Diverging Opinions*

James Andreoni
University of California, San Diego

Tymofiy Mylovanov
Penn State University

April 28, 2010

Abstract

People often see the same evidence but draw opposite conclusions, becoming increasingly polarized over time. Prior discussions of this puzzle have focused on the failure of Bayesian reasoning or on the possibility of non-common priors. By contrast, we offer a model of diverging opinions in a Bayesian environment with common priors. In our model, people hold different private information and disagreement arises and continually increases as additional public information becomes available. We support our model with an experiment and argue for a useful reframing of the puzzle: Why can't individuals share their private information on all dimensions of an issue, but instead concede to agree to disagree?

Keywords: polarization of opinions, Bayesian beliefs, updating of beliefs

*We are deeply grateful to Nageeb Ali, Julia L. Evans, Navin Kartik, Georg Noldeke, Matthew Rabin, Joel Sobel, Leeat Yariv, and Muhamet Yildiz for the many helpful comments and conversations. We gratefully acknowledge financial support from the National Science Foundation and from the German Science Foundation (DFG) through SFB/TR 15 "Governance and the Efficiency of Economic Systems." We are thankful to Ben Cowan, Megan Ritz, and Justin Rao for expert research assistance. A significant part of this paper has been written while the second author was at the University of Bonn.

1 Introduction

We see many examples in the world around us in which individuals see the same information and draw opposite conclusions, and additional information only results in increased polarity. This is true even among educated experts, such as on a divided supreme court, or among academics locked in ideological struggles. We observe diverging opinions about sundry issues, such as gun control, social welfare benefits, affirmative action, the war in Iraq, and the death penalty.

How can two people see the same information and draw opposite conclusions? How can additional public information persistently draw two people into a stronger disagreement? Presumably, new information should decrease uncertainty about the optimal policy and move individual opinions closer to each other. Thus, diverging opinions seem to contradict theories of rational inference.[1]

The existing explanations of this phenomenon typically rest on one of two assumptions. First, diverging opinions may result from psychological biases, which can be formalized as non-Bayesian updating (see, e.g., Rabin and Shrag [40]). Second, the potential for disagreement can be introduced through different individual prior beliefs about the underlying state of nature (Dixit and Weibull [11]).

In this paper we provide an alternative explanation. We argue that diverging opinions may come from rational actors who filter information through a different model, and that these different models may themselves be the result of Bayesian inference with common initial priors. For instance, perhaps some personal experiences—or even introspection—have led one person to believe that people are basically honest, while different experiences have lead another person to believe others are basically opportunistic. When welfare rolls swell dramatically after an expansion of eligibility rules, the person who feels people are honest may see that this demonstrates the need, while the person who believes people are opportunistic may see that it demonstrates the consequences of moral hazard. Given their models of the world, developed rationally from personal experience, each may be updating appropriately given the evidence. Diverging views about gun laws could also fit this framework. People supporting liberal gun laws may believe, based on private evidence, that the expanding gun ownership will deter violence, while gun opponents may believe that expanding ownership will lead to violence. A tragic school shooting will provide each side with evidence they need to strengthen their view that their favored changes would improve

[1]Certainly, divergence of opinions can occur as a result of conflicting preferences. In this paper, we focus on situations in which people have same preference over the optimal course of action and divergence of opinions is potentially more intriguing.

society.

A way to formalize these ideas is to assume that information about the true state of nature has more than one dimension. For instance, one dimension of the information tells the person which model is more likely (honest versus opportunistic). The beliefs about this model color the interpretation of later information (rise in welfare rolls), and this leads to different conclusions (welfare is beneficial versus welfare is detrimental).

