Distinguishing Bounded Rationality from Overconfidence in Financial Markets - Theory and Experimental Results[1]

Shimon Kogan[2]
Haas School of Business
University of California at Berkeley

November 18, 2004

[1]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, Stefano DellaVigna, Teck Ho, Botond Koszegi, Lars Lochstoer, Rich Lyons, Ulrike Malmendier, Barbara Mellers, Finance Faculty Lunch participants, seminar participants at the London Business School and the Economics Department at University of California, Berkeley, for their valuable comments. I am particularly indebted to Jonathan Berk, Shachar Kariv, John Morgan, Jacob Sagi and Nancy Wallace for their insightful suggestions and continuous support. All errors are of course mine alone.

[2]Haas School of Business, U.C. Berkeley, S545 Student Services Bldg., #1900, Berkeley, CA 94720, tel: (510) 642-5295, fax: (510) 643-1420, email: kogan@haas.berkeley.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, price errors and volatility 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 observationally equivalent 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 about 40% of *excess* volume is attributable to strategic response to errors, while the remaining is attributable to overconfidence. If one looks at price errors or price volatility, similar results are obtained. Further, we show that the distribution of estimated individual level overconfidence is linked to the observed price reversals, present only in the overconfidence-induced sessions. Additional findings are discussed.

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 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: 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' ac-

[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.

tions [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 turn is randomly chosen and each subject is paid according to the accuracy of her guess 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 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 return and price volatility indexes.

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

[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.

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

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 that 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, price error and volatility 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. Specifically, we find that about 40% of *excess* volume index is attributable to strategic response to errors, while the remaining is attributable to overconfidence. If one looks at price error or price volatility indexes, 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 allows us to back out participants' subjective confidence in their information *at each stage* of the game, calculate best-responses and estimate the magnitude of errors.[6] Fitting the model on both treatments confirms that subjects are on average overconfident in the OCT (around 11%) but not in the BLT and that mean-zero errors are present in all sessions and are economically significant.[7] Further, the NABB model can account for virtu-

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

[7] Average overconfidence of 11% means that subjects perceive the probability of being perfectly informed to be on average 61% while in fact it is 50% (since half of the subjects

ally all the deviations from the fully-rational model predictions, observed in the BLT.

We utilize the model to provide dynamic estimations of individual subjective probabilities (to the best of our knowledge, this is the first experiment do so without using direct elicitation or restricting beliefs).[8,9] We show that contrary to the common practice in modeling overconfidence, which assumes a constant miscalibration level, not all overconfident individuals are alike: subjects that are very likely to be perfectly informed tend to be somewhat underconfident while subjects that are very unlikely to be perfectly informed tend to be overconfident.[10] Because of that, the information of the *poorly* informed is *overweighted* relative to the information of the well-informed, resulting in price index reversals and negative return index autocorrelation, observed in the OCT but not in the BLT.

Last, we explore the change in beliefs in the course of the session and find that in the BLT subjects update their type probability beliefs (how likely they are to receive a precise signal) and become better calibrated over time. The same learning process is weaker in the OCT, which suggests that miscalibration is not easily eradicated by market interaction and may thus affect the long term behavior of markets.

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 Theory

2.1 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

are perfectly informed by design).

[8]Elicitation may either influence the way subjects act or, at the other extreme, have little to do with how they actually form their decisions [see Deaves, Luders and Luo (2003)].

[9]Ignoring heterogeneity or forcing it to take a particular form may be too strong an assumption, as we show later.

[10]This is consistent with findings in other studies, see discussion in the Results section.

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:

- 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 obtain 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, denoted by $\widetilde{q_i}$, about 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.[11] The game consists of 3 turns: at the beginning of each turn, t, both players simultaneously submit an action, $a_{i,t}$.[12] At the end of each turn, both players' estimates are announced.[13] 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: $E\left(v|I_{i,t}\right) \in \arg\max E\left[u_i(\pi_i(a_{i,t}, v))|I_{i,t}\right]$, 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

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

[12] 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.

[13] 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.

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.

2.1.1 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 PBE where players optimal actions are:* $a_{i,t}^* = E(v|I_{i,t})\ \forall i, t$.

Proof In turn 1, optimal actions are:[14]

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

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:

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

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

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[15], $a^*_{i,3} = E(v|\{s_i, q_i, a^*_{j,1}, a^*_{j,2}\}) = E(v|\{s_i, q_i, s_j, q_j\}) = a^*_{j,3}$

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

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

Proposition 2 *The 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 (see table 1). 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:* $Vol_t = |a_{i,t} - a_{j,t}|$

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

Turn 1	Turn 2	Turn 3
Own signal	Own signal	Own signal
	Other's signal	Other's signal
	Own signal precision	Own signal precision
		Other's signal precision

Table 1: Aggregation of information in the game

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.*

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

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 the realized signal error of the imperfectly informed trader and in 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.

2.1.2 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 $0 \leq \widetilde{q}_i \leq 1 \; \forall i$.[16]

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).[17] Now, we can rewrite volume and price error indexes in the case of miscalibration, denoting them with superscript MC.

Volume index**:**

- 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: $PE_1^{MC} = \left|\frac{e}{2}\right|$

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

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

- 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: $PE_3^{MC} = \left| \frac{2\left(1_{\widetilde{s}_i > \widetilde{s}_j}\widetilde{s}_i + 1_{\widetilde{s}_j > \widetilde{s}_i}\widetilde{s}_j\right) + 1_{\widetilde{s}_j = \widetilde{s}_j}(\widetilde{q}_i s_i + (1-\widetilde{q}_i)s_j + (1-\widetilde{q}_j)s_i + \widetilde{q}_j s_j)}{2} - v \right| =$
  $\left| 1_{\widetilde{s}_i > \widetilde{s}_j}\widetilde{s}_i + 1_{\widetilde{s}_j > \widetilde{s}_i}\widetilde{s}_j + \frac{1_{\widetilde{s}_j = \widetilde{s}_j}(\widetilde{q}_i s_i + (1-\widetilde{q}_i)s_j + (1-\widetilde{q}_j)s_i + \widetilde{q}_j s_j)}{2} - v \right|$

While the link between subjective beliefs and the objective quantities may take on many different deterministic or stochastic forms, hereafter we consider a particular specification. We assume that miscalibration is a linear function of objective beliefs: $MC_i = \delta + \gamma q_i$ where $0 < \gamma \leq -1, 0 < \delta < 1, \delta + \gamma \leq 0$ are known constants and are the same for all subjects. This specification and our estimation ($\widehat{\delta} = 0.2950, \widehat{\gamma} = -0.4711$) is supported by many studies in the calibration literature (see for example Lichtenstein, Fischhoff and Phillips (1977)). The intuition is that subjects that are very likely to be perfectly informed are underconfident while subjects that are very unlikely to be perfectly informed are overconfident.

Proposition 8 *Given the form of individual overconfidence assumed above, the volume index is (weakly) increasing in average overconfidence for all turns (a.s.) iff* $q_i + q_j \geq \frac{1-2\delta}{1+\gamma}$.

Proof. For turns 1 and 3 the proof is trivial.
For turn 2, recall that $Vol_2^{OC} = |(1 - (q_i + q_j) - (MC_i + MC_j))(s_j - s_i)|$
Denoting $(q_i + q_j) \equiv q_{ij}$ and $(MC_i + MC_j) \equiv MC_{ij}$, and squaring both sides of the expression,
$\left(Vol_2^{OC}\right)^2 = (1 - q_{ij} - MC_{ij})^2(s_j - s_i)^2 = (1 - q_{ij} - MC_{ij})^2(e_j)^2$
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 iff
$q_{ij} + MC_{ij} - 1 > 0$
Recall that $MC_i = \delta + \gamma q_i$. Thus, we obtain that $q_{ij} > \frac{1-2\delta}{1+\gamma}$ ■

Proposition 9 *Given the form of individual overconfidence assumed above, the expected price error index is (weakly) increasing in dispersion of overconfidence for all turns.*

Proof See page 40.

3 Experimental Design

3.1 General

The experiment was run at the Haas School of Business: a total of 12 sessions were conducted in which 72 subjects participated. Out of those, 5 were Baseline Treatment (BLT) and 7 were Overconfidence Treatment (OCT).[18] 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 [see Appendix] aloud and answered questions in private. 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]

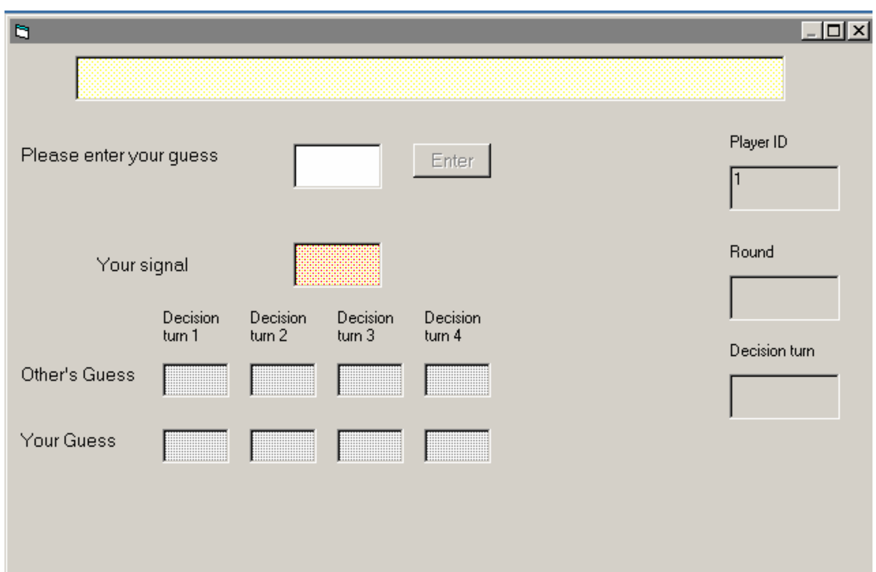


Figure 1: Interface screen shot

[18] The order of treatments was determined randomly.

[19] The application developed by the author for this experiment is available upon request.

