Accounting for Noise in the Microfoundations of Information Aggregation *

Sera Linardi

California Institute of Technology†

Abstract

This paper shows that the basic unit of information aggregation described by the Geanakoplos and Polemarchakis (1982) posterior revision process does not always produce public statistics that are closer to the full information posterior than the common prior. I study this process of back and forth communication between two individuals with private signals by introducing white noise into payoff computations, defining the evolution of common knowledge, and providing conjectures on the resulting public statistics. I then develop a computational method to ex-ante rank information structures on likelihood of beating the prior. Through a laboratory experiment I find that subjects behave in ways consistent with the noise model: even though initial revisions bring reports closer to each other and to the full information posterior, subsequent revisions have little effect. A comparison of report errors across four information structures provides evidence that the ex-ante ranking matters; more than half of subjects' reports are further from the full information posterior than the common prior in the two lowest ranked structures.

JEL: C91, D82, D83, C73

Keywords: Information aggregation, noisy communication protocols, common knowledge.

*This is a preliminary draft, comments are welcome. This research is funded by Caltech Social Information Science Laboratory. I am especially thankful to John Ledyard and Tom Palfrey for their guidance throughout the project. I thank seminar participants at ESA (Tuscon, 2008, 2009), the SISL-Yahoo! Theory Workshop, and the WEAI 2009 Graduate Student Dissertation Workshop in Vancouver for helpful comments.

†Division of Humanities and Social Sciences, 1200 E. California Blvd, MC 228-77, Pasadena, CA 91125 USA. Phone: (626) 395-8772. Fax: (626)405-9841. Email: slinardi@hss.caltech.edu.

1 Introduction

One of the most common human interactions consists of individuals inferring information from the behavior of others and reacting to it. The assumption that people do this correctly serves as a theoretical basis for a large variety of economic mechanisms under uncertainty. Under rational expectations equilibrium (REE), prices are supposed to aggregate all the information about the states of the world. Those who are 'outsiders' to the information in the system become 'insiders' by simply observing these statistics (Plott, 2000). This has led to a large literature on the design and use of economic mechanisms for information aggregation. However, increasing experimental evidence shows that prices in markets and other forms of public statistics (e.g prices in market, odds in parimutuel, and numbers in polls)[1] only approximately converges to the full information posterior and that 'increased complexity is accompanied by increased noise' (Sunder, 1995).

Noisy and approximate aggregation implies that an outside observer may not always be better off abandoning the common prior in favor of the public statistics produced by the interactions of insiders. Various institutional components such as aggregate uncertainty (Plott and Sunder 1988, Lundholm 1991, Ackert and Church 1998), number of assets (O'Brien and Srivastava, 1991), and signal informativeness (Goeree et al, 2008) seems to contribute to informational efficiency. In this paper I capture the interaction of these three components by analyzing information structures. Focusing on a basic aggregation process, I provide a method of ranking information structures on the relative likelihood that public statistics improve predictions.

I study the Geanakoplos and Polemarchakis (1982) posterior revision process which describes how common knowledge of the full information posterior is reached instantaneously when two individuals with private signals communicate. This setup is fundamentally different from the more commonly studied information cascades setting since individuals here have repeated opportunities to revise their previous actions after receiving feedback from others' action. The posterior revision process forms the microfoundation of theoretical information aggregation; extensions by McKelvey and Page (1995) and Nielsen et al (1995) describe how dispersed information is incorporated into market

[1] Some example are Plott and Sunder (1988) for double auctions, Plott and Roust (2006) for parimutuel, McKelvey and Page (1995) for iterative polling, and Ledyard, Hanson, Ishikida (2009) for poll with Market Scoring Rule. See Healy, Ledyard, Linardi, and Lowery (2009) for experimental comparisons of the aggregation properties of the four mechanisms here.

prices[2] and the insights from this paper continues to be influential in recent work.[3]

I perturb this basic unit of aggregation by introducing white noise into payoff computations. The information content of reports and the speed at which beliefs converge become dependent on the information structure and noise level. Each step of revising and reporting introduces new noise into the system, thus limiting the marginal impact of additional revisions. In some information structures, a low level of perturbation is sufficient to result in public statistics that are worse approximation of the full information posterior than the common prior. When the expected distance of reports from the full information posterior is monotonically increasing in the noise level, there is a unique solution to the probability of improving predictions in an information structure. An outsider with no information about the noise level can use this result to ex-ante rank a set of information structures on the likelihood of improving prediction using the posterior revision output.

I ex-ante rank four information structures on the likelihood of prediction improvement. I then provide the first in-depth experimental investigation of the dynamics of posterior revision by testing the process in these four structures.[4] I find that subjects behave in ways consistent with the noise model: even though initial revisions bring reports closer to each other and to the full information posterior, subsequent revisions have little effect. Report errors remain significant throughout the revision process. Comparison of these errors across the four information structures provides evidence that the ex-ante ranking matters. I show that fragility to mistakes is not simply a function of states: a three state information structure that theoretically tolerates more noise indeed resulted in more frequent and larger prediction improvement compared to a two state structure that tolerates less noise.

The rest of the paper proceeds as follows. Section 2 defines information structures and the posterior revision process and provide a metric to measure aggregation success. Section 3 analyzes the changes in the process when stochastic noise is present. Section 3.1 presents a general model of noisy common knowledge evolution and the resulting posterior revision output. A computational method to rank information structure for noise tolerance follows. Section 3.2 assumes particular behavioral model

[2]See Feigenbaum, Fortnow, Pennock, and Sami (2005), Chen, Mullen, and Chu (2006), and Ostrovsky (2009) for connection between the posterior revision process and prediction markets.

[3]See Hanson (2002), Menager(2005), Acemoglu, Chernozhukov and Yildiz (2005), Liu (2008), and Zimper (2009)

[4]Weber (2001) dirty face game is related to the common knowledge inference process but does not directly depict the posterior revision process.

with white noise payoff perturbances and derives testable conjectures. Section 4 summarizes the hypotheses and presents the experimental design. Results are presented in Section 5 and organized into the following units of analysis: individual behavior (5.1), behavior within group (5.2), and outcome of the posterior revision process (5.3). Section 6 concludes.

2 Posterior revision as the microfoundation of aggregation

An information aggregation setup can be described follows: Θ is a finite set of states with an associated probability distribution $p \in \Delta_\Theta$, where Δ_Θ is the set of all possible probability distributions over Θ. State $\theta \in \Theta$ occurs with probability $p(\theta)$. *Insiders* $i \in \{A, B, .., I\}$ privately observe signal $s_i \in S_i$ which is drawn with probability $q_i(s_i|\theta)$. The distribution of signals across all insiders is denoted as $s \in S = S_A \times .. \times S_I$ and drawn with probability $q(s|\theta) = \prod_{i=1}^{I} q_i(s_i|\theta)$. As is standard in the literature, we consider only setups with states that are distinguishable by signals, $q(s|\theta) \neq q(s|\theta')$ if $\theta \neq \theta'$.

The effect of various components of this setup on information aggregation have been studied separately. O'Brien and Srivastava (1991) show that aggregation degenerates as number of assets $|\Theta|$ increases. Plott, Witt, and Yang (2003) show that aggregation improves as the private signal becomes more informative. However, the tradeoff between these components remain unknown. For example, is aggregation better when there are three states with informative signals or two states with less informative signals? How do the prior probabilities of the states affect aggregation? The interaction between these three components can be captured by analyzing the information structure as whole.

Definition 1. *Let $(\Omega, 2^\Omega, \psi)$ be a probability space where $\Omega = \Theta \times S$ and $\psi : 2^\Omega \to [0, 1]$ is the probability measure that represents the common prior. $\psi(\theta, s)$ is defined as $q(s|\theta)p(\theta)$, $\psi(\theta|s) = \sum_s \psi(\theta, s)$, and $\psi(s) = \sum_\theta q(s|\theta)$. We call ψ an* ***information structure****.*

Whether an information aggregation mechanism takes the form of a market, a parimutuel, or repeated polling, the central mechanism is as follows: individuals holding private information choose actions, which influences the formation of a public signal. The announcement of this public signal allows individuals to make inferences about other individuals' private information, which induces them to revise their actions. These new actions produce the next public signal, one that in theory contains more information than the previous public signal.[5] Communication between two Insiders makes up the most basic unit of this process: an individual A chooses an action, another (B) observes the action, updates her information, and chooses her own action, which in turn induces A to revise his action. This is the Geanakoplos and Polemarchakis (1985) posterior revision process. When this basic unit works well and instantaneously, this process produces the full information posterior of A's and B's combined private signals. Observers of this statistic incorporate A and B's private signals with their own information before taking further actions, thus advancing the aggregation process (McKelvey and Page, 1995 and Nielsen et al, 1995). However, this statistic may not be the full information posterior if the posterior revision process works slowly or noisily, and errors may accumulate throughout the aggregation process.

Geanakoplos and Polemarchakis (1985) assume that individuals are communicating truthfully; the incentives for communicating is undefined. McKelvey and Page (1990) augments the process with the quadratic scoring rule (qPSR) and show that myopic truthful reporting at every step by every subject is a Bayes-Nash equilibrium.[6] The posterior revision process studied here follows their precedence in utilizing qPSR.[7] Here individuals are directly assumed to be myopic: they choose a_i^t to maximize period t's payoff.

Definition 2. *Let $i \in \{A, B\}$. Define the incentivized T-step* ***posterior revision process*** *as $PR^T = (\Delta_\Theta \times \emptyset, Y_i, \gamma)$. $\Delta_\Theta \times \emptyset$ is the finite set of actions that i can choose from. The action of individual i is $a_i = (a_i^1, ...a_i^T)$ where $a_i^t \in \Delta_\Theta$ when t is odd and $i = A$ or when t is even and $i = B$; $a_i^t = \emptyset$ otherwise. A scoring rule assigns a score to a probability report a_i^t given the realization of the state of the world $\tilde{\theta}$, which happens after $t = T$. $Y_i(.) = \sum_{t=1}^{T} Q_{\tilde{\theta}}(a_i^t)$ is the state contingent payoff function for player i. $Q(.)$ is the quadratic scoring rule defined as:*

$$Q_{\tilde{\theta}}(a_i^t) = 1 - \sum_{\theta \in \Theta} (a_i^t(\theta) - \eta(\theta|\tilde{\theta}))^2$$

where $\eta(\theta|\tilde{\theta}) = 1$ when $\theta = \tilde{\theta}$ and 0 otherwise. $Q_{\tilde{\theta}}(a_i^t) = 0$ when $a_i^t = \emptyset$. The outcome function γ

[5] Contrast this with information cascades where an individual only takes a single action that he will not be able to revise.

[6] See McKelvey and Page's extension of the posterior revision process to a three person iterative poll.

[7] For favorable properties of QPSR in terms of handling risk aversion, accuracy, and calibration, see Sonnemans and Offerman 2001 and Palfrey and Wang 2009.

produces a probability distribution over Θ from the actions taken by the players; $\gamma(a_A, a_B) = a^T_{I(T)}$ where $I(T) = A$ when T is odd and $I(T) = B$ when T is even.