This paper will present a simple theoretical environment that incorporates these ideas and in which information has two dimensions. We show that situations can arise in which people with different private information about one dimension can view the same public information about the other dimension and come to opposite conclusions. Furthermore, as the amount of public information increases, the agents disagree with probability one: An infinite sequence of public signals removes all uncertainty about one of the dimensions and exaggerates the effect of the disagreement about the other dimension on the belief about the optimal action.[2]

How relevant is our model as an explanation of diverging opinions? There are two potential criticisms. First, numerous studies have documented that individuals are notoriously bad at Bayesian updating. Hence, building our model from the assumption that individuals are Bayesian updaters may be problematic. Second, one might argue that the real puzzle is not the emergence of disagreement among individuals but rather its persistence after the disagreement becomes common knowledge. Since the agents in our model share a common prior about the state of nature and the distribution of signals, the disagreement should disappear if individuals were able to communicate (Aumann [2], Geanakoplos and Polemarchakis [19]).

To address these issues, we test our model with a laboratory experiment. We conduct the experimental condition that mirrors the situation considered by the theory. In addition, we also conduct a control treatment in which the private information is followed by the sequence of public signals providing information about *both* dimensions and the theory prescribes that the agents' posterior beliefs should converge with certainty.

As predicted, we observe that disagreement is more likely to occur when public information is drawn on only one dimension, rather than on both, and beliefs are likely to become more polarized with additional information. Furthermore, when

[2]Our model satisfies the assumptions in Blackwell and Dubins [4] whose results imply that the agents' posterior beliefs about the distribution of future *signals* (on one of the dimensions) converge. It is the agents' posterior beliefs about the optimal *action* that fail to converge.

public information is on only one dimension, the disagreement tends to be *larger* for the subjects whose behavior is more *consistent* with Bayesian updating. That is, the failure of subjects to adhere to Bayesian updating does not exacerbate disagreement. This observation addresses the first criticism mentioned above.

In addition, in the final round in each session, we provide subjects with information about the actions of others in the previous rounds. If the subjects believe others to be rational, this information is sufficient to infer their private information and should eliminate disagreement.[3] Surprisingly, we find that, despite giving subjects the common information they need to reach full agreement, a sizable minority of our subjects maintain their opposing views. In particular, this suggests that the persistence of disagreement might be caused by reasons *other* than non-common priors.

Notice that the common prior assumption has two methodological advantages in our context. First, it makes it easier to control subjects' information in the laboratory: in the experiment, we simply show the subjects the urns, representing possible states, from which signals are drawn. The alternative of non-common priors would require either deception or selection of subjects based on the beliefs acquired outside the lab. Second, the common prior assumption allows for a clear test of the common knowledge of rationality among subjects. Without this assumption, we would not be able to interpret the observation that disagreement may continue to persist despite sufficient information that should eliminate it.

In summary, the results in this paper suggest the following story. The initial disagreement may arise and continue to increase in the light of commonly observed new evidence because the public has different models of the world. Furthermore, as confirmed by our experiment, even if there is sufficient information to infer the models of others, the public may fail to do so and the disagreement can persist and become common knowledge.

If our multi-dimensional-information theory is correct, it changes the focus of the puzzle about common knowledge of disagreement. Rather than ask why people may be non-Bayesian, ignore relevant information, or have non-common priors, as in the previous literature, we must instead ask why they may have different models and why they don't share all the information that shaped their individual models? That is, what keeps them from communicating enough to draw their "world views" together?

Although we do not have a formal model addressing the failure of agreement in our experiment, we can envision several alternative answers. First, people may doubt

[3]Under the alternative assumption of non-common priors, the subjects would be expected to continue to disagree and testing for common knowledge of agreement would not make much sense.

the rationality of others. Second, people may fail to make a correct inference from communication and the information about the actions of others. Finally, people may forget, misinterpret, or doubt the experimenter's description of the signal structure of other subjects, thus, discounting the information received from them. Which of these explanations, if any, is correct is the empirical question requiring further experimental work.

Outside of the laboratory, the set of reasons that may prevent people from sharing their models of the world is richer. First, there may be barriers or biases that keep people from conveying what they know. For instance, people may face technological or cognitive limitations to information sharing or processing, especially if the information is vast. Moreover, some private information may be difficult to communicate, such as that coming from introspection, personal experience, trusted advisors, religious principles, family heritage, or cultural identity.