3.2 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]

[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.

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

3.4 Assignment

Pairing into markets was randomly determined while ensuring that exactly one subject is perfectly informed while the other 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]

3.5 Actions and payoff

At the beginning of each turn t subjects simultaneously submit their reports $a_{i,t}$ by entering a number on their screen. No restrictions are imposed on the value the report can take. Upon receiving submissions from both subjects,

[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 report 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.

the turn comes to an end and no changes are accepted. At that point subjects are informed of each others' report 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:

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

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]

4 Model independent results

Our main findings can be divided into

- individual level

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

- subjects seem to incorporate their signal precision, in addition signal realization, into their estimates
- subjects seem to exhibit overconfidence and react to others' propensity for making mistakes

- market level
  - markets seem to aggregate and disseminate information under both treatments but not with the same degree of efficiency
  - the volume, price error and price volatility index levels are in excess of the fully rational model prediction in *both* treatments
  - comparing the levels of volume, price error and price volatility indexes across treatments we find that they are generally higher in the OCT
  - the price index exhibits reversals in the OCT but *not* in the BLT

4.1 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.

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

- 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 showed in the theory section, in the absence of errors, subject i's turn 2 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

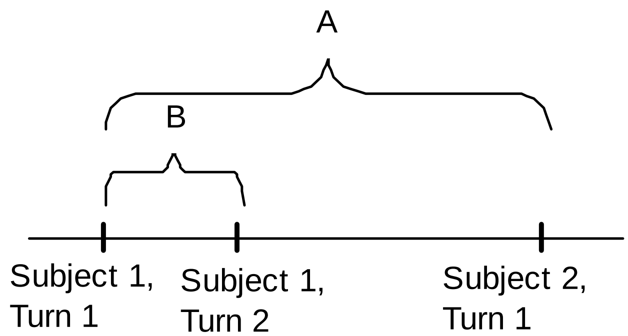


Figure 2: Adjustment rate

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 reflects a unit-free measure of weighting on 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 ("FRNO") 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 which is consistent with our intuition of the game:

- Objective confidence is monotonically related to adjustment rates *in both treatments*; that is, higher objective probability corresponds to

[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.

lower adjustments. This suggests that subjects incorporated 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 with the effect of overconfidence
- Subjects adjust *uniformly less* (in both treatments) than predicted by the FRNO model, suggesting that players discount others' actions. This is consistent with strategic response to others' errors.

4.2 Market-level results

4.2.1 Measures

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

- Volume index represents the extent to which players diverge in their estimates and is denoted by: $Vol_{r,t} = |a_{i,r,t} - a_{j,r,t}|$.[30]
- Price index represents the average estimates and is denoted by: $P_{r,t} = \frac{a_{i,r,t}+a_{j,r,t}}{2}$.
- Price error index represents the distance between the average estimate and the fundamental value and is denoted by: $PE_{r,t} = |v_r - P_{r,t}|$.
- Price volatility index represents the rate at which the average of estimates change across *turns* and is denoted by:

$$PStd_r = \left[\frac{\sum_{t=1}^{4} \left(P_{r,t} - \frac{1}{4} \sum_{t=1}^{4} P_{r,t} \right)^2}{3} \right]^{1/2}.$$

4.2.2 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

[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.

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 2 provides average levels (across market instances and rounds) of volume and price indexed for both treatments while table 3 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.

4.2.3 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 FRNO 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 figure 5).[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 4 for a formal cross-treatment significance test). Specifically, we find

[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 that predicted by the FRNO in turn 1. This is to be expected since the precision of information should not affect turn 1 estimates.

that about 40% of *excess* volume index is attributable to strategic response to errors, while the remaining is attributable to overconfidence. If one looks at price error or price volatility indexes, similar results are obtained (see table 2).[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.

4.2.4 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 9). 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 5 summarizes the results. First, there is no indication of unconditional return predictability in either of the treatments. Second, we find negative serial autocrrelation 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

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

the existence of overconfidence, they would fail to properly offset the "negative externality" brought about by the poorly informed yet overconfident players.

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. We describe the model in the next section.

5 Econometric Model

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 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 a fellow players' errors both result in increasing the 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 rational 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 dynamic 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 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:

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

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 $(t-1)$), and $e_{i,t} \sim N(0, \sigma_t^2)$ is turn t realized error.

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

[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.

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

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.

5.1 Optimal Actions

In this part we derive expressions for best-responses and actions for a generic player i during turns 1-3. 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 40 for details) that:

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

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

[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)].

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 report, $a_{2,1}$ (recall that in the fully-rational model $E(s_2|a_{2,1}) = a_{2,1}$), for the following parameter values: $s_1 = 500, Y = 30, \sigma_1 = 5$.

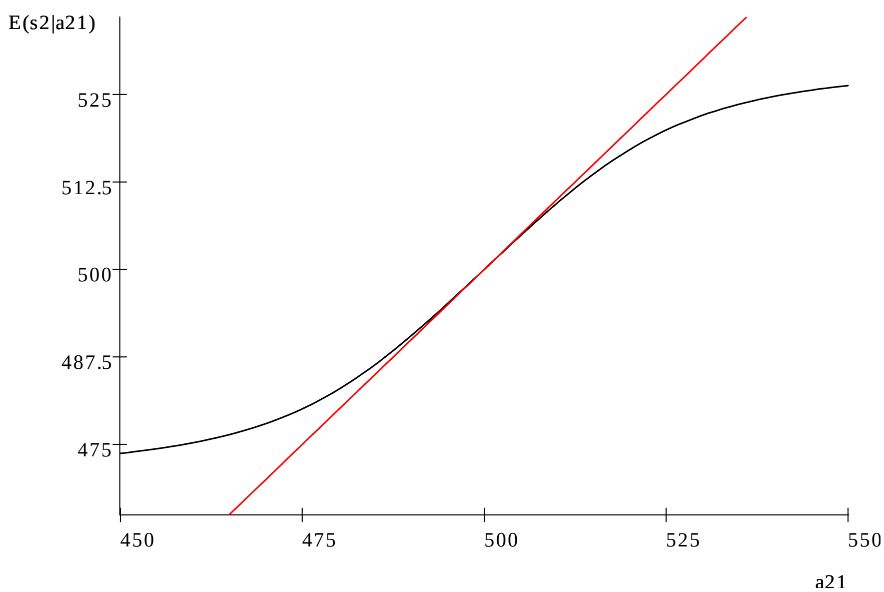


Figure 3: Turn 2 best response adjustment term effect - numerical example

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.

In turn 3, it is possible for each player to make inferences about the subjective probability of the player with whom they are paired, in addition to improving the inference about the other player's signal. Therefore, $a^*_{i,3} = \int\limits_0^1 \int\limits_{s_i-Y}^{s_i+Y} \Pr(s_j, q_j|I_{i3})(Ind(q_1 \geq q_2)s_i + Ind(q_1 < q_2)s_j)ds_j dq_j$.

Defining:

$f_{1,2} = \frac{2\sigma_1^2(\phi(a_{2,1};s_1-Y,\sigma_1)-\phi(a_{2,1};s_1+Y,\sigma_1))}{\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{2,1})\right)-\mathrm{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{2,1}-s_1-Y)\right)}$

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

$C_{1,3} = \frac{\phi(a_{2,1}-s_2;0,\sigma_1)\phi^2(a_{2,2}-q_2s_2-(1-q_2)(a_{11}+f_{2,2});0,\sigma_2)}{\int\limits_{s_1-Y}^{s_1+Y}(\phi(a_{2,1}-s_2;0,\sigma_1)\phi(a_{2,2}-q_2s_2-(1-q_2)(a_{1,1}+f_{2,2});0,\sigma_2))ds_2}$

$C_{2,3} = \frac{\phi(a_{1,1}-s_1;0,\sigma_1)\phi^2(a_{1,2}-q_1s_1-(1-q_1)(a_{2,1}+f_{1,2});0,\sigma_2)}{\int\limits_{s_2-Y}^{s_2+Y}(\phi(a_{1,1}-s_1;0,\sigma_1)\phi(a_1-q_1s_1-(1-q_1)(a_{2,1}+f_{1,2});0,\sigma_2))ds_1}$

We obtain that (see page 41 for details):

$$a_{1,3} = \int\limits_0^1 \int\limits_{s_1-Y}^{s_1+Y} \left(\frac{C_{1,3}}{\int\limits_0^1 (C_{1,3})\, dq_2} (Ind(q_1 \geq q_2)s_1 + Ind(q_1 < q_2)s_2) \right) ds_2 dq_2 + e_{1,3} \tag{10}$$

$$a_{2,3} = \int\limits_0^1 \int\limits_{s_2-Y}^{s_2+Y} \left(\frac{C_{2,3}}{\int\limits_0^1 (C_{2,3})\, dq_1} (Ind(q_2 \geq q_1)s_2 + Ind(q_2 < q_1)s_1) \right) ds_1 dq_1 + e_{2,3} \tag{11}$$

Due to the computational load involved (compared to the little marginal benefit it is likely to bring) we did not proceed to calculate turn 4 optimal actions.

5.2 Maximum Likelihood Estimation

In order to understand the estimation procedure we start by clarifying the somewhat complex structure of the data collected in this experiment. The observations are organized in the following way:

- Level

	Market			**Individual**
Turn	Price	Price error	Volume	Reports
Round	Price volatility	Return autocorrelation		Subjective probability

  - Treatment (BLT and OCT)
  - Session (5 BLT and 7 OCT)

- Time

  - Round (10 in each session)
    - Turn (4 in each round)

- Unit of analysis

  - Market (3 per turn)
  - Subjects (2 per market)

The data recorded is organized as follows:

We intend to estimate subjects' *individual* beliefs across rounds, their optimal action and the variance of errors across turns. The specification of our model assumes that :

1. Each error is conditionally independent of all other errors.
2. Subjective beliefs are constant across pairs of subsequent rounds. That is, subjective beliefs in rounds 1 and 2, 3 and 4, 5 and 6, 7 and 8, 9 and 10 are the same.
3. Within each treatment, for a given stage of the game, the distribution of errors is constant (across sessions).