Lemma 2.1 shosw that myopic payoff optimizers facing a quadratic scoring rule report their current period beliefs, and hence, by Geanakoplos and Polemarchakis (1982), the second step of the incentivized T-step posterior revision process produces the full information posterior.

Lemma 2.1. *Let $p^t_i \in \Delta_\Theta$ denote i's belief at time t. A myopic payoff optimizer i facing a payoff function of $\sum_{t=1}^{T} Q_{\tilde{\theta}}(a^t_i)$ maximizes his payoff by reporting $a^t_i = p^t_i$.*

All proofs are in the Appendix.

Indicate the private signal realizations as $\tilde{s} = \tilde{s}_A\tilde{s}_B$ and the full information posterior as $p_{\tilde{s}}$. The information structure ψ and behavioral model is common knowledge[8].

Theorem 2.2. *Geanakoplos and Polemarchakis (1982). Let $i \in \{A, B\}$ be two myopic payoff maximizers participating in PR^T. If $\tilde{s}$ occured, then for all PR^T where $T \geq 2$, $\gamma(.) = p_{\tilde{s}}$.*

The intution is provided below, see the original paper for the full proof. Under Bayesian updating, $p^1_A(\theta) = \psi(\theta|\tilde{s}_A)$ and $p^1_B(\theta) = \psi(\theta|\tilde{s}_B)$. By Lemma 2.1 at step 1, $a^1_A = p^1_A$. Under common knowledge of Bayesian updating, $Pr(s_A|a^1_A = p^1_A) = 1$ for $s_A = \tilde{s}_A$ and 0 otherwise. Hence $p^2_B(\theta) = \psi(\theta|\tilde{s}_A\tilde{s}_B) = p_{\tilde{s}}$ and by Lemma 2.1 at step 2 $a^2_B = p_{\tilde{s}}$. Under common knowledge of Bayesian updating, $Pr(s_B|a^2_A = p^1_A) = 1$ for $s_B = \tilde{s}_B$ and 0 otherwise. Hence $p^3_A(\theta) = \psi(\theta|\tilde{s}_A\tilde{s}_B) = p_{\tilde{s}} = p^2_B(\theta)$, indicating that beliefs have converged. As before $a^3_A = p^3_A(\theta) = p_{\tilde{s}}$ and from this step onwards beliefs and reports remain at the full information posterior.

I now define a metric to measure aggregation success. Let $\omega^n = (\theta^n s^n) \in \Omega$ indicate the nth independent draw from $(\Omega, 2^\Omega, \psi)$. Let $p_{s^n} \in \Delta_\Theta$ indicate the full information posterior over Θ computed from s^n where $p_{s^n}(\theta) = \psi(\theta|s^n)$. Let $\alpha^n \in \Delta_\Theta$ be the observed output of PR^T when ω_n occured. Denote the set of full information posteriors and observed output distribution from N realizations of ω^n as $D^N = \{(p_{s^1}, \alpha^1), ..., , (p_{s^n}, \alpha^n), ..(p_{s^N}, \alpha^N)\}$. Following convention, the loss of approximating p_{s^n} with α^n is measured with the Euclidean distance. We refer to $d(p_{s^n}, \alpha^n) = (\sum_\theta (p_{s^n}(\theta) - \alpha^n(\theta))^2)^{\frac{1}{2}}$ as the *report error* of α^n.

[8]Common knowledge describes what A knows that B knows, B knows that A knows, A knows that B knows that A knows, and so on.

Define an *Outsider* as an individual $O \notin \{A, B\}$ who knows only of the common prior p and does not receive any private signal s_i. O does not participate in the posterior revision but chooses whether to use the output from the process in place of p to make predictions about Θ. Let *baseline error* be the average error over D^N from following the prior in ψ; this quantity indicates the loss suffered by the Outsider when no attempts to aggregate information are made.

$$BaseErr(D^N, \psi) = \frac{1}{N} \sum_{n=1}^{N} d(p_{s^n}, p)$$

By following PR^T's nth output in information structure ψ, the Outsider achieves an *error reduction* of $d(p_{s^n}, p) - d(p_{s^n}, \alpha^n)$. An error reduction of 0.4 in an information structure where the baseline error is 0.8 indicates a prediction improvement. This is not true for an information structure where the baseline error is 0.3. In the metric defined below, the size of error reduction is evaluated relative to the baseline error.

Definition 3. ***Prediction Improvement*** *of the nth output in information structure ψ is the error reduction of the nth output normalized by the baseline error of ψ.*

$$PI(n, D^N, \psi) = \frac{d(p_{s^n}, p) - d(p_{s^n}, \alpha^n)}{BaseErr(D^N, \psi)}$$

We write the average prediction improvement as $\bar{PI}(D^N, \psi)$.

This metric makes it easy to compare aggregation across different information structures. $\bar{PI}(D^N, \psi) = 1$ indicates that $\alpha_n = p_{s^n}$ for all $n \in \{1, .., N\}$. $\bar{PI}(D^N, \psi) \leq 0$ indicates that the error of the aggregation process in D^N is on average larger than the baseline error. In this case, the Outsider who naively abandons his prior would have been better off without the aggregation process.

With myopic payoff optimizers, an Outsider to the system can always improve upon the common prior by following the output of the posterior revision process. As we will see in the next section, this is not the case when stochastic noise affects A and B's actions.

3 Noisy Posterior Revision

Experimental evidence has shown that instead of the exact and instantaneous aggregation predicted by theory, various information aggregation mechanisms produce statistics that are only approximately and noisily converging to the full information posterior. Three types of explanations have received the most attention: systematic deviations from Bayesian updating,[9] the effect of risk attitudes,[10] and the existence of stochastic noise. There is mixed evidence on the type of systematic deviations from Bayes rule that best fit the information aggregation setting, and the assumption of risk neutrality has been found to be reasonable in laboratory experiment with small stakes (Holt and Laury, 2002 and Offerman, 2006); I therefore focus on the third explanation, the existence of stochastic noise.

Stochastic noise can be present in the perception of a private signal, in the belief updating process, or in payoff perception. In an experimental setting, private signals are often simple (for example, subjects draw a colored ball from an unknown urn) and the information remains with the subject for the duration of the experiment. Stochastic noise is more likely to be present in belief updating or in payoff perception. Modelling stochastic noise in payoff perception presents two advantages. First this allows the failures of aggregation to be examined while retaining the assumption of Bayesian updating, which plays a crucial role in the theoretical literature. Second, there is strong evidence of stochastic action in simple dominant strategy experiments where Bayesian updating does not play a role (Camerer 1989, Starmer and Sugden 1989, Hey and Orme 1994, Ballinger and Wilcox 1997),[11] suggesting that noise in payoff perception exists independently of noise in belief updating.

In the first subsection I describe the general structure of the posterior revision process with stochastic action. In the second subsection I assume that payoff disturbances can be represented with white noise ($\epsilon \sim \mathcal{N}(0, \sigma)$) and draw conjectures that will be tested in the experiment in Section 4.

3.1 Model of noisy common knowledge evolution

Let $f(\sigma)$ be the distribution from which payoff perturbances are drawn from where $\sigma = 0$ indicates no noise and $\sigma = \infty$ indicates pure noise. The noise level σ is a random variable drawn from an

[9] Camerer (1995) provide an overview of experimental evidence of deviations from Bayesian updating such as overweighting one's prior relative to private information (conservatism), neglecting one's prior when considering private information (base rate fallacy) (Kahneman and Tversky 1973), and overweighting one's private information relative to information from others (Hung and Plott 2001, Goeree et al. 2008).

[10] Risk aversion has an effect of 'flattenening' reported probabilities towards equal likelihood (Harrison and Rustrom, 2008).

[11] These studies found that more than 25% of subjects reversed their choices when presented with the same binary choice problem for the second time.

unknown distribution $g(.)$ at every time the true state of the world and private signal $\omega = (\theta, s_A s_B)$ is realized. Let F and G be the respective cumulative density functions for f and g. A and B, as Insiders, learn of s_i and the noise level σ that affects them; O, as the Outsider, does not. Think of A and B as two salespersons in a busy mall who have observed customers attitudes towards a product but are distracted by activities in the mall while discussing it; O is the CEO at the headquarters who is trying to infer customers attitudes from the communication between A and B but is unaware of the context in which this communication took place.

Let $Pr_\sigma(a_j^t|p_{s_j}^t)$ be the standard quantal response function of Luce (1957) denoting the probability of choosing an action a_j^t given beliefs $p_{s_j}^t$ and noise level σ. As before myopic players optimize only their current period payoffs, but here this computation is affected by stochastic disturbances.[12] With stochastic action, reports now could have been induced by any private signal. There is no longer the one to one mapping of reports to private signal that completed the aggregation process by the second step of the posterior revision process (Theorem 2.2). The state of common knowledge among A and B has to be enlarged to contain each person's hypothetical beliefs about Ω given each realization of the private signal.

Definition 4. *Denote* ***common knowledge*** *among myopic σ optimizers A and B at step t of the posterior revision process as $K^t = (K_A^t, K_B^t)$. K_A^t is the set $\{\psi_{s_A}^t\}_{s_A \in S_A}$ where $\psi_{s_A}^t$ is A's posterior over Ω at time t if A's private signal is s_A. Similarly, K_B^t is the common knowledge at time t about B's possible beliefs. The initial state of common knowledge is $K^1 = (K_A^1, K_B^1) = (\{\psi_{s_A}^1\}_{s_A \in S_A}, \{\psi_{s_B}^1\}_{s_B \in S_B})$ where $\psi_{s_i}^1(\theta, s_j) = \psi(\theta, s_j|s_i)$.*

A general behavioral model of common knowledge evolution is described below; the next section provides details under the assumption of white noise payoff disturbances. When actions are stochastic, individuals gradually learn about their partner's signals by intepreting the history of past actions within the context of the noise level. After observing A's action, B can update his probability distribution over A's private signals by comparing the likelihood of A's report given A's hypothetical beliefs. Notice that A is aware of B's updating process without being aware of the private signal that initialized B's beliefs. The resulting evolution of common knowledge is a path dependent;

[12] Because players are myopic, this is not an equilibrium model such as McKelvey and Palfrey (1995, 1998) quantal response equilibrium (QRE).

A is aware that the inaccuracies of his earlier reports affects B's beliefs which in turn affects A's interpretation of B's subsequent reports.