Second, there may be barriers or biases that keep people from seeing information that others are sharing. These again may be technological, as above, but the barriers may also be strategic. There may be vested interests or ulterior motives, orthogonal to the issue of disagreement, that make opponents skeptical of any private information offered. Financial, reputational, ideological, or social incentives could cause people to willfully misrepresent beliefs.[4]

A third and potentially most interesting reason that people may not share their individual models is that there may be a basic unawareness that there is anything at all to share. There are, in fact, many things that people come to know, believe, and understand without awareness that they have done so. The psychologists' name for such unawareness is *implicit learning* (see, e.g., Axel Cleeremans [8]). In particular, it is possible that people who believe others are basically trustworthy, or that gun owners are overwhelmingly responsible, may be unaware of how this belief was formed, why they hold it, and are unable to explain this belief to others. People may instead simply agree to disagree.

There is a growing theoretical literature that studies reasons for divergent opin-

[4] Another reason that is related to the first two is that the individuals' models may be functions of preferences rather than facts. For instance, suppose there are two dimensions of information, say A and B, and two individuals who have received different information on the two dimensions. If one person's tastes make dimension A irrelevant to his decision, and for the other B is irrelevant, sharing their beliefs about the optimal decision may not lead to agreement. Rather, they would also need to share information on their preferences, and this may be difficult to accomplish for reasons like those stated in the last two paragraphs. This alternative was suggested to us by Joel Sobel, and is similar to the ideas presented in Santos-Pinto and Sobel [41].

ions. One possibility is that people pay preferential attention to evidence that supports their view and be overly dismissive of evidence that contradicts it; this phenomenon is called *confirmatory bias*.[5] Rabin and Schrag [40] capture this idea in a non-Bayesian model with uncertainty about a one-dimensional state of nature in which agents may *misinterpret* signals. An alternative approach is to assume that people differ in the weight they place on importance of public evidence; it has been formalized, e.g., in Gerber and Green [20] and Harris and Raviv [21].

Dixit and Weibull [11] offer a model with one-dimensional uncertainty about the state of nature and *non-common priors*. In their model, inference about the optimal policy can diverge when signals about policy outcomes are non-monotone in the policy choice, that is, the same signal can have multiple interpretations. The substantive difference is that individuals eventually learn the state and disagreement is temporary in Dixit and Weibull [11], whereas this is not the case in our model.

In our model, private signals are used to interpret the implication of the public signals and different private signals induce distinct interpretations. A similar idea is present in a number of papers, including Acemoglu, Chernozhukov, and Yildiz [1], Kandel and Pearson [26], Kondor [31], Kim and Verrecchia [28]. These papers consider Bayesian models in which there is uncertainty about the state of nature and individuals have *different priors* or *different private information* that determines *interpretation* of public signals. Consider, for instance, a recent contribution by Acemoglu, Chernozhukov, and Yildiz [1]. The key distinction between our models is as follows. In their model, disagreement between different individuals occurs even if these individuals expect to learn the state with certainty. By contrast, this result is impossible in our model due to the common prior assumption; if there is disagreement, then individuals necessarily fail to learn the state.

Our paper complements this body of work, first, by offering a very simple and transparent model of diverging opinions and, second, by experimentally demonstrating the possibility of disagreement and its *persistence* in an environment with *common priors*.

Finally, the general theme of this paper is related to the questions addressed in the work of Cripps, Ely, Mailath, and Samuelson [10] (CEMS) and Sethi and Yildiz [44]. CEMS provide conditions under which individuals who privately learn the value of a parameter will also learn it commonly. In a model with heterogenous priors

[5] Rabin and Schrag [40] provide a thoughtful summary of this research. See also Gerber and Green [20] for a review from the political science perspective and Nickerson [36] from the psychological perspective.

and private information, Sethi and Yildiz [44] provide conditions under which private information is aggregated through repeated communication.

The remainder of the paper is organized as follows. Next we briefly review the background literature. The model is presented in section 3. Section 4 describes our experimental design. The experimental results are provided in section 5. Section 6 concludes.