While assumption (1) follows from the specification of the econometric model, assumption (2) should be clarified. In principle, since each subject holds beliefs that may evolve over time, we should allow the estimates to vary across rounds.[37] However, doing so would entail estimating $(N+3)$ parameters using only $3N$ observations (where $N = 30$ in the BLT and

[37] We defer explicitly modeling the updating process (due to its complexity) for future work. In the perfectly rational model, though, as long as subjects' subjective beliefs in round 1 are not the same they should update the probability to 0 or 1 at end of that round.

$N = 42$ in the OCT). Instead, we group rounds together so that we estimate $(N + 3)$ parameters using $6N$ observations, allowing for a more favorable ratio of data to degrees of freedom. As a result, we will estimate the model independently for 5×2 subsets of the data (5 groups of pair rounds and 2 treatments).

This leads naturally to the following likelihood function:

$$L = \prod_{r=r_0}^{r_1} \prod_{i=1}^{N} \phi\left(a_{r,1,i} - a^*_{r,1,i}; 0, \sigma_{t,1}\right) \phi(a_{r,2,i} - a^*_{r,2,i}(\sigma_{r,1}, I_{r,i,2}, \widetilde{q}_{r,i}); 0, \sigma_{r,2})$$
$$\phi(a_{r,3,i} - a^*_{r,3,i}(\sigma_{r,1}, \sigma_{r,2}, I_{r,i,3}, \widetilde{q}_{r,i}); 0, \sigma_{r,3})$$

Where i indexes subjects (N is the number of subjects across sessions), r indexes the round (with $r_0(r_1)$ being the first (last) round included in the estimation) and the numbers $1, 2, 3$ denote the turn. For example, $a_{r,1,i}$ refers to action submitted in round r, turn 1, by player i. We fit this model using observed actions during the first three turns (within each round) while separating the data into treatments and into subsets of rounds, as explained below.

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.[38] Since these procedures find local minima, results may be sensitive to starting points. To avoid biasing our estimations, we specify a starting point for subjective probabilities of .5 for all subjects 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.

6 Model estimation results

Our results can be organized around three main areas:

1. Test of the NABB model against the two main alternative models – Rational Expectations (RE) or Private Information (PI) – suggests

[38] Matlab code is available upon request.

that the NABB model fits the observed behavior better than the other two.

2. Miscalibration

   - We find strong evidence to suggest that estimated subjective probabilities, which are measured indirectly, represent quantities that can indeed be thought of as players' beliefs.
   - Comparing miscalibration across treatments we find that:
     - Aggregate overconfidence in the OCT (of about 11%) but not in the BLT; that is, we confirm the absence of overconfidence in the BLT and its existence in the OCT.
     - The distribution of miscalibration across subjects is different. Specifically, subjects that have low confidence are overconfident while subjects that exhibit high confidence are somewhat underconfident.
     - In the OCT, subjects exhibit "stickier" beliefs: while subjects' beliefs in the BLT seemed to become more accurate in the course of the session. Subjects in the overconfidence treatment exhibit persistent levels of miscalibration. This seems to suggest that the source of miscalibration influences not only the prior distribution but also the efficiency with which it is updated as a result of market interaction.

3. Market outcomes:

   - Using individual level estimates of miscalibration we construct two measures of aggregate miscalibration (mean and dispersion) within each market instance. We find that:
     - Controlling for treatment effect, there is positive association between average miscalibration and volume / price error indexes, which is consistent with our intuition.
     - Interestingly, *dispersion* of miscalibration has a significant affect on these measures.
     - Price volatility index seems to be uncorrelated with these miscalibration measures.

6.1 Measures

- Miscalibration
  Miscalibration, denoted by $MC_{i,r} = \widetilde{q}_{i,r} - q_{i,r}$ where $\widetilde{q}_{i,r}(q_{i,r})$, is the subjective (objective) probability of being perfectly informed. It captures the degree to which individual beliefs deviate from their objective level (recall that we interpret positive miscalibration as representing overconfidence and negative miscalibration as representing underconfidence).

- If miscalibration is heterogeneous then aggregate measures are needed. We suggest two natural candidates:

  - Mean miscalibration within a market instance: $MMC_r = \frac{MC_{i,r}+MC_{j,r}}{2}$
  - Dispersion of miscalibration within a market instance: $DMC_r = |MC_{i,r} - MC_{j,r}|$

6.2 Testing the NABB Model

In this section we show that the NABB model performs better than either the Private Information (PI) or the perfectly Rational Expectations (RE) models.[39] We compare these models' implications on individual behavior The PI model implies that each subject relies only on their information, completely disregarding the actions of others. The RE 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 report submitted by player j in the previous turn.

[39] Notice that the QRE model nests both the RE and the PI models: if players are perfectly calibrated and make no errors, we obtain the RE; if subjects are all extremely overconfident and/or make submit extremely noisy actions we would obtain the PI. The question we address is how much better does the QRE model perform relative to either of the other models.

We conduct the following procedure: for each of the models, we generate predicted individual decisions and calculate the absolute deviations between those and observed decisions. We record the absolute value of errors and use a non-parametric procedure to test the null that median errors are the same across models. Table 6 reports median errors and p values when comparing NABB with PI and RE, one at a time.[40,41]

We see that the NABB performs much better than PI across the turns. The NABB model performs (weakly) better than the RE model in all rounds, although statistical significance can only be established for turn 2. This may be because both models (NABB and RE) provide predictions that are closer together over the turns.

6.3 Miscalibration

6.3.1 Estimated probabilities as subjective beliefs

To calculate miscalibration, we need to obtain both subjective and objective probabilities for each subject. While the former come from the econometric model estimations, the latter is calculated in the following way: using outcomes from the initial phase, we determine the probability that each of the subjects is classified as perfectly informed, i.e. is ranked in the upper half of the group. For example, the probability that a subject obtaining 4 in the die throw will be ranked above average among a group of 6 others is 0.6553. Likewise, we bootstrap from the observed scores on the SAT quiz to calculate objective probabilities.[42]

In order for our estimates to represent miscalibration, we need to ensure that estimated probabilities represent *beliefs*. We suggest two ways to do that, one is cross sectional and the other is intertemporal. First, we test whether the correlation between objective and subjective probabilities is different from zero at each round of the game. Second, we ask whether *across rounds* individual subjective probabilities exhibit persistence, as we would expect if those were beliefs.

Performing a standard correlation test (using Fisher transformation) we find that the correlation between objective and estimated probabilities (ranging from 0.46 to 0.77 in the BLT and 0.17 to 0.36 in the OCT) is

[40] Note that in turn 1 all three models suggest that one would simply submit their signal and therefore no comparison can be made.

[41] Since we have not calculated the theoretical QRE prices for turn 4 we do not include comparison of that turn.

[42] The correlation between subjective probabilities and realized types (0 imperfectly informed, 1 perfectly informed) is the same across treatment and equal to 0.8.

		Subjective belief in round $\mathbf{r+1}$	
		Above	Below
Subjective belief in round r	Above	A	B
	Below	C	D

positive and statistically different from zero for all but one round in both treatments, see table 7.

To provide additional support for the notion that these quantities are informative, we form a test of the intertemporal behavior of individual estimated beliefs. Formally, we test the null that subjective probabilities are *iid* draws from an unspecified distribution and thus do not exhibit any persistence. Using a modification of the binomial test we match the estimated beliefs across subsequent rounds (1&2 vs. 3&4, 3&4 vs. 5&6, etc.) and count the number of times subjects display above (below) median beliefs in a given round but then below (above) median beliefs in the next round, i.e., switch.[43] To visualize the procedure, consider assigning each subject into one of these groups:

The rank is being used since no a priori probability can be assigned to the various *absolute* ranges of outcomes. Thus, the test compares the number of $A+D$ cases to the number of $B+C$ cases. Formally, under the null, $\frac{A+D}{A+B+C+D} = \frac{B+C}{A+B+C+D} = \frac{1}{2}$, or:

$$\Pr(\widetilde{q}_{r+1,i} > Med(\widetilde{q}_{r+1})|\widetilde{q}_{r,i} > Med(\widetilde{q}_r)) + \Pr(\widetilde{q}_{r+1,i} < Med(\widetilde{q}_{r+1})|\widetilde{q}_{r,i} < Med(\widetilde{q}_r)) =$$
$$\Pr(\widetilde{q}_{r+1,i} < Med(\widetilde{q}_{r+1})|\widetilde{q}_{r,i} > Med(\widetilde{q}_r)) + \Pr(\widetilde{q}_{r+1,i} > Med(\widetilde{q}_{r+1})|\widetilde{q}_{r,i} < Med(\widetilde{q}_r)) = .5$$

where $Med(\widetilde{q}_r)$ is the median subjective probability in round r. That is, under the null, the probability that subject i would exhibit above median beliefs in round $r+1$ is .5 and *does not depend* on her beliefs in round r. We administer the test separately for each treatment by constructing the following measure:

$$S = \sum_{i=1}^{N} \left(\Pr(\widetilde{q}_{r+1,i} < Med(\widetilde{q}_{r+1})|\widetilde{q}_{r,i} > Med(\widetilde{q}_r)) + \Pr(\widetilde{q}_{r+1,i} > Med(\widetilde{q}_{r+1})|\widetilde{q}_{r,i} < Med(\widetilde{q}_r))\right)$$

[43] We use separate medians for each series.

Table 8 reports the results; we can reject the null that estimated beliefs are random in favor of the alternative that they exhibit (at least some) persistence at high significance levels (above 95%) for all cases.

6.3.2 Miscalibration levels

Having shown that estimated probabilities represent subjective beliefs, we turn to discuss miscalibration. We first seek to establish that the conjectured treatment effect induces (overall) overconfidence in the OCT but not the BLT. Table 9 reports mean miscalibration (across subjects) and tests of the null that it is equal to zero, using both parametric (after verifying normality with a Jarque-Bera test) and non-parametric (Wilcoxon sign-rank test) procedures.