Definition 5. *An **update rule** is a function $L(K^t, a_i^t, \sigma) = K^{t+1}$ which takes in the current state of common knowledge K^t, noise level σ, in addition to the most recent report a_i^t and returns the next period state of common knowledge K^{t+1}. The update rule for the posterior revision process is*

$$\begin{aligned} L(K^t, a^t, \sigma) &= \begin{cases} K_A^t, \{\psi_{s_B}^{t+1}\}_{s_B \in S_B} & \text{if } t \text{ is odd} \\ \{\psi_{s_A}^{t+1}\}_{s_A \in S_A}, K_B^t & \text{otherwise} \end{cases} \\ \text{where } \psi_{s_i}^{t+1}(\theta, s_j) &= \frac{\psi_{s_i}^t(\theta, s_j) Pr_\sigma(a_j^t | p_{s_j}^t)}{\sum_{s_j'} \psi_{s_i}^t(\theta, s_j) Pr_\sigma(a_j^t | p_{s_j'}^t)} \\ \text{where } p_{s_j}^t &= \sum_{s_i} \psi_{s_j}^t(\theta, s_i) \end{aligned}$$

What is the output of this communication process? As before the true private signal realization is denoted as $\tilde{s} = \tilde{s}_A \tilde{s}_B$, and A's true belief is distinguished from A's hypothetical belief as $\psi_{\tilde{s}_A}^t$ instead of $\psi_{s_A}^t$. The probability of observing a certain action at step t of this process depends on the history of actions that has preceeded it. The expected error is therefore:

$$\mathbb{E}Err_{PR^T}(\tilde{s}, \sigma) = \mathbb{E}_{a_A^1} .. \mathbb{E}_{a_{I(T)}^T} \prod_{t=1}^{T} Pr_\sigma(a_{I(t)}^t | p_{s_{I(t)}}^t)[d(p_{\tilde{s}}, a_{I(T)}^T)]$$

Suppose T is odd, and hence the last person reporting is A. Then $a_{I(T)}^T$ is chosen with $Pr_\sigma(a_A^T | p_{\tilde{s}_A}^T)$ where $p_{\tilde{s}_A}^T = \sum_{s_B} \psi_{\tilde{s}_A}^T(\theta, s_B)$ where $\psi_{\tilde{s}_A}^T \in K^T = L(K^{T-1}, a^{T-1}, \sigma)$.

Definition 6. *The expected improvement of using the output of the posterior process in place of the common prior in information structure ψ is:*

$$EI_{PR^T}(\psi | \sigma) = \mathbb{E}_s \mathbb{E}_{a_A^1} .. \mathbb{E}_{a_{I(T)}^T} \prod_{t=1}^{T} Pr_\sigma(a_{I(t)}^t | p_{s_{I(t)}}^t) \frac{[d(p_s, p) - d(p_s, a_{I(T)}^T)]}{\mathbb{E}BaseErr(\psi)}$$

where $\mathbb{E}_s BaseErr(\psi) = \sum_{s \in S} \psi(s) d(p_s, p)$.

Given a set of different information structures, how can an Outsider decide where to ignore the output of an aggregation process and stay with the common prior? If expected report error $\mathbb{E}ErrPR^T(s, \sigma)$ is an increasing function in σ, there exists a unique noise level σ^* that solves $\mathbb{E}_s \mathbb{E}ErrPR^T(s, \sigma) = \mathbb{E}BaseErr(\psi)$. The lower σ^* is, the more accuracy is required of the Insiders

(A, B) in order to produce public statistics that approximate the true state of the world more accurately than the prior. This concept of difficulty as 'noise-tolerance' assigns a single number to each information structure, allowing comparison of aggregation difficulty to be made across information structures with different number of assets, prior probabilities, private signal informativeness, and number of private signal draws. This computational method will be used in Section 4.1 to ex-ante rank the performance of the posterior revision process across four different information structures.

Proposition 3.1. *Suppose $EI_{PR^T}(\psi|\sigma)$ is increasing in σ. For two probability spaces $(\Omega_1, 2^{\Omega_1}, \psi_1)$ and $(\Omega_2, 2^{\Omega_2}, \psi_2)$, if σ_1^* and σ_2^* are the respective solution to $\mathbb{E}_s\mathbb{E}ErrPR^T(s,\sigma) = \mathbb{E}BaseErr(\psi)$ in ψ_1 and ψ_2, then $\sigma_1^* > \sigma_2^*$ implies that for the posterior revision process, $Pr(PI(D^N, \psi_1) > 0) > Pr(PI(D^N, \psi_2) > 0)$ as $N \to \infty$.*

If $EI_{PR^T}(\psi_1|\sigma)$ weakly first order stochastically dominates $EI_{PR^T}(\psi_2|\sigma)$, $\sigma_1^* > \sigma_2^*$ also implies larger average prediction improvement in ψ_1 than in ψ_2 as N gets large.

In the next section I assume a particular behavioral model with white noise perturbances and arrive at several conjectures that I will test in the experiment. Readers interested in the experiment can continue directly to Section 4.

3.2 Posterior revision with white noise perturbances

I derive several conjectures under the assumption of a simple behavior model. Let the payoff perturbations be white noise of level σ, or in other words, let $\epsilon \in \mathbb{R}$ be a random variable distributed with $\mathcal{N}(0, \sigma)$. Individual i is aware of his true beliefs which he has arrived at by updating his old belief using Bayes rule. Assume that when faced with a quadratic scoring rule i is choosing only between reporting his true beliefs $p_i^t \in \Delta_\Theta$ or other probability distributions $\alpha \in \Delta_\Theta$. The region of error that induces i to report α when his true beliefs is p_i^t is:

$$R(\alpha|p_i^t) = \left\{(\epsilon, \epsilon') \in \mathbb{R}^2 : \mathbb{E}_{p_i^t} Q(\alpha) + \epsilon \geq \mathbb{E}_{p_i^t} Q(p_i^t) + \epsilon'\right\} \tag{1}$$

By Eq. 1, α is chosen when the difference in the payoff perturbance $\epsilon - \epsilon'$ is large enough to overcome the difference of i's expected payoff between reporting p_i^t and a_i^t. Since the difference of expected payoff between reporting p_i^t and a_i^t under the quadratic scoring rule is the Euclidean distance between them (Selten, 1998), it is simple to compute the probability distribution of actions given current beliefs.

Definition 7. *A **myopic** σ **optimizer** i faces payoff perturbation drawn from distribution F_σ, which is a normal distribution with mean 0 and variance σ. The probability that i reports α when p_i^t is the true belief is:*

$$Pr_\sigma(\alpha|p_i^t) = \frac{\int\limits_{R(\alpha|p_i^t)} dF_\sigma(\epsilon)}{\sum\limits_{a'\in\Delta_\Theta} \int\limits_{R(a'|p_i^t)} dF_\sigma(\epsilon)} \tag{2}$$

where $R(\alpha|p_i^t)$ be the set of payoff perturbances that such that the current period payoff of reporting α exceeds the payoff of reporting the truth p_i^t.

Lemma 3.2. *Let $F'_\sigma(d(p_i^t,\alpha))$ be the probability of drawing errors larger than $d(p_i^t,\alpha)$ under error distribution $\mathcal{N}(0,2\sigma)$.*
(i) The probability that a myopic σ optimizer i with beliefs p_i^t reports α under quadratic scoring rule $Q(.)$ is:

$$Pr_\sigma(\alpha|p_i^t) = \frac{F'_\sigma(d(p_i^t,\alpha))}{\sum\limits_{a\in\Delta_\Theta} F'_\sigma(d(p_i^t,a))}$$

(ii) For any $0<\sigma<\infty$, $Pr_\sigma(\alpha|p_i^t)$ is decreasing in $d(p_i^t,\alpha)$

How does this probability distribution change as a function of the noise level? Since the noise level is the variance of a centered normal distribution, we can make use of the following property of the normal distribution.

Proposition 3.3.

$$\forall \delta, k \in \mathbb{R}_+,\ \frac{F'_\sigma(\delta)}{F'_{\sigma+k}(\delta)} < 1$$

This ratio is decreasing in δ.

Proposition 3.3 provides Lemma 3.4 with the following property: if i is more likely to report α and less likely to report β when noise level σ decreases, α must be closer to i's true belief p_i^t than β. As a consequence, the average distance of reports to true beliefs is an inverse function of noise level.

Lemma 3.4. *Let $\sigma < \sigma'$.*

$$(i) Pr_\sigma(\alpha|p_i^t) > Pr_{\sigma'}(\alpha|p_i^t) \text{ and } Pr_\sigma(\beta|p_i^t) < Pr_{\sigma'}(\beta|p_i^t) \Rightarrow d(p_i^t,\alpha) < d(p_i^t,\beta)$$

$$(ii) \sum Pr_\sigma(\alpha|p_i^t) d(p_i^t,\alpha) < \sum Pr_{\sigma'}(\alpha|p_i^t) d(p_i^t,\alpha)$$

The rest of this section describes the dynamics of the posterior revision process under this behavioral model and draw conjectures that will be tested in the next section.

As in Definition 4, the initial state of common knowledge is $K^1 = (K_A^1, K_B^1) = (\{\psi_{s_A}^1\}_{s_A \in S_A}, \{\psi_{s_B}^1\}_{s_B \in S_B})$ where $\psi_{s_i}^1(\theta, s_j) = \psi(\theta, s_j|s_i)$.

At step 1, A's true belief about Θ is $p_{\tilde{s}_A}^1 = \sum_{s_B} \psi_{\tilde{s}_A}^1(\theta, s_B)$ where $\psi_{\tilde{s}_A}^1 \in K_A^1$. By Lemma 3.2 A announces $a_A^1 \in \Delta_\Theta$ with probability $Pr_\sigma(a_A^1|p_{\tilde{s}_A}^1)$. B takes no action, $a_B^1 = \emptyset$.

B draws inference about s_A from a_A^1 as follows:

$$Pr(s_A|a_A^1, K^1, s_B, \sigma) = \frac{Pr_\sigma(a_A^1|p_{s_A}^1)\psi_{s_B}^1(s_A)}{\sum\limits_{s'_A} Pr_\sigma(a_A^1|p_{s'_A}^1)\psi_{s_B}^1(s'_A)}$$

where each $p_{s'_A}^1$ is computed from $\psi_{s'_A}^1 \in K_A^1$. In the extreme case where $\sigma = \infty$, $Pr_\sigma(a_A^1|p_{s_A}^1) = Pr_\sigma(a_A^1|p_{s'_A}^1)$ for all $s_A, s'_A \in S_A$ and B can not update his belief over s_A regardless of a_A^1. For $\sigma < \infty$, since $Pr_\sigma(a_A^1|p_{s_A}^1)$ is decreasing in $d(a_A^1|p_{s_A}^1)$ by Lemma 3.2(ii), on average B increases the probability assigned to $\tilde{s}_A$ after seeing a_A^1.

Conjecture 1: On average, beliefs converge with revisions.

Updated beliefs about s_A drives the updating for the rest of the probability distribution.

$$\psi_{s_B}^2(\theta s_A) = \frac{\psi_{s_B}^1(\theta s_A)}{\sum\limits_{\theta'} \psi_{s_B}^1(\theta' s_A)} Pr(s_A|a_A^1, K^1, s_B, \sigma)$$

The common knowledge of posteriors is now updated to $K^2 = (K_A^1, \{\psi_{s_B}^2\}_{s_B \in S_B})$. The increase in the probability B assigns to $\tilde{s}_A$ results in second step beliefs $\psi_{s_B}^2$ that are closer to $\psi_{\tilde{s}_A s_B}$ than beliefs in the first step $\psi_{s_B}^1$. This is true as well for B's actual belief $\psi_{\tilde{s}_B}^2$, which results in $p_{\tilde{s}_B}^2$ that is closer to full information posterior $p_{\tilde{s}}$ than $p_{\tilde{s}_B}^1$.

$$p_{\tilde{s}_B}^2(\theta) = \frac{\sum\limits_{s_A} \psi_{\tilde{s}_B}^1(\theta s_A) Pr_\sigma(a_A^1|p_{s_A}^1)}{\sum\limits_{s'_A} \psi_{\tilde{s}_B}^1(\theta s'_A) Pr_\sigma(a_A^1|p_{s'_A}^1)}$$

Conjecture 2a: On average, report error decreases with revision.