2 Background

Much of the evidence on diverging opinions comes from psychological studies.[6] In the seminal experiment by Lord, Ross and Lepper [33] subjects who were selected because of differing views on the death penalty were pulled further apart after reading the same essay about the death penalty. This type of result has been replicated in numerous studies.[7] Furthermore, in the classic work on the formation of public opinion, Zaller [49] suggests that political disagreement is more likely to occur among more aware and better informed individuals. He proposes that partisan preferences cause voters to "filter out" information that does not confirm their predispositions. Other psychological theories are cognitive dissonance (Festinger [16]) and motivated reasoning (Kunda [32]). These theories postulate that people have preferences and make choices over what they believe. In economics, these ideas have been formalized, for example, in Brunnermeier and Parker [6] and Yariv [48].

By contrast, Nisbett and Ross [38] downplay preference-driven interpretations and suggest that different cognitive biases could be an equally plausible explanation. Supporting this view, Evans [15], for instance, argues that confirmatory bias is caused by a selective processing of information and inability to think about alternative explanations. Eil and Rao [12] suggest that asymmetric updating may be motivated by image concerns. They find individuals incorporate "good news" into their priors more accurately than "bad news," regardless of whether the new information confirms or

[6]We refer the reader to the surveys of the literature by Barberis and Thaler [3], Gerber and Green [20], Hirshleifer [23], Narasimhan et al [35], and Rabin [39].

[7]Similar results are obtained by Houston and Fazzio [25] and Schuetto and Fazzio [42] in the context of capital punishment, Katz and Feldman [27] and Sigelman and Sigelman [45] in the context of presidential debates, Kinder and Mebane [29] in the context of evaluation of the state of economy, and Sears [43] in the context of the credibility of the source of factual information. Nickerson [36] provides a survey of related evidence; additional references can also be found in Greber and Green [20]. Finally, a recent study by Westen et al [46] finds further support for the effect of prior political attitudes on the interpretation of available evidence in an fMRI study.

disconfirms those priors.

Other explanations include ‘conditional reference frame,’ in which subjects choose to assume that the prior hypothesis is true (Koehler [30]); pragmatism and error avoidance (Friedrich [18]), in which people’s inferential strategies have goals other than finding out the truth, such as identifying potential rewards and avoiding costly mistakes; and educational effects (Nickerson [36]), according to which people learn to seek confirming evidence and ignore disconfirming evidence because of incentives in the educational process.

Several authors argue that confirmatory bias might have some adaptive value from the evolutionary perspective (Byrne and Kurland [7], Nickerson [36]) or may be an outcome of adaptive learning (Hopkin [24]).[8]

Hellman and Cover [22] and Wilson [47] consider single-person decision problems in which the decision maker has limited capacity to remember signals and has to find an optimal way of storing information. The optimal decision rule ignores some of the evidence with positive probability, implying that the posterior beliefs of two decision makers may diverge if they have different priors.[9] In a conceptually different approach, Epstein [13] and Epstein, Noor, and Sandroni [14] provide axiomatization of non-Bayesian updating that captures information processing biases.

3 Model

The model we offer is intentionally made as simple as possible. It can be generalized to signal distributions other than the one considered below but we think that the current model is sufficient to clarify the features that generate diverging opinions. In our model, private and public signals are complements: the value of either public or private signals alone is zero.[10] To illustrate our model, imagine that two players play one shot of matching pennies. There are two outside observes who receive noisy information about the players’ moves and are asked to bet on the winner of the game. Imagine that the observers receive different private information about the move of player 1. This information is not helpful in determining the winner of the game and hence does not affect their opinions. We now let them observe a public signal

[8]A similar idea appears in Compte and Postlewaite [9], who demonstrate that an information processing bias leading to overconfidence might be welfare improving.

[9]Similar results are obtained in Monte [34] who studies a strategic setting with imperfect information.

[10]Tilman, Hernando-Veciana, and Krähmer [5] study signals that are complements and substitutes and show that complementary signal structures are *not* non-generic.

about the move of player 2. Together with private information, this signal is valuable. Furthermore, the observers with different private information will now diverge in their opinions about who is the more likely winner of the game. Thus, arrival of public information can cause divergence of opinions.