Both tests point to the same conclusion: average miscalibration is close to zero in the BLT but not in the OCT. In particular, there is clear indication of average overconfidence in the OCT, when we average across all rounds, or in 6 out of the 10 separate rounds (when these are considered separately).

Additional support that a systematic treatment effect is present can be obtained by comparing the overall *distribution* of miscalibration across treatments, depicted in figure 6. We see that miscalibration is more dispersed in the OCT. A 2-way analysis of variance (ANOVA), in which one dimension is the stage of the game (round) and the other is the treatment, confirms that (see table 10); treatment has a strong effect on the variability of miscalibration ($p = .03$), while round does not. In the next section we characterize the *way* in which variability is different across treatments.

6.3.3 Distribution of individual miscalibration

Further characterization of the difference between the two treatments is achieved by relating subjective and objective probabilities. These plots are very common in the miscalibration literature [see Alpert and Raiffa (1977) for a review]. In these studies, subjects are presented with a set of two-alternative general-knowledge questions and are asked to assign a probability that the answer they picked is correct.[44] Then, the frequency of correct answers within each group of stated confidence level ($0.5 - 0.6, 0.6 - 0.7$, etc.) is calculated. If subjects are well-calibrated, frequencies should match subjective probabilities and thus lie on a 45 degree line; observations that land below (above) the line represent overconfidence (underconfidence).

[44] We are presenting here one common variant of many; this is not meant to be a comprehensive survey of this vast literature.

We construct a similar plot while averaging across objective *probabilities* instead of frequencies (see figure 7). First, we find striking resemblance to plots generated by previous studies mentioned above. The fact that these studies have used direct elicitation of beliefs, which is very different from the approach taken here, yet arrived at overall similar results provides further validation of the subjective probability results generated by our econometric model. Next, we observe that in both treatments low (high) confidence levels are associated with underconfidence (overconfidence). For example, subjects holding subjective beliefs of .9 have objective probability of around .6 in the OCT and around .7 in the BLT. However, the observations in the BLT are closer to the 45 degree line than in the OCT.

We can test this difference by specifying a simple linear relationship between miscalibration and the subjective probabilities: $\widetilde{q}_i = \delta + \gamma q_i + e_i$ (which can also be written as $MC_i = \delta + (\gamma - 1)q_i$).[45] Results of estimating this relationship separately for both treatments are in table 11. Under the null that subjects are well-calibrated we would expected to find $\delta = 0$ and $\gamma = 1$. For both treatments we can reject the null. Notice that rejecting the null does not necessarily indicate aggregate overconfidence. *Unconditionally*, subjects are on average overconfident in the OCT but not in the BLT. *Conditionally*, in both treatments, subjects that have less precise information are overconfident while those who have more precise information are somewhat underconfident. However, the conditional relation is not the same across treatments, as suggested by the F-test value ($F = 22.12$). In particular, the estimated slope in the BLT is closer to 1 than in the OCT suggesting that there is tighter relationship, in that treatment, between subjective and objective probabilities. These results are important in explaining the cross-treatment market-level evidence we present in a following section.

6.3.4 Learning

We now turn to answer the questions: does the treatment affect the way subjects update their beliefs. This is a particularly important question since much of the focus of economics and finance is on equilibrium results. If individuals start with too-high average beliefs but update these beliefs in the equilibration process, through market interaction, it may be that the eventual consequences of miscalibration are marginal. However, as suggested by Gervais and Odean (2001) this need not happen: traders may become *more* overconfident in the process of trading.

[45] Recall that $OC_i \equiv \widetilde{q}_i - q_i$ and rearrange the original expression.

The results we present below suggest otherwise - in the OCT, subjects' beliefs seem to exhibit higher persistence than in the base-line treatment. To show that, we start by looking at how systematic both positive and negative miscalibration is in the two treatments. To test that, we average estimated miscalibration in early rounds $(1-4)$ and late rounds $(7-10)$ and count the number of cases where subjects 'flipped' from being over(under) confident early on to being under(over) confident later on in the session (see table 12 for summary of the results).

First, miscalibration is generally systematic - only a quarter of the subjects switch from being overconfident to being underconfident. Second, in the BLT persistence is not dependent on the initial sign of miscalibration (positive or negative) while in the OCT subjects that are initially overconfident are *less* likely to switch than those that are initially underconfident.

Next, we examine the extent to which changes of beliefs are in 'the right direction'. That is, whether subjects that early on are very miscalibrated become less so at the close of the session. To do that, we divide subjects into one of three broad categories (according to initial miscalibration): underconfident (miscalibration between -1 and $-\frac{1}{3}$), well calibrated (miscalibration between $-\frac{1}{3}$ and $\frac{1}{3}$) and overconfident (miscalibration between $\frac{1}{3}$ and 1). Then we calculate the fraction of cases in which improvement / no change / deterioration was detected (see table 13 for details).

Comparing the results across treatments we see that the relative fraction of cases where improvement was detected is consistently (for all initial classifications) higher in the BLT than in the OCT. The most important difference lies in the initially overconfident group of subjects: while 83% of them became less overconfident over-time in the BLT, only 67% of the same group became less overconfident in the OCT. Moreover, 27% of the same group became even more overconfident in the OCT!

Formally testing the hypothesis that learning is not affected by treatment is done by regressing beginning-of-session miscalibration, for each subject, on end-of-session miscalibration. A flatter regression line slope would represent greater improvement in calibration; if by the end of the session all subjects are well-calibrated the line should have a slope (and intercept) of zero.

Result of estimating $MC_i^{6-10} = a + b_1 MC_i^{1-4} + b_2(Dummy_i * MC_i^{1-4}) + e_i$, where the variable $Dummy_i$ takes on a value of 0 (1) in the BLT (OCT), are reported in table 14. These results indicate that learning is present in *both* treatments – the coefficient on initial miscalibration is statistically different from 1, indicating that subjects become less miscalibrated in the course of the session. At the same time, treatment influences the *rate* of calibration: the coefficient on the interaction between slope and treatment is both positive and significantly different from zero (albeit at 6% significance level). Taken together, we provide strong support to the notion that treatment has an affect on the updating process of beliefs as well as on their initial level.

6.4 Aggregation of miscalibration

Given that miscalibration is not distributed evenly across subjects, we explore whether *dispersion* of miscalibration, in addition to its average level, can be linked to market outcomes (e.g., volume and price error indexes). To test that, we construct two *aggregate* measures of miscalibration: average miscalibration (denoted by MMC) and the dispersion of miscalibration (denoted by DMC) and calculate them for each market instance. Recall that since subjects are matched randomly at the beginning of each round, we obtain a wide cross-distribution of individual miscalibration (see figure 8). We regress (using robust regressions) the treatment dummy variable, mean miscalibration and dispersion of miscalibration on average volume index (turns 2 and 3), price error index (turns 2 and 3) and price standard deviation index:

$$Vol_i = a + b_1 Dummy_i + b_2 MMC_i + b_3 DMC_i + e_i$$

$$PE_i = a + b_1 Dummy_i + b_2 MMC_i + b_3 DMC_i + e_i$$

$$PStdv_i = a + b_1 Dummy_i + b_2 MMC_i + b_3 DMC_i + e_i$$

Findings are summarized in table 15.

Consistent with our intuition, we find that after controlling for treatment effect, higher average level of miscalibration increases volume and price indexes levels. We also find that *dispersion* of miscalibration has a strong effect on these quantities. However, the circumstances under which these two arise may be different; it is possible for a group of traders to have no average overconfidence but still have high differences in their overconfidence level. This would result in deviation from the fully-rational predictions in the direction suggested by overconfidence based model. Studying the relation between mean miscalibration, dispersion of miscalibration, and price volatility index we find that these measures do not show up significantly.

7 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 allows controls for the presence of overconfidence, and on econometric model that nests both biases. As a result, we are able to estimate subjects' beliefs throughout the course of the experiment and quantify the magnitude of errors.

At the individual level, we show that the initial miscalibration is on average positive in the overconfidence treatment and zero in the baseline treatment, serving to establish a causal relation between market outcomes and individual attributes. We use the dynamic feature of subjects' confidence to demonstrate that the rate of learning (as expressed through the update of beliefs about one's own type) is slower in the overconfidence treatment. Further, we illustrate that miscalibration is not spread uniformly, suggesting a role for miscalibration dispersion in explaining some of the results.

At the market level, we find that subjects strategically respond to others' mistakes and that this feature generates a pattern of volume, informational efficiency and volatility index levels similar to those predicted by models based on overconfidence. Nonetheless, most canonical predictions linking investors' overconfidence to markets is borne out: we find a higher volume and lower informational efficiency index levels in the overconfidence treatment, but similar price volatility index levels across treatments.

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 return index. Preliminary results suggest intriguing dynamics, in the spirit of Llorente, Michaely, Saar and Wang (2002). While we have documented that treatment effects on the updating of beliefs, further exploration into the role market interaction plays in learning is likely to be fruitful. 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. Last, since our game does not depend on a large number of participants, one can stress-test this setup to see how well it performs as an information gathering mechanism in 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", 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

[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

[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] Kahnman 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] 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

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

[38] 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

[39] 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

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

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

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

[43] 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

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

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

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

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

[48] 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

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

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

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

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

8 Appendix

8.1 Proofs

Proof of proposition 9 For turns 1 and 3 the proof is obvious. Using the same notation as before we let $q_{-ij} = q_i - q_j$, $MC_{-ij} = MC_i - MC_j$ and noting that $1 \leq q_i - q_j \leq 0$, $1 \leq q_i + MC_i - q_j - MC_j \leq -1$ we get that:
$E(PE_2) = E(\frac{e}{2})E\left(|1 - q_i + q_j|\right) = E(\frac{e}{2})E\left(\left|1 - q_{-ij}\right|\right)$
$E(PE_2^{MC}) = E(\frac{e}{2})E\left(|1 - q_{-ij} - MC_{-ij}|\right) = E(\frac{e}{2})E\left(|1 - q_{-ij}(1+\gamma)|\right)$
$E(PE_2) < E(PE_2^{MC})$ if $E\left(|1 - q_{-ij}(1+\gamma)|\right) > E\left(\left|1 - q_{-ij}\right|\right)$. Since $0 < 1+\gamma < 1$ and $0 < q_{-ij} < 1$, both arguments are positive, therefore, $E\left(|1 - q_{-ij}(1+\gamma)|\right) - E\left(\left|1 - q_{-ij}\right|\right) = E\left(1 - q_{-ij}(1+\gamma)\right) - E\left(1 - q_{-ij}\right) = q_{-ij}(-\gamma) > 0$.
Notice that we obtain that price errors increase in γ which captures the dispersion of overconfidence in the population.