However, the speed at which $p^2_{\tilde{s}_B}$ moves towards $p_{\tilde{s}}$ depends on the probability B assigned to $\tilde{s}_A$, which in turn depends on the difference between $Pr_\sigma(a^1_A|p^1_{\tilde{s}_A})$ and $Pr_\sigma(a^1_A|p^1_{s_A})$ where $s_A \neq \tilde{s}_A$. For a large range of $\sigma > 0$, $p^2_{\tilde{s}_B} \neq p_{\tilde{s}}$. The probability distribution B announces at $t = 2$ is chosen with probability $Pr_\sigma(a^2_B|p^2_{\tilde{s}_B})$. a^2_B not only communicates $\tilde{s}_B$ to A, but also introduces new noise into the system. The report error at the second step of the posterior revision process with myopic σ optimizer is:

$$\mathbb{E}Err_{PR^2}(\tilde{s}, \sigma) = \mathbb{E}_{a^1_A}\mathbb{E}_{a^2_A} Pr_\sigma(a^1_A|p^1_{s_A}) Pr_\sigma(a^2_B|p^2_{\tilde{s}_B}) d(p_{\tilde{s}}, a^2_B)$$

Conjecture 2b: Report error remains persistently above zero.

The inference that A is able to draw on s_B from a^2_B is affected that A's earlier influenced on B's belief through her first report a^1_A.

$$Pr_\sigma(s_B|a^2_B, K^2_B, \psi^2_{s_A}) = \frac{Pr_\sigma(a^1_B|p^2_{s_B})\psi^2_{s_A}(s_B)}{\sum\limits_{s'_B} Pr_\sigma(a^2_B|p^2_{s'_B})\psi^2_{s_A}(s'_B)}$$

Due to the information from A's preceeding report, there is less difference between two posteriors $p^2_{s_B}$ and $p^2_{s'_B}$ computed from K^2_B than there previously was in K^1_B. This limits the confidence with which A updates. However, on average A increases the probability assigned to $\tilde{s}_B$ after seeing a^2_B and moves closer to $\psi_{\tilde{s}_A\tilde{s}_B}$.

Conjecture 3: Additional revisions have decreasing marginal impact on belief convergence and reduction in report error.

For an arbitrary step number t, how does average report error behave as a function of the noise level? Suppose A has just made the most recent report a^t_A. Because the noise level is common knowledge, given the same state of common knowledge K^{t-1} and report a^t_A, B will move further towards the signal s_A implied by a^t_A under a lower noise level is low. If the signal s_A implied by a^t_A is more likely to be $\tilde{s}_A$ as noise level decreases, B's belief will be closer to the full information posterior as noise level decreases.

Lemma 3.4 states that a reduction in noise level results in higher frequency of reports that are closer to beliefs and smaller average distance of reports to beliefs. This implies that when noise is lower, A's report a_A^t are more likely to be closer to A's belief p_A^t. This in turn implies a higher frequency of 'correct' reports: reports that lead B to conclude that A's signal must have been $\tilde{s}_A$. This drives B's belief to be closer to the full information posterior as noise level decreases. B then chooses actions, and again by Lemma 3.4, B's action will be closer to his true beliefs (which is closer to the full information posterior) as noise level decreases.

Conjecture 4: Expected report error is increasing in σ. Therefore Proposition 3.1 holds: if information structure ψ_1 tolerates a higher noise level than information structure ψ_2 ($\sigma_1^* > \sigma_2^*$), the frequency of prediction improvement will be higher in ψ_1 than ψ_2.

The next section describes an experiment with four information structures that will test these conjectures.

4 Experiment

4.1 Experimental Design

The experiment follows the posterior revision process described in Section 2. Subjects are randomly paired and assigned either the first mover (A) or the second mover (B) role. The states of the world are represented as boxes X,Y,or Z, each with different quantities of Black and White balls.[13]. One of the boxes is randomly chosen with some prior probability. From that box a ball is privately drawn (with replacement) for each person.

Table 1 describes the four information structures tested in the experiment. Each row lists the probability that a particular box is chosen and the conditional probability of drawing a Black ball from that box. The information structures are selected to fulfill several criterion. First, a three state structure with a difficulty ranking between a pair of two state structures is chosen to stress test the ranking across different number of states. Second, the four structures are chosen so that $EI_{PR}(\psi, \sigma)$ is separated with stochastic dominance for all but the lowest noise levels in order to test the impact of the difficulty ranking on the size of prediction improvement. Third, to minimize the possibility that the visual display of the probabilities cue subjects to the difficulty level, information structures with

[13]See appendix for software screenshot.

Info structure	$Pr(X)$	$Pr(Y)$	$Pr(Z)$	$Pr(Black\|X)$	$Pr(Black\|Y)$	$Pr(Black\|Z)$
2E	0.40	0.60		0.79	0.13	
2H	0.33	0.67		0.30	0.60	
3E	0.40	0.50	0.10	0.15	0.85	0.30
3H	0.25	0.50	0.25	0.40	0.60	0.80

Table 1: Marginal and conditional distributions of four information structures

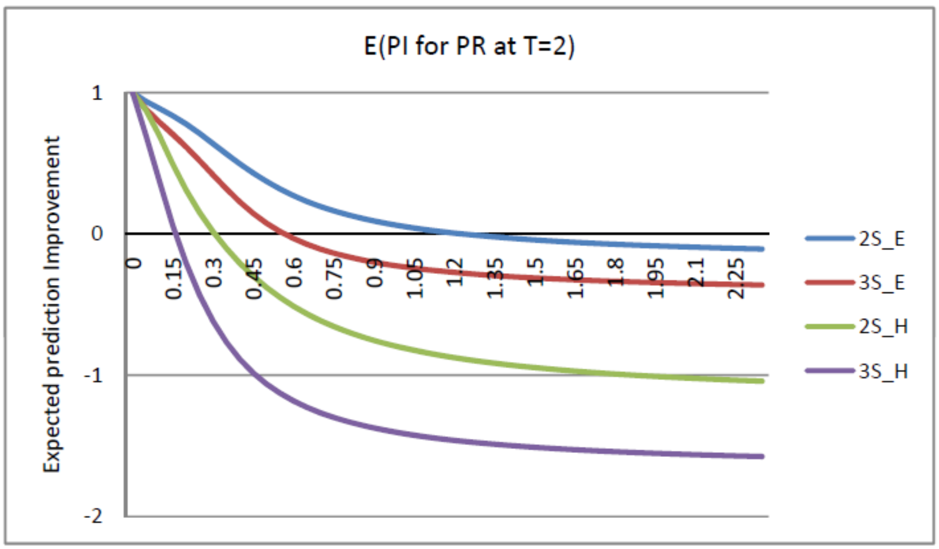

Figure 1: EI for theoretical distribution (y) against noise level σ (x)

equal likelihoods are excluded and two digit precision are included in every information structure.

Figure 1 plots Expected Improvement (Definition 6) of the four information structures against σ at $t = 2$. In these structures EI is an increasing function of σ and the graphs for each structure is separated from the others by stochastic dominance for all $\sigma > .1$.

Solving for $\mathbb{E}_s[d(p^s, p)] = \mathbb{E}_s \mathbb{E} Err_{PR^2}(s, \sigma)$ for $2E, 2H, 3E, 3H$ resulted in corresponding σ^* of $1.2, .55, .3, .15$. Expected error is slightly higher for $t = 3$, then went back to $t = 2$ level for $t = 4$, and remained constant at the 0.01 precision level afterwards. The ranking and the stochastic dominance is preserved for $t = 3, 4, 5$. The sensitivity of each information structure to noise implies the following difficulty ranking (from the structure where PI can be most easily achieved to the structure where it is hardest):

$$Difficulty\ Ranking:\ 2E < 3E < 2H < 3H$$

The experiment tests a five step posterior revision process that proceeds as follows. After seeing private signals, A and B are both asked to report the percentage chance that the ball they have observed came from Box X, Y, and Z .[14] A's report a_A^1 is shown to B, but B's report a_B^1 is not shown to A.[15] After observing a_A^1, B gets to revise his report and announces a_B^2 to A. A then revises and reports a_A^3. Another sequence of revisions from B (a_B^4), then A (a_A^5), follows before the identity of the chosen box is revealed.

To emphasize the number of revisions undergone by each player, I refer to a_A^1 and a_B^1 as *initial reports*, a_B^2 and a_A^3 as *first revisions*, and a_B^4 and a_A^5 as *second revisions*. The hypotheses tested in this experiment are as follows:

- **H1**: Actions are stochastic. The same individual will report different probability distributions when faced with the same private signal realization within the same information structure.
- **H2**: Beliefs are converging. Reports are updated in the direction of the partner's report, which is reflected in the decrease in average distance between a pair's reports with additional iteration of revision. Initial reports should be further apart than first revisions, which in turn are further apart than second revisions.
- **H3**: Reports are moving towards the full information posterior, however, due to the persistence of noise, the average distance from reports to the full information posterior remains significantly larger than zero. Hence the average initial report error is larger or equal to the average first revision error and the average first revision error is larger or equal to the second revision error.

[14]In a two state case subjects only need to enter the percentage chance for Box X (e.g Pr(Ball is from Box X)=85% implies Pr(Ball is from Box Y)=15%), and in the three state case subjects enter a percentage each for Box X and Box Y leaving the probability of Box Z to be automatically computed.

[15]B's report is elicited in order to learn about his initial belief and does not change the game since it is never shown to any players.

- **H4**: An additional iteration of revision has decreasing impact on both belief updating and report error. This is again an implication of persistent noise. Second revisions are only slightly closer together than first revisions; the impact of first revision on initial reports is much higher. Second revision error are also only slightly smaller than first revision error.

- **H5**: Prediction improvement is less frequent in information structures ranked as more difficult, or in other words, is more sensitive to noise. The likelihood of observing output closer to the full information posterior than the prior depends more on noise sensitivity than the number of states: prediction improvement in 3E is more frequent than in 2H.

- **H6**: Prediction improvement is smaller in information structures ranked as more difficult. Given two information structures, the posterior revision process achieves a smaller fraction of possible error reduction in the harder structure even if a larger number of possible states is present in the easier structure.

4.2 Implementation

The experiments were ran at California Institute of Technology in summer of 2009. Using within subject design, 30 subjects (in random pairings) participated in the 5 step posterior revision process in all four information structures from Table1. Data collection was conducted through three sessions with ten subjects each. Instructions (see Appendix) were distributed and read out loud by the experimenter.

For the posterior revision process, subjects interacted through a zTree software written for the experiment.[16] A software screenshot is included in the Appendix. Subjects were given thirty seconds to submit initial reports and twenty seconds in each opportunity to revise their reports. Once the time limit is reached, the experimenter announced 'Please enter your guess' at every fifteen-seconds interval until subjects enter their reports.