3.1 Environment

The state of nature $\theta = (\alpha, \beta)$ is a realization of a random variable $\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta})$, where $\tilde{\alpha}, \tilde{\beta} \in \{0, 1\}$. All states are equally likely. There are two Bayesian agents, each can take an action $a \in \{Even, Odd\}$. The payoff of an agent is

$$u(Even, \theta) = \begin{cases} 1, & \text{if } \theta \text{ equals } (0,0) \text{ or } (1,1); \\ 0, & \text{otherwise;} \end{cases} \tag{1}$$

$$u(Odd, \theta) = \begin{cases} 1, & \text{if } \theta \text{ equals } (1,0) \text{ or } (0,1); \\ 0, & \text{otherwise,} \end{cases}$$

independently of the action taken by the other agent.

Agents do not know the state and observe two signals, $\tilde{a}, \tilde{b} \in \{0, 1\}$, that are distributed independently conditional on the state with $\Pr(\tilde{a} = \alpha | \alpha) = p_\alpha > 1/2$ and $\Pr(\tilde{b} = \beta | \beta) = p_\beta > 1/2$.

There are infinitely many periods, $t = 0, 1, ...$ In period zero, the agents privately observe independent realizations of signal $\tilde{a}$. Starting from the first period, the agents commonly observe a realization of signal $\tilde{a}$ or $\tilde{b}$ in each period. We will consider two settings: in one all public signals are $\tilde{b}$, whereas in the other public signals are of both types, $\tilde{b}$ in odd periods and $\tilde{a}$ in even periods.

3.2 Disagreement about the optimal action

Let *type 0* and *type 1* denote the agents who observe private signal $a = 0$ and $a = 1$ respectively. After observing $a = 1$, *type 1* believes that $\alpha = 1$ is more likely, while his beliefs about β are unaffected. However, (1) implies that this type is indifferent about which action to take. A similar argument applies to type 0. Hence, although different signals in the first period might lead to distinct beliefs about α, they cannot create a disagreement about the optimal action.

Nevertheless, private signals can be used to interpret future signals about the other dimension and can lead to disagreement. We say that type 0 and type 1 *disagree* about the optimal action if they strictly prefer different actions.[11] Imagine

[11] We say that the different types *weakly disagree* about the optimal action if one type is indifferent

that in the second period both types observe $b = 1$. Now, type 1 believes that the state $\theta = (1, 1)$ is more likely and the optimal course of action is $a =$*Even*, while type 0 disagrees.

The independence of the signals conditional on the state and their binomial distribution implies that the agent's posterior beliefs depend only on the difference between the number of realizations of different signals, but not their order. Let t_a and t_b be the number of respective public signals and k_a and k_b be the number of realizations of the corresponding signals equal to one. We define

$$\begin{aligned} \delta_a^0 &= 2k_a - t_a - 1, \\ \delta_a^1 &= 2k_a - t_a + 1, \\ \delta_b &= 2k_b - t_b. \end{aligned}$$

Remark 1. If $\delta_b = 0$, both types believe that both actions are equally likely to be optimal.

Remark 2. Different types disagree about the optimal action if and only if (i) $\delta_b \neq 0$ and (ii) $\delta_a^1 = 1$ and $\delta_a^0 = -1$.

3.3 Public signals $\tilde{b}$

If all public signals are $\tilde{b}$, we have $\delta_a^1 = 1$ and $\delta_a^0 = -1$ for any number of signals. Then, the probability of disagreement between types 0 and 1 is equal to $1 - \Pr(\delta_b = 0)$.

If the number of public signals is odd, then $\delta_b \neq 0$, in which case different types disagree with probability one. If, however, the number of public signals $\tilde{b}$ is even, $t_b = 2N$, $N > 0$, the probability of $\delta_b = 0$ is equal to

$$\Pr(\delta_b = 0) = \frac{(2N)!}{(N!)^2}(p_\beta(1 - p_\beta))^N.$$

This expression is decreasing in N and converges to zero as $N \to \infty$. Hence,

Proposition 1. *If all public signals are $\tilde{b}$, the probability of disagreement is*

1. *one if the number of public signals is odd;*
2. *positive and increasing in N if the number of public signals is even.*

about which action to take and the other type believes that one of the actions is more likely to be optimal.