Derivation of $\mathbf{a}_{1,2}, \mathbf{a}_{2,3}$ 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:
- $\Pr(a_{2,1}|s_2, s_1) = \Pr(a_{2,1}|s_2) = \phi(a_{21} - s_2; 0, \sigma_1)$
- $\Pr(s_2|s_1) = \frac{1}{2Y}$

$$\Pr(a_{21}|s_1) = \int\limits_{s_1-Y}^{s_1+Y} \frac{1}{2Y}\phi(a_{21} - s_2; 0, \sigma_1) ds_2 =$$
$$\frac{\sqrt{4\pi}\operatorname{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}Y\sqrt{2} - \frac{1}{2}s_1\sqrt{2} + \frac{1}{2}a_{21}\sqrt{2}\right)\right)}{\frac{1}{8\sqrt{\pi}Y}}$$
$$-\frac{\sqrt{4\pi}\operatorname{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}a_{21}\sqrt{2} - \frac{1}{2}s_1\sqrt{2} - \frac{1}{2}Y\sqrt{2}\right)\right)}{\frac{1}{8\sqrt{\pi}Y}} =$$
$$\frac{1}{4Y}\left(\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}\left(Y - s_1 + a_{21}\right)\right) - \operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}\left(a_{21} - s_1 - Y\right)\right)\right)$$
Collecting these terms we obtain that:
$$E(s_2|a_{21}, s_1) = \int_{s_1-Y}^{s_1+Y} s_2 \Pr(s_2|a_{21}, s_1) ds_2 =$$

$$\int_{s_1-Y}^{s_1+Y} s_2 \frac{\phi(a_{21}-s_2;0,\sigma_1)\frac{1}{2Y}}{\frac{1}{4Y}\left(\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)\right)} ds_2 =$$

$$-\frac{\sqrt{2}\left(\frac{1}{2}\sigma_1\sqrt{2}e^{-\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\left(-\sigma_1\sqrt{2}\exp\left(\frac{a_{21}}{\sigma_1^2}(Y+s_1)-\frac{1}{2\sigma_1^2}(Y+s_1)^2\right)-\sqrt{\pi}a_{21}e^{\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\operatorname{erf}\left(\frac{1}{2}\frac{a_{21}}{\sigma_1}\sqrt{2}-\frac{1}{2\sigma_1}(Y+s_1)\sqrt{2}\right)\right)-\right)}{\sqrt{\pi}\sigma_1\operatorname{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}a_{21}\sqrt{2}-\frac{1}{2}s_1\sqrt{2}-\frac{1}{2}Y\sqrt{2}\right)\right)-\sqrt{\pi}\sigma_1\operatorname{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}Y\sqrt{2}-\frac{1}{2}s_1\sqrt{2}+\frac{1}{2}a_{21}\sqrt{2}\right)\right)}+$$

$$\frac{\sqrt{2}\left(\frac{1}{2}\sigma_1\sqrt{2}e^{-\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\left(-\sigma_1\sqrt{2}\exp\left(\frac{a_{21}}{\sigma_1^2}(s_1-Y)-\frac{1}{2\sigma_1^2}(s_1-Y)^2\right)-\sqrt{\pi}a_{21}e^{\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\operatorname{erf}\left(\frac{1}{2}\frac{a_{21}}{\sigma_1}\sqrt{2}-\frac{1}{2\sigma_1}\sqrt{2}(s_1-Y)\right)\right)\right)}{\sqrt{\pi}\sigma_1\operatorname{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}a_{21}\sqrt{2}-\frac{1}{2}s_1\sqrt{2}-\frac{1}{2}Y\sqrt{2}\right)\right)-\sqrt{\pi}\sigma_1\operatorname{erf}\left(\frac{1}{\sigma_1}\left(\frac{1}{2}Y\sqrt{2}-\frac{1}{2}s_1\sqrt{2}+\frac{1}{2}a_{21}\sqrt{2}\right)\right)} =$$

$$a_{21}+\frac{e^{-\frac{1}{2}\frac{a_{21}^2}{\sigma_1^2}}\left(\sigma_1\sqrt{2}\left(\exp\left(\frac{a_{21}}{\sigma_1^2}(s_1-Y)-\frac{1}{2\sigma_1^2}(s_1-Y)^2\right)-\exp\left(\frac{a_{21}}{\sigma_1^2}(Y+s_1)-\frac{1}{2\sigma_1^2}(Y+s_1)^2\right)\right)\right)}{-\sqrt{\pi}\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)+\sqrt{\pi}\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)} =$$

$$a_{21} + \frac{2\sigma_1^2(\phi(a_{21};s_1-Y,\sigma_1)-\phi(a_{21};s_1+Y,\sigma_1))}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}$$

Derivation of $\mathbf{C}_{1,3}, \mathbf{C}_{2,3}$ Calculating the terms we get that:

$\Pr(s_j, q_j|I_{it}) = \Pr(q_j|s_j, I_{it}) \Pr(s_j|I)$

$\Pr(q_j|s_j, I_{it}) = \Pr(q_2|s_2, a_{22}, a_{21}, a_{11}, a_{12}, s_1)$

Since $q_2 = \frac{a_{22}-a_{11}-f_{22}-e_{22}}{s_2-a_{11}-f_{22}}$, we get that $\Pr(q_2|s_2, a_{22}, a_{21}, a_{11}, a_{12}, s_1) = \phi(a_{22} - s_2q_2 - (1-q_2)(a_{11}+f_{22}); 0, \sigma_2)$

$\Pr(s_j|I) = \Pr(s_2|a_{22}, a_{21}, a_{11}, a_{12}, s_1) = \frac{\Pr(a_{22}|s_2,a_{21},s_1,a_{11},a_{12}) \Pr(s_2|a_{21},s_1,a_{11},a_{12})}{\Pr(a_{22}|a_{21},s_1,a_{11},a_{12})}$

Since

$a_{12} = a_{12}(s_1, a_{21}, e_{12})$

we condition on s_1 and a_{21}

e_{12} is uncorrelated with any other random variable (conditional on s_1 and a_{21})

We can omit a_{12} from the information set.

Also, since $a_{21} = s_2 + e_{21}$ and e_{21} is mean zero, uncorrelated r.v:

$\Pr(a_{22}|s_2, a_{21}, s_1, a_{11}) = \Pr(a_{22}|s_2, s_1, a_{11})$.

$\Pr(s_2|a_{21}, s_1, a_{11}) = \Pr(s_2|a_{21}, s_1)$

Thus:

$\Pr(s_2|a_{22}, a_{21}, a_{11}, s_1) = \frac{\Pr(a_{22}|s_2,s_1,a_{11}) \Pr(s_2|a_{21},s_1)}{\Pr(a_{22}|a_{21},s_1,a_{11})}$

$f_{22} = \frac{2\sigma_1^2(\phi(a_{11};s_2-Y,\sigma_1)-\phi(a_{11};s_2+Y,\sigma_1))}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_2+a_{11})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{11}-s_2-Y)\right)}$

$f_{12} = \frac{2\sigma_1^2(\phi(a_{21};s_1-Y,\sigma_1)-\phi(a_{21};s_1+Y,\sigma_1))}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}$

$\Pr(a_{22}|s_2, s_1, a_{11}) = \Pr(a_{22}|s_2, a_{11}) = \phi(a_{22} - q_2s_2 - (1-q_2)(a_{11}+f_{22}); 0, \sigma_2)$

From previous part with obtain that:

$\Pr(s_2|a_{21}, s_1) = \frac{2Y\phi(a_{21}-s_2;0,\sigma_1)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}$

$\Pr(a_{22}|a_{21}, s_1, a_{11}) =$
$\int\limits_{s_1-Y}^{s_1+Y} \frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)} ds_2$
Thus:
$\Pr(s_2|a_{22}, a_{21}, a_{11}, s_1) =$
$$\frac{\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}}{\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\right)ds_2}$$
Recall that $\Pr(q_2|a_{22}, a_{21}, a_{11}, s_1) = \phi(a_{22} - s_2q_2 - (1 - q_2)(a_{11} + f_{22}); 0, \sigma_2)$, therefore:
$\Pr(s_2, q_2|a_{22}, a_{21}, a_{11}, a_{12}, s_1) =$
$$\frac{\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\phi(a_{22}-s_2q_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\right)ds_2} =$$
$$\frac{\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi^2(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}}{\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\right)ds_2} \equiv B_{13}$$
Thus,
$$a_{13}^* = \int\limits_0^1\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{B_{13}}{\int\limits_0^1\int\limits_{s_1-Y}^{s_1+Y}(B_{13})ds_2dq_2}\left(Ind(q_1 \geq q_2)s_1 + Ind(q_1 < q_2)s_2\right)\right)ds_2dq_2.$$
Rearranging the terms we get:
$$\int\limits_0^1\int\limits_{s_1-Y}^{s_1+Y}\frac{\frac{\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi^2(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}}{\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\right)ds_2}}{\int\limits_0^1\left(\frac{\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi^2(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\right)ds_2}{\int\limits_{s_1-Y}^{s_1+Y}\left(\frac{2Y\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(Y-s_1+a_{21})\right)-\operatorname{erf}\left(\frac{\sqrt{2}}{2\sigma_1}(a_{21}-s_1-Y)\right)}\right)ds_2}\right)dq_2}$$
$(Ind(q_1 \geq q_2)s_1 + Ind(q_1 < q_2)s_2)\, ds_2dq_2$
Letting:
$$C_{13} = \frac{\phi(a_{21}-s_2;0,\sigma_1)\phi^2(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2)}{\int\limits_{s_1-Y}^{s_1+Y}(\phi(a_{21}-s_2;0,\sigma_1)\phi(a_{22}-q_2s_2-(1-q_2)(a_{11}+f_{22});0,\sigma_2))ds_2}$$

$$C_{23} = \frac{\phi(a_{11}-s_1;0,\sigma_1)\phi^2(a_{12}-q_1 s_1-(1-q_1)(a_{21}+f_{12});0,\sigma_2)}{\int\limits_{s_2-Y}^{s_2+Y}(\phi(a_{11}-s_1;0,\sigma_1)\phi(a_1-q_1 s_1-(1-q_1)(a_{21}+f_{12});0,\sigma_2))ds_1}$$