Each report were rewarded with Quadratic Proper Scoring Rule with 25 cents for a correct answer. When a round concludes, subjects observe their payoff for that round and proceeds to the next round where they will be randomly rematched and shown a new information structure and corresponding

[16] Pilot tests of the software were completed in fall 2008, the data is available upon request.

private signals. Each experiment consisted of 22 rounds of posterior revisions and take on average 75 minutes; each subject made a total 66 reports and earn between $19 to $22.

For the parameter set, a random number generator generated 11 realizations of boxes and ball draws for the four information structures, resulting in 44 (not necessarily unique) aggregation problems. Another random number generator ordered the 44 problems, grouping the first 22 as one parameter set and the last 22 as another. Experimental sessions alternated between the two parameter sets.[17]

5 Results

The presentation of results follows the hypotheses listed in Section 4.1. In the first two subsections, reports are pooled from all information structures, resulting in 660 observations each of initial reports, first revisions, and second revisions from 30 subjects who each participated in 22 rounds of posterior revisions with a random partner. Section 5.1 focuses on initial reports (a_A^1 and a_B^1) to test the hypothesis of stochastic behavior (H1). Section 5.2 analyzes the impact of first (a_B^2 and a_A^3) and second (a_B^4 and a_A^5) revisions on group behavior to test for evidence that reports are converging (H2), report errors decrease with revisions (H3), and additional revisions have decreasing impact on both belief convergence and report error (H4).

The next two subsections compare the relative performance of the posterior revision process across the four information structures in Table 1. In Section 5.3, the raw report errors and the baseline errors in each information structure are reported separately before combined into the prediction improvement metric. This is followed by tests that the theoretical ranking predicts relative frequency and size of prediction improvement across the four structures. Section 5.4 provides a joint test of hypothesis H3 to H6 through logistic regressions on the probability of positive prediction improvement and ordinary least squares regressions on the size of the improvement.

[17] Recalculating σ^* for the experimental distribution of $2E, 2H, 3E, 3H$ which differed slightly from the theoretical distribution produced σ^* of $1.1, .55, .3, .15$, preserving the ranking of the theoretical distribution.

5.1 Individual behavior: initial reports

Within 22 rounds of posterior revisions, subjects face 2*4=8 unique initial scenarios.[18] Subjects are therefore making decision within the same scenario 2-4 times. When payoff disturbances are present in subjects' utility computation, subjects will give different responses to the same scenario. Indeed the data shows that a subject who faces the same scenario reports a different probability distribution 81% of the time.

I now provide a maximum likelihood estimate of the noise level. Assuming that subjects use Bayes rule to update the common prior to p_i^1 and that payoff disturbances are distributed $N(0, \sigma)$, the distribution of distance between report a_i^1 and the initial belief p_i^1 follows a stochastic process that depends on σ. Given private information s_i and information structure ψ, a subject's initial belief over Θ is:

$$p_i^1(\theta) = \frac{\psi(\theta|s_i)}{\psi(s_i)}$$

Let d_α denotes the Euclidean distance of α to p_i^1. The probability that player i chooses $a_i^1 = \alpha$ that is d_α away from p_i^1 is

$$Pr(d_\alpha|\sigma) = \frac{F'_\sigma(d_\alpha)}{\sum_{a' \in \Delta_\Theta} F'_\sigma(d_{a'})}$$

Let d_{a_n} denote the distance between the nth initial report and the initial belief from the nth private signal draw. Assuming independence across reports, the likelihood of observing deviations of $\{d_{a_1}, ..., d_{a_n}, ..., d_{a660}\}$ in our experiment is:

$$l(d|\sigma) = \prod_{n=1}^{660} Pr(d_{\alpha_n}|\sigma)$$

The estimation result of this model is given in Table 2. As comparison, we include the estimation result of the standard model rational behavior where $a_i^1 = p_i^1$. The likelihood is much lower, allowing us to reject the hypothesis that initial reports are chosen with $\sigma = 0$.

Estimating each information structure separately, we find that the MLE estimate for the noise parameter σ is 0.37 for 2E, 0.36 for 3E, and 0.29 for 2H and 3H. All estimates of σ indicate that

[18] There are only four information structures and two possible private information (a black ball or a white ball.

	σ	logL
Stochastic	0.334	-1583.58
Deterministic		-2097.516

Table 2: Model estimates of variance of stochastic disturbance.

noise levels are above 0.1, where Expected Improvement are separated by stochastic dominance across the four information structures (Figure 1).

Could risk preferences have been responsible for the deviation between initial report and initial beliefs? If subjects are updating using Bayes rule, no initial reports will indicate equal likelihood or degenerate probabilities. Hence reports of equal likelihood can be taken as evidence of risk aversion and degenerate probabilities could be taken as evidence of risk loving behavior. However, out of 660 initial reports, there were only 21 reports of 50-50 and 11 reports where each of the three states receives at least 30% chance. There were even less degenerate reports; less than 2% of initial reports place 95% or more weight on a single state. This is consistent with the previous findings that the assumption of risk neutrality is appropriate when subjects face small payoffs in the laboratory (Holt and Laury 2002, Harrison et al 2009).

Result 1: *Initial reports indicate stochastic behavior.*

5.2 Group Behavior: Belief Updating

What effect does the posterior revision process have on group behavior? First of all, did subjects update their initial reports? The first row of Table 3 tabulated the fraction of reports that were revised. Column 1 shows that 79% of first revisions (a_A^3 for A and a_B^2 for B) are different from the initial reports (a_A^1, a_B^1). The average distance between first revisions and initial reports is 0.1645 (see first column of the second row).[19]

Column 2 shows that 68% of second revisions (a_A^5 for A and a_B^4 for B) are different from first revisions. The magnitude of the second revision is 0.1189, which is smaller than the first revision.

[19] $(\sum d(a_A^3, a_A^1) + d(a_B^2, a_B^1))/660 = 0.1645$. This is equivalent to changing initial report of equal likelihood (.5, .5) or $(0.33, 0.33, 0.33)$ into $(0.38, 0.62)$ in a 2 state case and $(0.47, 0.25, 0.28)$ in a 3 state case.

Impact of Revision	1st Revision from Initial Report	2nd Revision from 1st Revision	2nd Revision from Initial Report
Revised	79%	68%	86%
Size of revision	0.1645	0.1189	0.2091
Standard Error	(0.0140)	(0.0308)	(0.0148)
Revised towards partner	54.5%	37.4%	62.8%
Decrease in variance	0.1425	0.0308	0.1733
Standard Error	(0.0140)	(0.0096)	(0.0148)
Revised towards p^s	55.5%	35.2%	56.5%
Decrease in error	0.0639	-0.0036	0.0602
Standard Error	(0.0077)	(0.0060)	(0.0085)

Table 3: Impact of revision on belief convergence

14% of initial reports are never revised. Within the entire 5 step posterior revision process, the average distance between each person's final report and their initial report is 0.2091. This suggested that some fraction of the second revisions moved first revisions back towards initial reports.

But are belief are converging? Following Palfrey and Wang (2009)[20], Row 3 of Table 3 count the fraction of time that revisions reduce the distance between the probabilities reported by a pair of subjects reports and Row 4 reports the average reduction of distance from each revision.

The average distance between initial reports ($d(a_A^1, a_B^1)$) is 0.348 with a standard error of 0.0155. Row 4 of Table 3 shows that more than half (54.5%) of first revisions are closer together than initial reports were, but less than half (37.4%) of the second revisions continue to reduce the distance between first revisions. However, on average convergence is positive and significant; Row 5 reports that the average reduction in distance of the first revision, $d(a_A^1, a_B^1) - d(a_A^3, a_B^2)$, is 0.1425 while the average reduction in distance from the second revision, $d(a_A^3, a_B^2) - d(a_A^5, a_B^4)$, is 0.0308. We can reject the hypothesis that beliefs were not converging with the first revision at $p < 0.001$ level and with the second revision at the $p < .005$ level. This means that we can reject a model where

[20]See Table 7 and 8 in their paper. The quantities are not directly comparable since subjects in Palfrey and Wang (2009) are reporting beliefs about strategic play while subjects here are reporting beliefs about an objective true state of the world.

subjects are only following their private information.

Result 2: Subjects are revising reports towards each other

Are beliefs converging to the full information posterior? Row 7 reports the fraction of time that a revision brings report closer to a full information posterior. The average distance of initial reports to the full information posterior is 0.301 with a standard error of 0.008. 55.5% of first revisions reduces the report error, however, only 37.4% of the second revisions continue to reduce this error. A total of 56.5% of final reports are closer to the full information posterior than initial reports.

Row 8 compares the distance to the full information posterior before and after a revision, $d(p^s, a_i^t) - d(p^s, a_i^{t+2}$. First revision reduces report error by 0.0636, however, the second revision for the second revision it is -0.0036. We can therefore reject the hypothesis that first revision are not converging to the full information posterior at the $p < .001$ level but not for the second revision (p=0.545).

Result 3: First revisions decrease report error, however, the impact of second revisions on report error is statistically insignificant.

After both A and B have had one opportunity to revise their reports, the average report error is 0.237[21] and the average distance between these first revisions is 0.2054. This leave a lot of room for improvement for the second revision. However, comparing the total change in the posterior revision process (third column of Table 3) to the change from the first revision (first column), it appears that the second revision contributes little to decreasing report error and reducing within group variance compared to the first revisions. Formally, we test that:

$$d(a_A^1, a_B^1) - d(a_A^3, a_B^2) > d(a_A^3, a_B^2) - d(a_A^5, a_B^4)$$

$$d(p_s, a_A^1) - d(p_s, a_A^3) > d(p_s, a_A^3) - d(p_s, a_A^5) \text{ } and \text{ } d(p_s, a_A^1) - d(p_s, a_A^3) > d(p_s, a_A^3) - d(p_s, a_A^5)$$

Both statements are true with at the $p < 0.001$ level.

Result 4: Additional iteration of revision has decreasing impact on both belief updating and report error.

[21]This is equivalent to predicting $\frac{1}{3}, \frac{2}{3}$ when the true probability is the equal likelihood $\frac{1}{2}, \frac{1}{2}$ or $\frac{1}{2}, .15, .35$ when the true probability is the equal likelihood $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

Information Structure	2E	3E	2H	3H
Initial Report Error ($t = 1$)	0.343	0.368	0.256	0.245
Standard Error	(0.028)	(0.023)	(0.023)	(0.017)
Final Report Error ($t = 5$)	0.275	0.264	0.197	0.208
Standard Error	(0.026)	(0.021)	(0.022)	(0.016)
Total Error Reduction	0.068	0.103	0.059	0.037
Standard Error	(0.028)	(0.025)	(0.025)	(0.018)
Baseline error	0.56	0.54	0.17	0.18
Standard Error	(0.020)	(0.015)	(0.015)	(0.012)
N	85	80	80	85

Table 4: Average distance of reports to full information posterior

5.3 Aggregation Performance: Prediction Improvement

Four information structures are tested in this experiment, 2E and 2H with two states of the world, and 3E and 3H with three states of the world (Table 1). By solving for the maximum noise level tolerated by each information structure, we arrived at an ex-ante difficulty ranking of $2E < 3E < 2H < 3H$, where 3H is the structure where reports are least likely to beat the prior. This section tests if actual prediction improvement follows the ranking predicted by the noise tolerance model.[22]

Row 1 of Table 4 reports the Euclidean distance between the full information posterior and a_A^1, the first report in the posterior revision process. Row 3 reports this distance for the last report in the process. The report error in 2E and 3E is initially higher ($p < 0.05$) than the error in 2H and 3H and remain higher at the end of the process ($p < 0.05$). However, the difference of report error of 2E from 3E is not statistically significant and neither is 2H from 3H ($p > 0.5$ for both). Row 5 shows the reduction in error from the 5 step process; consistent with H3, this reduction is positive and significant in all the information structure ($p < 0.05$). Row 7 reports the distance of the common prior to the full information posterior. Again, 2E and 3E, despite having different numbers of states, have similar baseline errors which are significantly larger than the baseline errors of 2H and 3H.

