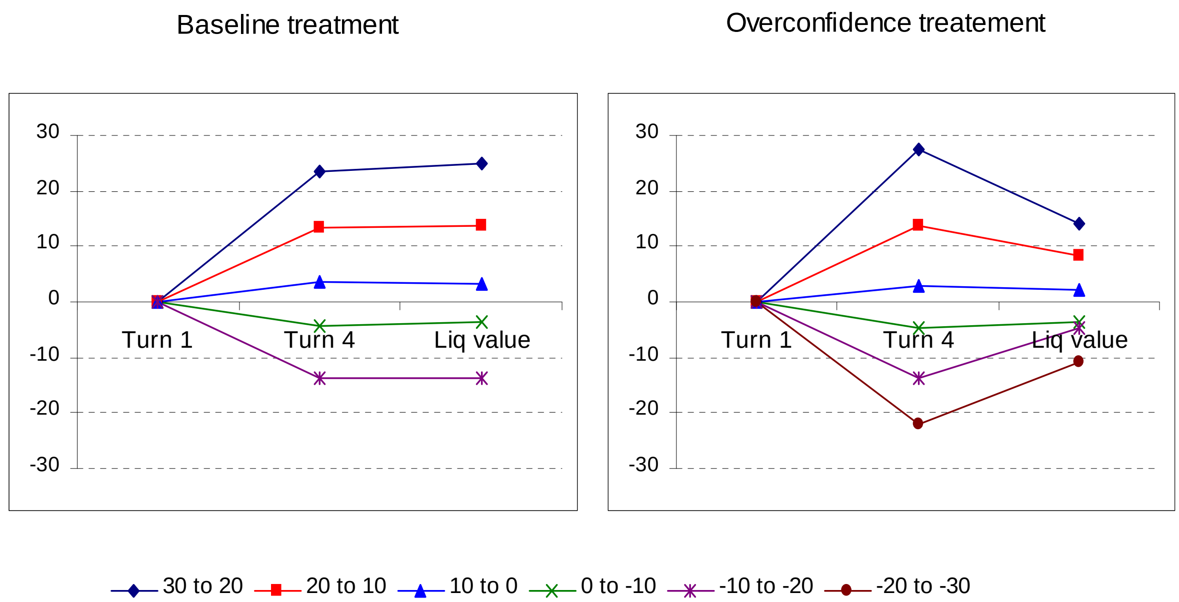
9.3 Figures (not included in the text)

Figure 4: Adjustment rate across treatments by initial probabilities



Note: this figure reports average adjustment rates, defined as the ratio of own estimate change from turn 1 to turn 2, over the absolute difference between subjects' turn 1 estimates (see figure 2). Each column represents the average adjustment rate across all subjects within a given range of objective probabilities (0 to $\frac{1}{3}$, $\frac{1}{3}$ to $\frac{2}{3}$, and $\frac{2}{3}$ to 1), grouped by treatment (BLT or OCT). In addition, we plot the fully-rational-no-overconfidence (RE) predicted average adjustment rate for each of the subgroups.

Figure 5: Return index autocorrelation



Notes: these figures show the price index in turn 1, turn 4, and the fundamental value (labeled "Liq value") such that all market instances are grouped by the change from turn 1 to turn 4. For example, the top line in each of the figures represents the average price index level of all market instances in which price index increased by 20 to 30 points from turn 1 to turn 4. The top (bottom) panel represents data gathered in the BLT (OCT). Prices are normalized to be equal to zero in turn 1.