We get that

$$a_{1,3} = \int\limits_0^1 \int\limits_{s_1-Y}^{s_1+Y} \left(\frac{C_{13}}{\int\limits_0^1 (C_{13})\, dq_2} (Ind(q_1 \geq q_2)s_1 + Ind(q_1 < q_2)s_2) \right) ds_2 dq_2 + e_{1,3}$$

$$a_{2,3} = \int\limits_0^1 \int\limits_{s_2-Y}^{s_2+Y} \left(\frac{C_{23}}{\int\limits_0^1 (C_{13})\, dq_2} (Ind(q_2 \geq q_1)s_2 + Ind(q_2 < q_1)s_1) \right) ds_1 dq_1 + e_{1,3} \quad (12)$$

8.2 Instructions

Below are experiment instructions with text that is being alternated across treatments appearing in brackets ('[]' for the OCT and '{}' for the BLT).

General

This is an experiment in economic decision making. Various research foundations have provided funds for it.

If you follow the instructions carefully and make good decisions, you may earn a considerable amount of money which will be paid to you in private and in cash at the end of the experiment.

There are 6 people in the room who are participating in this experiment. It is important that you do not talk to any of them until the experiment is over. The experiment will consist of 10 independent and identical rounds, in each of which you can earn points.

At the end of the experiment you will be paid an amount based on your total point earnings, which will depend on your decisions and on chance. Points will be converted to cash using an exchange rate of 1 point = 1 cent. Notice that the more points you earn, the more cash you will receive at the end of the experiment.

Initial

In this experiment you will receive either a perfect or imperfect signal (clue) about the value of a randomly drawn number and will be asked to submit your guesses about what that number is in a sequence of turns.

At the beginning of the experiment all participants will be asked to take the same quiz, consisting of 20 multiple-choice SAT questions, in 4 minutes.

[Upon completing the quiz, your score will be calculated based on the number of correct answers; the time it took you to complete the questioner will be used to resolve ties. Scores from all participants will be gathered to determine your relative position. Half of the people in this room, with the highest scores, will be ranked above average, while the other half, with the lowest scores, will be ranked below average.]

{At the beginning of the experiment you will be asked to toss a die. Outcomes from all participants will be gathered to determine your relative position. Half of the people in this room, with the highest draws, will be ranked above average, while the other half, with the lowest draws, will be ranked below average.}

Your rank will determine the type of signal you will receive. If you rank above average, you will receive a perfect signal throughout all 10 rounds. If you rank below average, you will receive an imperfect signal throughout all 10 rounds.

The absolute score and relative rank will not be reported to you.

Description of a Round

At the beginning of each round a single number will be randomly drawn by the computer. This number, which is equally likely to be any integer between 0 and 1,000 will not be revealed to anyone.

After drawing the number, an independent signal (about the drawn number) will be randomly selected by the computer for each player and will be displayed on their interface (the signal will come in the form of a number).

The signal that you receive will be one of two types: perfect signal or imperfect signal. If you receive a perfect signal, the signal reveals what the drawn number is (i.e., if the drawn number is Y, your signal is Y as well). If you receive an imperfect signal, the signal is equally likely to be any integer within 30 of the drawn number (i.e., if the drawn number is Y, your signal is equally likely to be any integer between Y-30 and Y+30).

At the beginning of each round you will be randomly paired with one other participant selected from the people in this room. The pairing will be such that one player receives a perfect signal while the other player receives an imperfect signal. That is, if you receive a perfect signal, the other player will receive an imperfect signal. Likewise, if you receive an imperfect signal, the other player will receive a perfect signal.

It will be your task, as well as task of the participant with whom you are paired, to try and guess the drawn number.

Description of a Decision Turn

Each round is divided into 4 decision turns in each of which you will be asked to submit your guess about the value of the drawn number. Throughout the turns the drawn number, your signal value, your signal type and the participant with whom you are paired, all remain the same.

At the beginning of each decision turn, your terminal will prompt you to submit a guess by entering it in the appropriate box. You have to submit a guess every single decision turn for the experiment to progress.

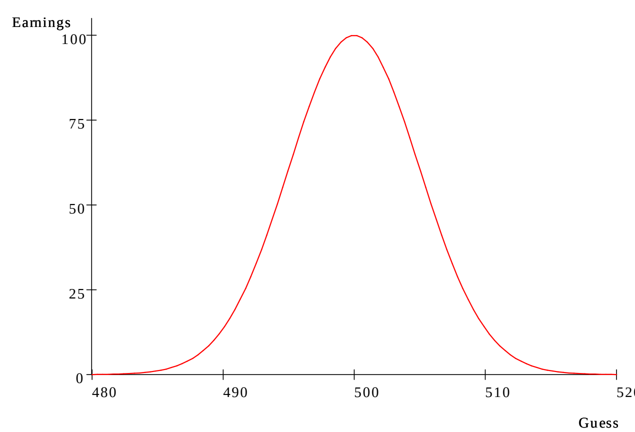
At the end of each decision turn, your terminal will display the other participant's guess, as submitted in that turn. Thus, from turn 2 and on, before submitting your guess, you will see the history of the guesses submitted in previous turns by you and the other player.

Calculating Earnings

At the end of each round, one of the four decision turns will be randomly selected. Your earnings for the round will be based on the accuracy of your guess, in that turn, relative to the drawn number.

Specifically, if the drawn number is Y and your guess is Z, your point earnings are: 100 times $exp\{-[(Y-Z)2]/50\}$. Thus, if you make a perfect guess of the drawn number, you will earn 100 points. The further away your guess is (from the drawn number), whether above or below it, the fewer points you earn. In short, the more precise your guesses are the more money you will earn at the end of the experiment.

To see that, consider an example where the drawn number is 500. Your earnings under different guesses are depicted below:



The sum of your point earnings across all rounds will be your point earnings for the experiment. These points will be converted into dollars using the ratio given above.

Starting the experiment

If you have any questions during the rest of the experiment, please raise your hand and a monitor will come to your desk and answer them in private. Otherwise, please follow the messages that will be displayed on your screen as they will guide you through the various stages of the experiment.

Upon completing the experiment further instructions will be provided.

8.3 Sample SAT questioner

Please indicate your answer to the following questions:

1. At bedtime the security blanket served the child as — with seemingly magical powers to ward off frightening phantasms.

A an arsenal

B an incentive

C a talisman

D a trademark

E a harbinger

2. Military victories brought tributes to the Aztec empire and, concomitantly, made it —, for Aztecs increasingly lived off the vanquished.

A indecisive

B pragmatic

C parasitic

D beneficent

E hospitable

3. Unlike sedentary people,—- often feel a sense of rootlessness instigated by the very traveling that defines them.

A athletes

B lobbyists

C itinerants

D dilettantes

E idealists

4. The researchers were — in recording stories of the town's African American community during the Depression, preserving even the smallest details.

A obstreperous

B apprehensive

C compensatory

D radicalized

E painstaking

5. WOOD : ROTTEN ::

A soil : sandy

B water : frozen
C paper : crumpled
D bread : moldy
E glass : broken

8.4 Tables (not included in the text)

Table 2: 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	5.52/8.32	4.71/6.76	4.71/6.76
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	4.19/6.43	3.79/6.38	3.79/6.38
Obs-RE	1.46/0.84	1.32/4.26	2.73/7.60	2.81/6.79
Turn	1 to 4			
		Price volatility index		
Obs	4.63/5.33			
RE	3.30/3.67			
NABB	3.57/2.77			
Obs-RE	1.33/1.66			

Notes: this table reports the average levels of volume, price error and price volatility 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 3: Change in price error and volume indexes

	Turn 1 vs 2	Turn 2 vs 3	Turn 3 vs 4	Turn 2 vs 4
(p values)				
		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 4: Price error and volume 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
Turn	1 to 4			
	Price volatility index			
All	0.6673			
Rounds 1-3	0.2541			
Rounds 4-7	0.8968			
Rounds 8-10	0.5713			

Notes: this table reports non-parametric (Mann-Whitney) test results of the null that median observed price error, volume and price volatility 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 5: 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 6: Comparing the QRE, RE and PI models

Turns	1	2	3
Median error			
NABB	1.0	1.5	2.6
PI	1.0	5.0	5.3
RE	1.0	3.6	3.5
Median error equality test (p-value)			
NABB vs. PI	NA	0.000	0.006
NABB vs. RE	NA	0.000	0.228

Notes: the upper part of this table describes median prediction error obtained by comparing observed decisions at each of the turns to the predicted decisions derived from the Noisy Actions Biased Beliefs (NABB), Rational Expectations (RE), and Private Information (PI) models, one at a time. The bottom part of the table reports non-parametric (Mann-Whitney) test results of the null that the NABB prediction errors are equal to those obtained from either the PI or the RE models.

Table 7: Correlation between subjective and objective probabilities

Rounds	1&2	3&4	5&6	7&8	9&10
			BLT		
Pearson correlation	0.46	0.59	0.77	0.77	0.65
t value	2.60	3.50	5.35	5.27	4.07
			OCT		
Pearson correlation	0.20	0.37	0.32	0.29	0.39
t value	1.28	2.42	2.05	1.84	2.54

Notes: this table reports the correlation between participants' subjective probabilities of being the perfectly informed type, estimated from the NABB model, at each stage of the experiment (rounds 1&2, rounds 3&4, etc.), and their corresponding objective probabilities. T-test results of the null that the correlation is equal to zero are also reported.