[22] a_B^1, the report that is never observed by A, is not considered from this point onwards.

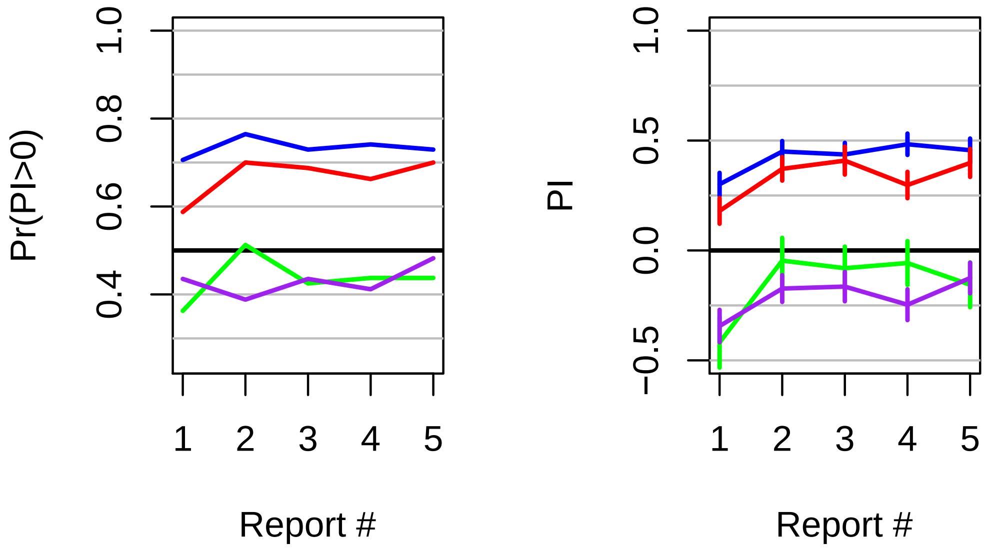

Figure 2: On x: Frequency (L) and size (R) of PI. On y: revision step (t). Blue: 2E, Red: 3E, Green: 2H, Purple: 3H

Let $n_\psi \in \{1, .., N_\psi\}$ refer to the nth posterior revision process for information structure $\psi \in \{2E, 3E, 2H, 3H\}$. Let p_{sn_ψ} and α_{n_ψ} respectively be the full information posterior from the signals drawn at n_ψ and the output at the tth iteration of the posterior revision process. Let p_ψ be the common prior associated with information structure ψ.

In order to evaluate the error reduction resulting from this aggregation process relative to errors without the process, the reduction in error is normalized by the average baseline error of ψ (Table 4 Row 4).

$$PI(n, \psi) = \frac{d(p_{sn_\psi}, p_\psi) - d(p_{sn_\psi}, \alpha_{n_\psi})}{\sum\limits_{n_\psi} d(p_{sn_\psi}, p_\psi)/N_\psi}$$

Figure 2 illustrates PI in each information structure as a function of the posterior revision step (t). The fraction of reports that beats the prior is on the left panel while the average size of prediction improvement and the standard error bars are presented on the right panel.

p-value	$2E > 3E$	$3E > 2H$	$2H > 3H$
fraction of $PI > 0$			
$t = 2$	0.11	0.00	0.01
$t = 5$	0.32	0.00	0.76
size of PI			
$t = 2$	0.221	0.006	0.229
$t = 5$	0.311	0.001	0.569

Table 5: One-sided binomial test (t-test) of fraction (size) of prediction improvement .

An Outsider is better off abandoning his prior for the output of the aggregation process if $PI(n, \psi) > 0$ (Definition 3). $PI(n, \psi) = 1$ indicates perfect aggregation. When no noise is present in this process, the report at step 2 onwards will be the full information posterior, and hence $PI(n, \psi) = 1$ in all information structure at $t \geq 2$ (Theorem 2.2). However, the noise tolerance model predicts that at $t \geq 2$, prediction improvement will be most frequent in 2E, then in 3E, then 2H, and finally least frequent in 3H.

Row 1 of Table 5 reports the pvalue of a one-sided binomial test that the frequency of prediction improvement at $t = 2$ is higher in 2E compared to 3E (Column 1), 3E compared to 2H (Column 2), and 2H compared to 3H (Column 3). The pvalue for $2E > 2H$, $2E > 3H$, and $3E > 3H$ are all < 0.01. The empirical ranking of $2E > 3E > 2H > 3H$ is consistent with the difficulty ranking and not with a model where the frequency of prediction improvement does not depend on the information structure. Row 2 reports the same test for $t = 5$. Although the differences between the two 2 state structures (2E,2H) and two 3 states structures (3E,3H) remain significant ($p < 0.001$), the difference between the two easy structures (2E,3E) and two hard structures (2H,3H) are no longer signficant.

Result 5: There is strong evidence that frequency of prediction improvement follows the ex-ante ranking of $2E \geq 3E > 2H > 3H$. for $t = 2$. The evidence for $t = 5$ is mixed: the ranking implied by the data is $2E \geq 3E > 3H \geq 2H$.

The stochastic dominance in the EI graph when $\sigma > 0.1$ suggests that the size of prediction improvement may also follow the difficulty ranking. Row 3 of Table 5 reports the pvalue of a t-test that the size of prediction improvement at $t = 2$ is larger in 2E compared to 3E (Column 1), and so on. Row 4 reports the same test for $t = 5$. The pvalue for t-test of $2E > 2H$, $2E > 3H$, and $3E > 3H$ for both $t = 2$ and $t = 5$ are less than 0.01. The size of prediction improvement at $t = 2$ follows the ex-ante ranking, but the difference between the pair 2E-3E and the pair 2H-3H is not statistically significant. The relative size of improvements constant remain for $t = 5$ except for the improvement in 3H, which is unexpected.

Result 6: At t=2, relative size of prediction improvement in the experiment is $2E \geq 3E > 2H \geq 3H$, which is consistent with the difficulty ranking. Again the evidence is mixed for $t = 5$ since the ranking implied by the experimental data is $2E \geq 3E > 3H \geq 2H$.

How does the difficulty ranking from this model of stochastic action compare to alternative models? In a model where subjects are perfectly rational, PI for all $t \geq 2$ is 1 in all information structures. If subjects ignore their private information, reports are centered around the prior and PI will remain around 0. Both imply a size ranking of $2H = 3E = 2H = 3H$, which is rejected by Result 5 and 6. A model where aggregation success depends on the number of possible states implies a ranking of $2E = 2H < 3E = 3H$, which is also rejected by Result 5 and 6. If subjects are following only their private information, beliefs will remain unrevised and the fraction of reports closer to p_s than the prior depends on the fraction of identical private signal draw (both A and B observe Black or White). This model results in the same difficulty ranking as the model of stochastic action ($2E > 3E > 2H > 3H$); this is expected since the slow updating of posterior is an implication of the stochastic action model. However Result 2,3, and 4 indicates that subjects are updating their posterior in response to their partners, thus rejecting the private information model.

5.4 Joint Test

As a joint test of Result 3,4,5, and 6, Figure 3 provide logistic regression on the probability of prediction improvement (Model 1-3) and OLS models (Model 4-6) on the size of improvement. Model 1 and 4 include a dummy variable for each of the information structure. The rest of the Model use the variable *Difficulty* where 1 indicates 2E and 4 indicates 3H to provide an easily interpretable interaction term for the difficulty ranking. *Same private signal* is 1 when the pair receive the same private signal and 0 otherwise. *First Revision* is a dummy variable which indicates

A and B's first opportunity to revise their report (a_B^2 and a_A^3). Second Revision equals 1 for a_B^4 and a_A^5. The intercept can therefore be interpreted as initial reports ($t = 1$) when private signals are different. The robust standard errors are clustered on individuals.

The increasingly negative coefficient on harder ranked information structures and the negative coefficient on *Difficulty* provide evidence that difficulty ranking predicts aggregation success. The coefficient on *Same private signal* is significant and positive across all models. When the realization of private signals are the same across all subjects, the slow updating that resulted from stochastic action has less of an effect on prediction improvement since an individual's initial belief is already closer to the full information posterior than the prior. The interaction of *Difficulty* and *Same private signal* in Model 3 reveals when private signals are the same, the posterior revision process fails to beat the common prior more frequently in harder ranked information structure. When private signals are different, the process fail equally frequent in all information structures, however the loss compared to the common prior is larger in harder ranked information structure (Model 6).

The coefficient on *First Revision* is significant and positive across all models, supporting H3. The interaction term in Model 3 shows that first revisions is less succesful in increasing the frequency of prediction improvement in harder ranked structures. However, this is not true of the size of prediction improvement (Model 6). Consistent with Result 4, the coefficient on *Second Revision* is not larger than *First Revision* across all six models, which supports H4. There is some evidence of learning: Model 6 indicates that subjects seem to improve their performance in harder information structure as they gain experience. The size of the effect is however quite small.

6 Discussion and Conclusion

Experimental evidence has long shown that mechanisms which are theoretically expected to fully aggregate dispersed private information (such as asset markets, polls, and other forecasting systems) in reality produce probability distributions that are only approximately converging towards the full information posterior. By explicitly modeling the stochastic component of individual behavior in the posterior revision process, I show the dependence of a basic unit of the group inference process on the underlying information structure. To measure the contribution of the aggregation process, I introduce a metric of prediction improvement which captures the fraction of the possible error

	Logistic Regression Y= (PI>0)			Ordinary Least Squares Y = PI		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	-2.381 ***	-1.607 ***	-3.893 ***	-0.812 ***	-0.586 ***	-0.284 *
	0.358	0.352	0.909	0.068	0.071	0.159
3E	-0.150			-0.016		
	0.259			0.039		
2H	-0.354			-0.166 **		
	0.317			0.081		
3H	-1.383 ***			-0.428 ***		
	0.260			0.067		
Difficulty		-0.454 ***	0.372		-0.144 ***	-0.260 ***
		0.085	0.281		0.024	0.072
Same private signals	3.809 ***	3.706 ***	5.514 ***	1.242 ***	1.228 ***	1.318 ***
	0.178	0.166	0.497	0.066	0.070	0.076
First Revision	0.578 ***	0.572 ***	1.818 ***	0.219 ***	0.219 ***	0.160 **
	0.184	0.183	0.689	0.032	0.032	0.074
Second Revision	0.427 ***	0.422 ***	0.960 **	0.206 ***	0.206 ***	0.169 ***
	0.137	0.135	0.393	0.032	0.031	0.061
Round	0.038 ***	0.033 **	0.087	0.011 ***	0.010 ***	-0.018 *
	0.015	0.013	0.060	0.004	0.004	0.010
Same privsig *Difficulty			-0.668 ***			-0.035
			0.163			0.037
FirstRevision*Difficulty			-0.433 *			0.024
			0.220			0.029
SecondRevision*Difficulty			-0.188			0.015
			0.134			0.029
Round*Difficulty			-0.017			0.011 **
			0.018			0.004
N	1972	1974	1970	1972	1974	1970
Likelihood Ratio	1171.4	1159.3	1186.9			
R2	0.598	0.594	0.604	0.437	0.434	0.434

* p<.1, ** p<.05, *** p<.01

Figure 3: Frequency and size of prediction improvement.

reduction that can be achieved in a given information structure. I then provide a method to rank a set of information structures over expected prediction improvement.