Table 8: Subjective probability persistance test

Rounds	1&2-3&4	3&4-5&6	5&6-7&8	7&8-9&10
		BLT		
switches	6^{***}	8^{**}	4^{***}	4^{***}
total	30	30	30	30
		OCT		
switches	8^{***}	8^{***}	8^{***}	10^{**}
total	42	42	42	36

Notes: this table reports the number of cases in which subjects 'switched' in their estimated miscalibration in the course of the session, i.e., moved from exhibiting above (below) median miscalibration in a given round to exhibiting below (above) median miscalibration in the subsequent round. Significance levels of the binomial test (null: the probability of observing 'switches' is equal to $\frac{1}{2}$) are denoted with superscript **(***), corresponding to significance levels of 5% (1%).

Table 9: Mean miscalibration

Rounds	1&2	3&4	5&6	7&8	9&10
BLT					
Mean	0.00	0.01	0.03	−0.06	−0.02
parametric (z value)	0.02	0.15	0.55	−1.10	−0.26
non-parametric (p value)	0.93	0.37	0.78	0.47	0.99
OCT					
Mean	0.16	0.14	0.07	0.07	0.13
parametric (z value)	2.10**	2.09**	0.91	1.01	1.80**
non-parametric (p value)	0.03**	0.05**	0.27	0.22	0.04**

Notes: this table reports the average miscalibration (across subjects) along with two test results of the null that the average level is equal to zero: one is a standard z-test, the other is non-parametric sign-test. Superscript ** represent significance levels greater than 5%.

Table 10: Miscalibration 2-way ANOVA

Source	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
Round	.2065	4	.05162	0.3	.876
Treatment	.0806	1	.0806	4.73	.030
Error	55.24	324	.1705		
Total	56.25	329			

Notes: this table reports the 2-way ANOVA test results of individual miscalibration when the data is classified by treatment (BLT vs. OCT) and round number.

Table 11: Relating subjective and objective probabilities

	$\widetilde{q}_i = \delta + \gamma q_i + e_i$		
	$\widehat{\delta}$	$\widehat{\gamma}$	R^2
BLT	0.1074^{**}	0.7726^{**}	0.4196
OCT	0.4251^{**}	0.3603^{**}	0.0978
F			22.12

Notes: this table reports regression results of subjects' objective (q_i) and subjective probabilities ($\widetilde{q}_i$) of being the perfectly informed type, conducted separately for each treatment (BLT and OCT). In addition, we test for equality of intercept and slope (F value). Superscript ** represents significance levels of 5%.

Table 12: Miscalibration persistence test by initial classification

Initial classification	# of observations	Fraction of subjects flipping
	BLT	
Overconfidence	15	0.250**
Underconfidence	15	0.250**
	OCT	
Overconfidence	29	0.208***
Underconfidence	13	0.300**

Notes: this table reports the number of subjects that moved from being overconfident (underconfident) in the early part of the experiment (rounds 1 through 4) to being underconfident (overconfident) in the late part of the experiment (rounds 7 through 10), under each treatment. For example, the row labeled "Overconfidence" under "BLT" suggest that 0.25 of all subjects who initially exhibited overconfidence in the BLT, exhibited underconfidence in the later part of the experiment. Binomial test results of the null that this fraction is equal to $\frac{1}{2}$ are denoted by superscript ** (***), representing significance levels of 5% (1%).

Table 13: Change in miscalibration

Initial classification	Worsened	No-change	Improved	Total
	BLT			
Underconfident (-1<MC<-1/3)	25%	25%	50%	4
Well calibrated (-1/3<MC<1/3)	35%	35%	30%	20
Overconfident (1/3<MC<1)	0%	17%	83%	6
Total (# of observations)	10	2	18	30
	OCT			
Underconfident (-1<MC<-1/3)	50%	17%	33%	6
Well calibrated (-1/3<MC<1/3)	52%	19%	29%	21
Overconfident (1/3<MC<1)	27%	6%	67%	15
Total (# of observations)	18	6	18	42

Notes: this table reports the fraction of subjects, in each treatment, that became less / equally / more miscalibrated in the course of the session (comparing average individual miscalibration in rounds 1 through 4 to rounds 7 through 10), classified by their initial level of miscalibration (denoted by "MC"). For example, the first column in the line labeled "Underconfident (-1<MC<-1/3)" suggests that 25% of the initially underconfident subject in the BLT became more underconfident, the second column in that line suggests that another 25% of these subjects maintained roughly the same level of miscalibration while the remaining 50% became less underconfident.

Table 14: Regressing initial on closing miscalibration

$MC_i^{7-10} = a + b_1 MC_i^{1-4} + b_2(Dummy_i * MC_i^{1-4}) + e_i$	$\widehat{a}$	$\widehat{b}_1$	$\widehat{b}_2$
Estimate	-0.0193	0.4424^{**}	0.2749^{*}
Standard error	0.0330	0.1490	0.1766

Notes: this table reports the regression results of subjects' initial miscalibration (averaged across rounds 1 through 4), denoted by MC_i^{1-4}, treatment dummy variable taking the value 1 in the OCT and 0 in the BLT, and final miscalibration (averaged across rounds 7 through 10), denoted by MC_i^{7-10}. Superscript * (**) denotes significance levels of 10% (5%).

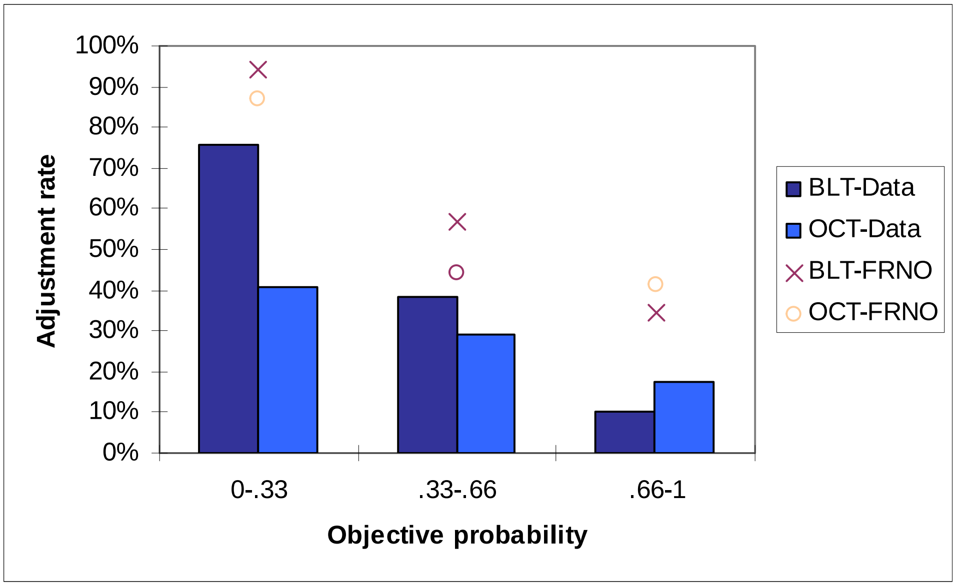
Table 15: Relating price errors and volume to mean and dispersion of miscalibration

Estimate (t value)	$\widehat{a}$	$\widehat{b}_1$	$\widehat{b}_2$	$\widehat{b}_3$
Volume index: $Vol_i = a + b_1 Dummy_i + b_2 MMC_i + b_3 DMC_i + e_i$				
(1)	5.7099	3.8135	6.3439	
	(8.6663)	(4.2808)	(3.7868)	
(2)	4.7376	3.7474		2.2387
	(5.9731)	(4.1047)		(1.9353)
(3)	4.721	3.1068	6.9482	2.7344
	(6.0386)	(3.3431)	(4.1494)	(2.3784)
Price Errors index: $PE_i = a + b_1 Dummy_i + b_2 MMC_i + b_3 DMC_i + e_i$				
(4)	2.9371	2.8734	2.2692	
	(7.7442)	(5.6033)	(2.3531)	
(5)	1.5095	2.2751		4.2813
	(3.365)	(4.4061)		(6.5437)
(6)	1.4108	1.8865	2.9941	4.5473
	(3.1875)	(3.5857)	(3.1583)	(6.9863)
Price volatility index: $PStdv_i = a + b_1 Dummy_i + b_2 MMC_i + b_3 DMC_i + e_i$				
(7)	4.5456	0.17172	−0.14642	
	(16.533)	(0.46192)	(−0.20944)	
(8)	4.2868	−0.0091144		0.68826
	(13.1)	(−0.024198)		(1.4421)
(9)	4.2865	−0.0084079	−0.03222	0.6877
	(12.986)	(−0.021428)	(−0.045571)	(1.4167)

Notes: this table reports robust regression results of treatment dummy variable, denoted $Dummy_i$, mean miscalibration across subjects in a given market, denoted by MMC_i, dispersion of miscalibration across subjects in a given market (absolute difference between their individual miscalibration), denoted by DMC_i, and three market measures: volume index, price error index and price volatility index. Estimates' t-value are reported in parenthesis.