I derive the difficulty rankings for four experimental environments using this method and test the rankings with a laboratory experiment where subjects participate in the posterior revision process with a random partner. The experimental evidence is consistent with conjectures from the model of noisy Bayesian updating. First, subjects are frequently revising their beliefs in the direction of their partner's reports. Second, first revisions decrease initial report error. Third, second revisions has little impact on both belief convergence and reduction of report error. Fourth, relative prediction improvement in the four information structures is consistent with the difficulty ranking derived from the noise tolerance approach. This is true for both the frequency and size of prediction improvement. However, there is little difference between the information structure pair 2E-3E and 2H-3H.

The result suggests that analyzing the tolerance of mechanisms and environments for stochastic noise may be a useful line of research. When noise is introduced into behavior, details of the environment that did not matter under models of deterministic behavior start to matter. This approach may yield a better understanding of the relationship between mechanisms and environments, which may help provide a way to anticipate the contribution of a mechanism when implemented in a particular setting.

7 Appendix

Lemma 2.1

Proof. Let p_i^t be the vector of i's belief over Θ. Myopic payoff optimizer i chooses a_i such that a_i^t maximizes $argmax_{a_i^t} Q(a_i^t|\tilde{\theta})$. The expected payoff for reporting a_i^t when belief is p_i^t is $\mathbb{E}_{p_i^t} Q(a_i^t) = \sum_\theta p_i^t(\theta)[1-(a_i^t(\theta)-p_i^t(\theta))^2]$; this is maximized when $a_i^t = p_i^t$. Hence $a_i^* = (p_i^1, ..p_i^T)$. □

Lemma 3.2

Proof. By Selten, 1998, $\mathbb{E}_{p_i^t} Q(p_i^t) - \mathbb{E}_{p_i^t} Q(\alpha) = d(p_i^t, \alpha)$. This simplifies Eq.1 into:

$$R(\alpha|p_i^t) = \left\{(\epsilon_\alpha, \epsilon_p) \in \mathbb{R}^2 : \epsilon_\alpha - \epsilon_p \geq d(p_i^t, \alpha)\right\}$$

Let $\epsilon = \epsilon_\alpha - \epsilon_p$. Since ϵ_α and ϵ_p are iid distributed with mean 0 and variance σ, ϵ is distributed with mean 0 and variance $\sigma + \sigma = 2\sigma$.[23] Substituting this we arrive at $R(\alpha|p_i^t) = \left\{\epsilon \in \mathbb{R} : \epsilon \geq d(p_i^t, \alpha)\right\}$. Therefore

$$\int\limits_{R(a|p_i^t)} dF_\sigma(\epsilon) = F'_\sigma(d(p_i^t, \alpha))$$

Substituting this into Eq 2, we arrive at $Pr_\sigma(\alpha|p_i^t)$ as defined above.

(ii) Let $\delta, \delta' \in \mathbb{R}_+$ where $\delta > \delta'$. When $0 < \sigma < \infty$, $F'_\sigma(\delta) > F'_\sigma(\delta')$. Since the numerator of $Pr_\sigma(\alpha|p_i^t)$ decreasing in δ while the denominator is constant, $Pr_\sigma(\alpha|p_i^t)$ must be decreasing in $d(p_i^t, \alpha)$. □

Proposition 3.3

Proof. When ϵ is distributed with $\mathcal{N}(0, \sigma)$, $F'(\delta)$ is $\frac{1}{\pi}\int_{\frac{\delta}{\sqrt{2}\sigma}}^\infty e^{-t^2} dt$ and its derivative with respect to δ is $-\frac{e^{-(\frac{\delta}{\sqrt{2}\sigma})^2}}{\sqrt{2}\pi\sigma}$. Hence,

$$\frac{F'_\sigma(\delta)}{F'_{\sigma+k}(\delta)} = \frac{\frac{1}{\pi}\int_{\frac{\delta}{\sqrt{2}\sigma}}^\infty e^{-t^2} dt}{\frac{1}{\pi}\int_{\frac{\delta}{\sqrt{2}(\sigma+k)}}^\infty e^{-t^2} dt} \tag{3}$$

[23]See Thm 4.2.12 Casella and Berger

Since $\delta, k \in \mathbb{R}_+$ implies $\frac{\delta}{\sqrt{2}\sigma} > \frac{\delta}{\sqrt{2}(\sigma+k)}$, Eq. 3 must be less than 1.

Taking the derivate of Eq. 3 with respect to δ we get

$$\frac{-\frac{e^{-(\frac{\delta}{\sqrt{2}\sigma})^2}}{\sqrt{2}\pi\sigma}\int_{\frac{\delta}{\sqrt{2}(\sigma+k)}}^{\infty} e^{-t^2}dt + \frac{e^{-(\frac{\delta}{\sqrt{2}(\sigma+k)})^2}}{\sqrt{2}\pi(\sigma+k)}\int_{\frac{\delta}{\sqrt{2}\sigma}}^{\infty} e^{-t^2}dt}{(\int_{\frac{\delta}{\sqrt{2}(\sigma+k)}}^{\infty} e^{-t^2}dt)^2}$$

Since the denominator is positive, the sign of the numerator (Eq.4) determines the derivative.

$$\frac{1}{\sqrt{2}\pi}[\frac{e^{-(\frac{\delta}{\sqrt{2}(\sigma+k)})^2}}{\sigma+k}\int_{\frac{\delta}{\sqrt{2}\sigma}}^{\infty} e^{-t^2}dt - \frac{e^{-(\frac{\delta}{\sqrt{2}\sigma})^2}}{\sigma}\int_{\frac{\delta}{\sqrt{2}(\sigma+k)}}^{\infty} e^{-t^2}dt] \tag{4}$$

We show by contradiction that the numerator is negative. Suppose this term is positive. Then:

$$\frac{e^{-(\frac{\delta}{\sqrt{2}\sigma})^2}}{e^{-(\frac{\delta}{\sqrt{2}(\sigma+k)})^2}}\frac{\sigma+k}{\sigma} < \frac{\int_{\frac{\delta}{\sqrt{2}\sigma}}^{\infty} e^{-t^2}dt}{\int_{\frac{\delta}{\sqrt{2}(\sigma+k)}}^{\infty} e^{-t^2}dt}$$

Which is a contradiction since the right hand side of the equation is less than 1 and the left hand side is greater than 1 since $e^{-(\frac{\delta}{\sqrt{2}\sigma})^2} > e^{-(\frac{\delta}{\sqrt{2}(\sigma+k)})^2}$. Hence Eq 4 must be negative, and Eq 3 is therefore decreasing in δ. □

Lemma 3.4

Proof. (i) As a shorthand we write $d_\alpha = d(p_i, \alpha) < d_\beta = d(p_i, \beta)$, and $d_a = d(p_i, a)$. $\sigma' = \sigma + k$ for some $k > 0$. By Lemma 3.2 we write:

$$\frac{F'_\sigma(d_a)}{\sum_a F'_\sigma(d_a)} > \frac{F'_{\sigma+k}(d_a)}{\sum_{a'} F'_{\sigma+k}(d_{a'})} \; and \; \frac{F'_\sigma(d_b)}{\sum_{a'} F'_\sigma(d_{a'})} < \frac{F'_{\sigma+k}(d_b)}{\sum_{a'} F'_{\sigma+k}(d_{a'})}$$

Putting together and simplifying:

$$\frac{F'_\sigma(d_a)}{F'_{\sigma+k}(d_a)} > \frac{\sum_{a'} F'_\sigma(d_{a'})}{\sum_{a'} F'_{\sigma+k}(d_{a'})} > \frac{F'_\sigma(d_b)}{F_{\sigma+k}(d_b)}$$

From Prop 3.3(ii) we know $\frac{F'_\sigma(\delta)}{F'_{\sigma+k}(\delta)}$ is decreasing in δ, which implies that $d_b > d_a$.

(ii) Let $c(\sigma, \sigma'|\alpha, p_i^t) = Pr_\sigma(\alpha|p_i^t) - Pr_{\sigma'}(\alpha|p_i^t)$. Partition Δ_Θ into $L+ = \{\alpha|c(\sigma, \sigma'|\alpha, p_i^t) \geq 0\}$ and $L- = \{\beta|c(\sigma, \sigma'|\alpha, p_i^t) < 0\}$. By (i), $d_L \equiv \max_{\alpha \in L+} d(p_i^t, \alpha) < d_H \equiv \min_{\beta \in L-} d(p_i^t, \beta)$. Since $\sum_\alpha Pr_\sigma(\alpha|p_i^t) = 1$, $\sum_\alpha c(\sigma, \sigma'|\alpha, p_i^t)$ must be 0, which implies $\sum_{\alpha \in L+} c(\sigma, \sigma'|\alpha, p_i^t) = -\sum_{\alpha \in L-} c(\sigma, \sigma'|\alpha, p_i^t) = C \in \mathbb{R}_+$. Therefore $\sum_{\alpha \in \Delta_\Theta} c(\sigma, \sigma'|\alpha, p_i^t)d(p_i^t, \alpha)$ must be smaller than $d_L k + d_H(-k) < 0$, which implies (ii). □

Proposition 7.1.

$$\lim_{N \Rightarrow \infty} \bar{P}I_{PR^T}(D^N, \psi) = \int EI_{PR^T}(\psi|\sigma)dg(\sigma)$$

The probability of achieving prediction improvement in a large sample is therefore equivalent to the probability of drawing noise level lower than σ^*:

$$\lim_{N \Rightarrow \infty} Pr(\bar{P}I(D^N, \psi) > 0) = G(\sigma^*)$$

Proof. Let N_s be the number of times s occured in D^N. As $N \to \infty$, by Law of Large Numbers $\frac{N_s}{N} \to \psi(s)$ and hence $BaseErr(D^N, \psi) \to \mathbb{E}_s BaseErr(\psi) = \sum_{s \in S} \psi(s)d(p_s, p)$. Since σ is redrawn each time from distribution $g(.)$, as $N \to \infty$, $\frac{1}{N}\sum_{n=1}^{N} ErrR(s^n) = \int \sum_s \psi(s)\mathbb{E}Err_{PR^T}(s,\sigma)dg(\sigma)$. Hence $\lim_{N \Rightarrow \infty} \bar{P}I_{PR^T}(D^N, \psi) = \int EI_{PR^T}(\psi|\sigma)dg(\sigma)$. □

Proposition 3.1

Proof. As before we assume σ is distributed with cumulative density function $G(.)$. Since $\mathbb{E}_s[\mathbb{E}_\alpha[d(p_s, \alpha)|\sigma, s]]$ is monotonically increasing in σ by assumption, $Pr(\int EI_{PR^T}(\psi|\sigma)dg(\sigma) > 0) = Pr(\sigma < \sigma^*) = G(\sigma^*)$. By Proposition 7.1 $\lim_{N \Rightarrow \infty} PI_{PR^T}(D^N, \psi) = \int EI_{PR^T}(\psi|\sigma)dg(\sigma)$, therefore $Pr(PI(D^N, \psi) > 0) \Rightarrow G(\sigma^*)$ as $N \to \infty$. Hence $\sigma_1^* > \sigma_2^*$ implies $\lim_{N \Rightarrow \infty} Pr(\bar{P}I(D^N, \psi_1) > 0) > \lim_{N \Rightarrow \infty} Pr(\bar{P}I(D^N, \psi_2) > 0)$ □

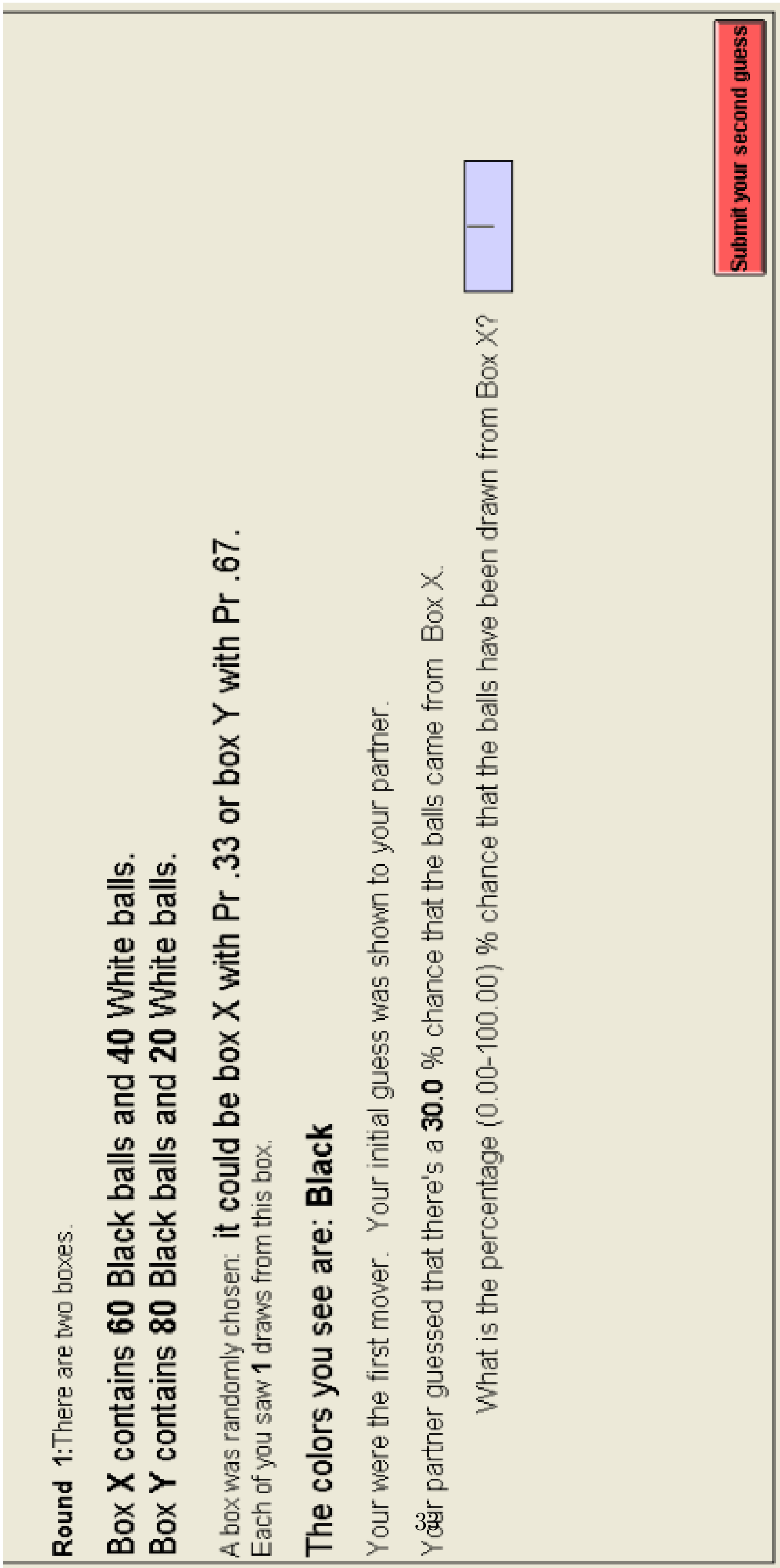

Figure 4: Screenshot of posterior revision software (ZTree)

References

Andersen, S., Fountain, J. (2009), Harrison, G., annd E. Rustrom, Estimating Subjective Probabilities, Working Paper 09-01,

Aumann, R. (1976),Agreeing to Disagree. *Annals of Statistics*, 4, 1236-39.

Ballinger, P. and N. Wilcox (1997), Decisions, error and heterogeneity. *Economic Journal* 107, 1090-1105

Brier, G. (1950), Verification of Forecasts Expressed In Terms of Probability. *Monthly Weather Review*, 78(1):13

Camerer, C. (1989), An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty* 2, 61-104

Camerer, C. (1995): Individual decision making. In: Kagel, J.H. & Roth, A.E. (eds.): Handbook of Experimental Economics, 587-703. Princeton, NJ: Princeton University Press.

Camerer, C. (1998): Can Asset Markets Be Manipulated? A Field Experiment with Racetrack Betting. *Journal of Political Economy* 106:3,, 457-482

Chen,Y., Mullen, T. and C.H. Chu (2006): An in-depth analysis of information markets with aggregate uncertainty. *Electronic Commerce Research*, 6: 2, p.201-221

Diamond, D. W. and R. E. Verrecchia (1981): Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics* 9, p.221-235

Feigenbaum, J., Fortnow, L., Pennock,D. and R. Sami (2005), Computation in a distributed information market. *Theoretical Computer Science*, 343(1-2):114-132

Forsythe,R., Nelson,F., Neumann, G., and J. Wright (1992), Anatomy of an Experimental Political Stock Market. *American Economic Review*, 82(5):1142-1161

Geanakoplos, J. and H. M. Polemarchakis (1982), "We Can't Disagree Forever." *Journal of Economic Theory*, 28(1): 192-200

Geanakoplos, J. (1989), Game Theory without Partitions, and Applications to Speculation and Consensus, *Cowles Foundation Discussion Paper* 914, Yale University, 1989.

Goeree, J.K., Palfrey, T.R., Rogers, B.W., and R.D McKelvey (2007) Self-Correcting Information Cascades. *Review of Economic Studies*, 74(3): 733-762

Grus and Ledyard (2005), experimental data.

Hey, J.D. and C. Orme (1994), Investigating generalisations of expected utility theory using experimental data, *Econometrica* 62, 1291-1326

Hey, J.D. (2005), "Why We Should Not Be Silent About Noise.", *Experimental Economics*, 8.4, 325-345

Healy, P, J. Ledyard, S.Linardi, and R.Lowery (2007), "Information Aggregation Mechanisms and Environmental Complexity", Working Paper.

Holt, C. and S. Laury (2002), Risk Aversion and Incentive Effects in Lottery Choices, *American Economic Review*, 92(5): 1644-55

Hung, A.A and C. R. Plott (2001), Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions, *American Economic Review*, 91(5): 1508-1520

Hung, A.A and J. Dominitz (2009), Empirical Models of Discrete Choice and Belief Updating in Observational Learning Experiments, *Journal of Economic Behavior and Organization*, 69(2):94-109

Luce, R.D (1959), Individual choice behavior: a theoretical analysis. New York: Wiley.

Ledyard,J., Hanson, R. and T. Ishikida (2009), An experimental test of combinatorial information markets, *Journal of Economic Behavior and Organization*, 69(2): 182-189

Lundholm, R. (1991), What affects the efficiency of a market? Some answers from the laboratory, *The Accounting Review* 66, pp. 486-515

McKelvey, R. and T.R. Palfrey (1995): Quantal Response Equilibria for Normal Form Games, *Games and Economic Behavior*, 10, pp.6-35

McKelvey, R. D. and T. R. Palfrey (1998), Quantal Response Equilibrium for Extensive Form Games, *Experimental Economics*, 1(1), 9-41.

McKelvey, R and T.Page (1986), Common Knowledge, Consensus and Aggregate Information, *Econometrica*, 54(1), 109-127.

McKelvey, R. and T. Page (1990), Public and Private Information: an Experimental Study of Information Pooling, *Econometrica*, 58(6), 1321-1339.

Nielsen, L., A. Brandenberger, J. Geanakoplos, R. McKelvey and T. Page (1990), Common Knowledge of an Aggregate of Expectations, *Econometrica*, 58, 1235-1239.

Offerman, T., Sonnemans J., van de Kuilen G. and P.P. Wakker (2009) A Truth-Serum for Non-

Bayesians: Correcting Proper Scoring Rules for Risk Attitudes, *Review of Economic Studies* 76, 1461-1489

Plott, C. and S. Sunder (1988), Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets, *Econometrica*, 56(5):1085-1118

Plott, C.R. (2000), Markets as information gathering tools, *Southern Economic Journal*, 67(1), pp. 215

Plott, C.R. and K. Roust (2006), Design and Testing of Information Aggregation Mechanisms: A Two-Stage Parimutuel IAM, Social Science Working Paper

Selten, S (1998), Axiomatic Characterization of the Quadratic Scoring Rule, *Experimental Economics*, 1(1), pp. 43-61

Starmer, C. and Sugden, R. (1989), Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty* 2, 159-178

Sunder, S. (1995): Experimental Asset Markets: a Survey, in J. Kagel and A. E. Roth (eds) Handbook of Experimental Economics, Princeton, NJ. 13

Toda, M. (1963), Measurement of Subjective Probability Distribution, Report 3, State College, Pennsylvania, Institute for Research, Division of Mathematical Psychology.

Weber, R. A. (2001), Behavior and learning in the dirty faces game. *Experimental Economics*, 4 (3), 229-242.

Winkler, R.L. (1969), Scoring Rules and the Evaluation of Probability Assessors, *Journal of the American Statistical Association* 64, 10731078

Zimper, A. (2009), Half empty, half full and why we can agree to disagree forever, *Journal of Economic Behavior and Organization*, 71(4): 283-299