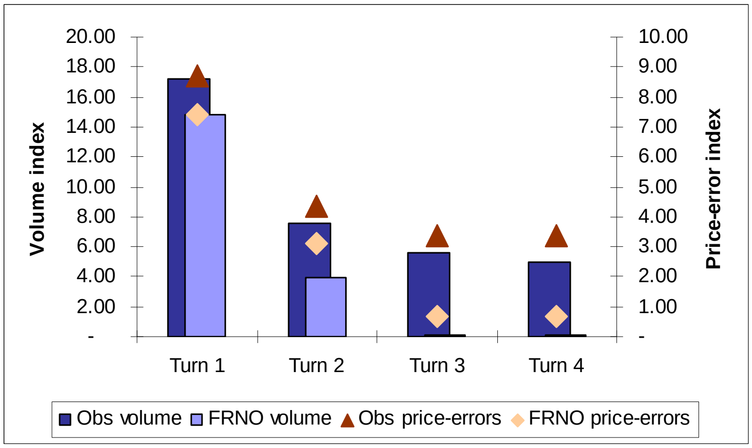
8.5 Figures (not included in the text)

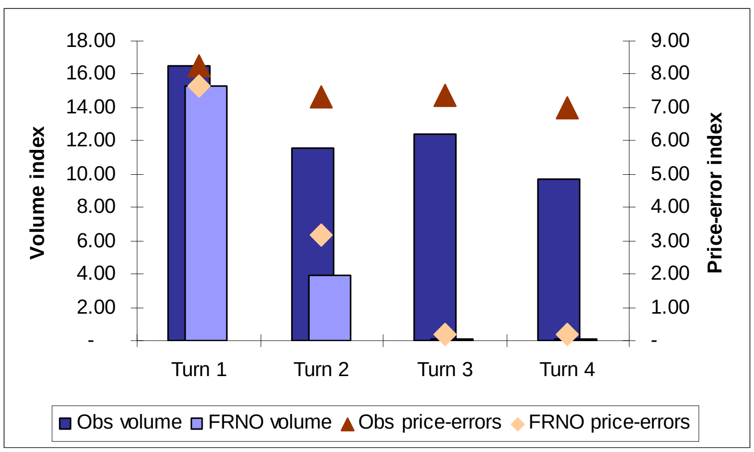
Figure 4: Adjustment rate across treatments by initial probabilities



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 (FRNO) predicted average adjustment rate for each of the subgroups.

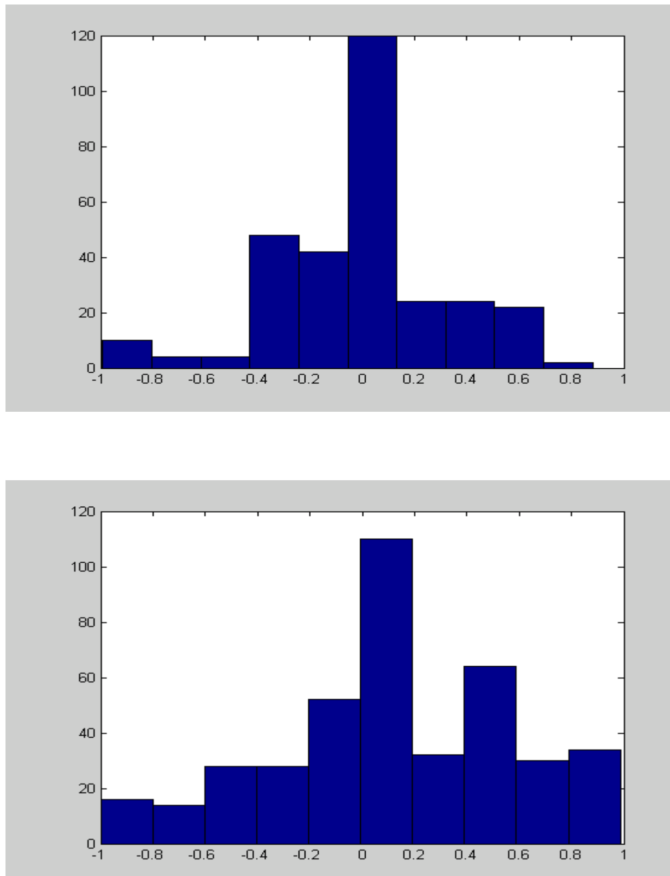
Figure 5: Adjustment rate across treatments by initial probabilities





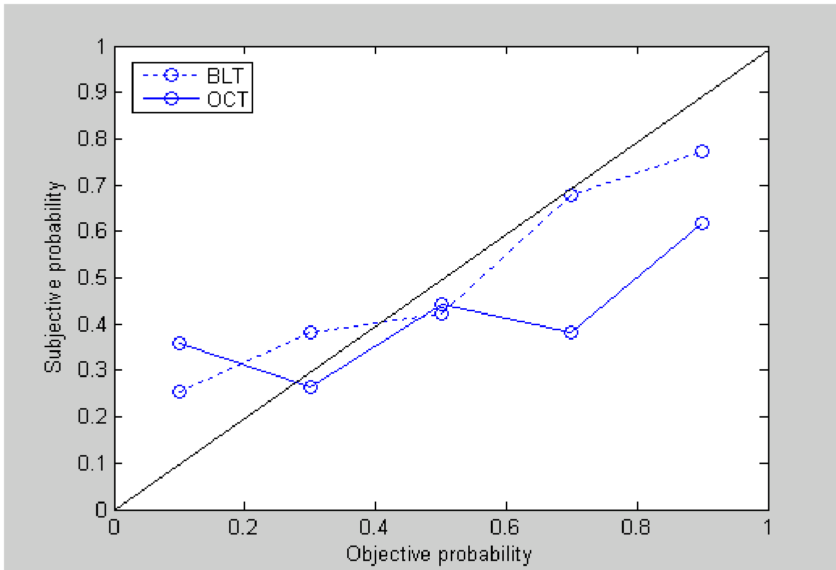
Notes: these figures depict the average (across all rounds and sessions) observed volume and price error indexes against the corresponding levels predicted by the fully-rational-no-overconfidence model (FRNO) for the BLT (top) and the OCT (bottom).

Figure 6: Individual miscalibration histograms



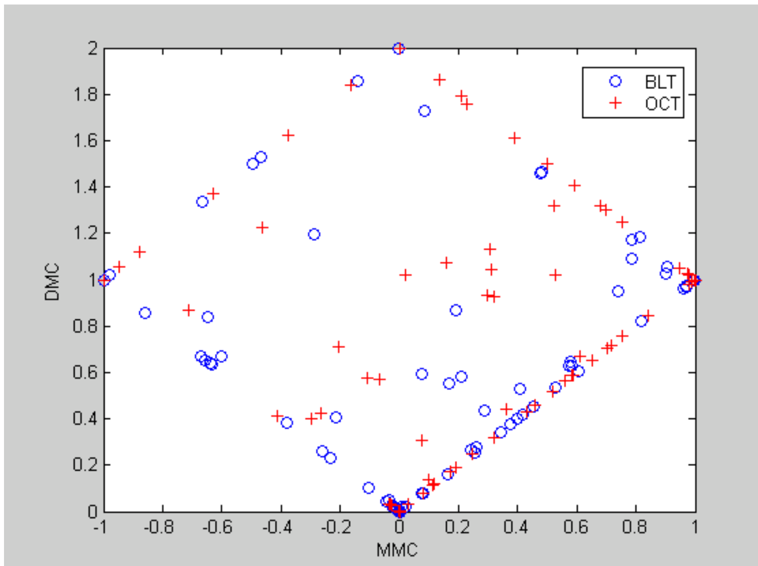
Notes: these figures plot the distribution of individual miscalibration for all rounds in the BLT (top) and the OCT (bottom). Positive values represent overconfidence while negative values represent underconfidence.

Figure 7: Objective vs. subjective probabily plot (BLT and OCT)



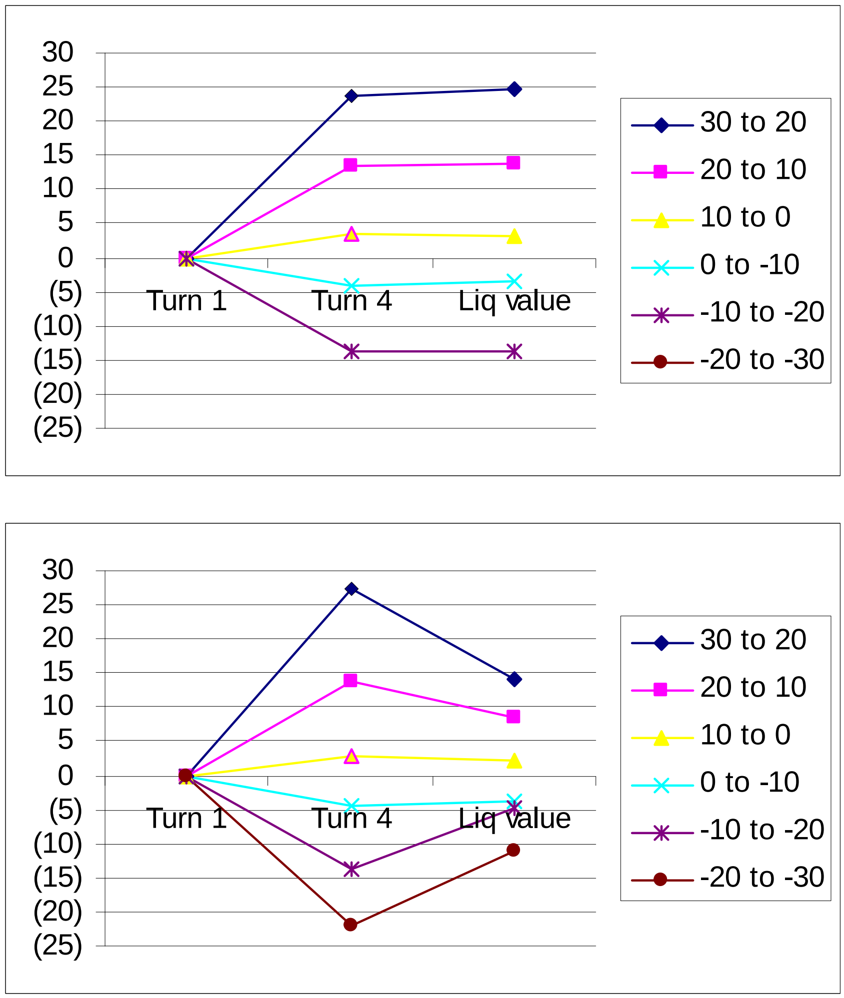
Notes: this figure contrasts subjects' objective probability of being the perfectly informed type with the corresponding subjective probability (obtained from the NABB model estimation). Observations are grouped by objective probability (horizontal axis). Each dot represents the average subjective probability for the group. The 45% line depicts the case of no-miscalibration.

Figure 8: Distribution of MMC and DMC across treatments



Notes: this figure depicts the distribution of all market instances by mean miscalibration (MMC) and dispersion of miscalibration (DMC).

Figure 9: 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.