As a measure of intensity of disagreement, we now consider the absolute value of the difference between the beliefs about the optimal action conditional on different private signals. Let

$$q^1(Even|\delta_b) = \Pr(\beta = 1|\delta_b)\Pr(\alpha = 1|a = 1) + (1 - \Pr(\beta = 1|\delta_b))(1 - \Pr(\alpha = 1|a = 1))$$

denote the probability that type 1 assigns to the event that the optimal action is $a =$*Even*, conditional on δ_b. Define $q^0(Even|\delta_b)$ analogously for type 0. The absolute value of the disagreement between the beliefs about the optimal action is

$$\Delta(\delta_b) = |q^1(Even|\delta_b) - q^0(Even|\delta_b)| = (2p_\alpha - 1)\left|2\frac{p_\beta^{\delta_b}}{p_\beta^{\delta_b} + (1-p_\beta)^{\delta_b}} - 1\right|.$$

The following proposition states that the expected value of disagreement, conditional on the realized state, is increasing in the number of signals.

Proposition 2. *For any $n > 0$, the expected absolute value of disagreement conditional on the realized state satisfies*

$$E\{\Delta(\delta_b)|t_b + 1, \alpha, \beta\} = E\{\Delta(\delta_b)|t_b, \alpha, \beta\}, \qquad \text{if } t_b = 2n + 1, \tag{2}$$

$$E\{\Delta(\delta_b)|t_b + 1, \alpha, \beta\} > E\{\Delta(\delta_b)|t_b, \alpha, \beta\}, \qquad \text{if } t_b = 2n. \tag{3}$$

Proof. See the appendix. □

3.4 Public signals $\tilde{a}$ and $\tilde{b}$

We now turn to the setting in which agents observe signal $\tilde{b}$ in odd periods and signal $\tilde{a}$ in even periods.

The probability of disagreement in this environment is non-monotone, due to discreteness of the signals, but it decreases after every four periods and converges to 0 as the number of signals becomes large. This is because the uncertainty about dimension α becomes small as more signals are realized. We can formally state this as follows.

Proposition 3. *Let public signals be of both types. Let $t_a > 0$ and $t_b > 0$ be the number of public signals of the respective types. Denote by $z(t_a, t_b)$ the probability of disagreement. Then,*

$$z(t_a, t_b) = \Pr(\delta_a^1 = 1|t_a)(1 - \Pr(\delta_b = 0|t_b))$$

and

$$z(t_a, t_b) > z(t_a + 2, t_b + 2).$$

Furthermore, $\lim_{t_a, t_b \to \infty} z(t_a, t_b) = 0$*.*

Proof. See the appendix. □

We now consider the absolute value of the difference between the beliefs about the optimal action conditional on different private signals. Let

$$q^1(Even|\delta_b, \delta_a^1) = \Pr(\beta = 1|\delta_b) \Pr(\alpha = 1|\delta_a^1) + (1 - \Pr(\beta = 1|\delta_b))(1 - \Pr(\alpha = 1|\delta_a^1))$$

be the probability that type 1 assigns to the event that the optimal action is $a =$*Even*, again conditional on δ_b. Define $q^0(Even|\delta_b, \delta_a^0)$ analogously for type 0.

The absolute value of the difference between beliefs about the optimal action of different types is equal to,

$$\Delta(\delta_b, \delta_a^1, \delta_a^0) = |q(Even|\delta_b, \delta_a^1) - q(Even|\delta_b, \delta_a^0)|.$$

The expected value of $\Delta(\delta_b, \delta_a^1, \delta_a^0)$ is non-monotone. Furthermore, it does not have to decrease every four periods. For example, if the signals about α are not very informative, i.e., p_α is close to $1/2$, while the signals about β are sufficiently informative, e.g., $p_\beta = 3/5$, then the expected value of $\Delta(\delta_b, \delta_a^1, \delta_a^0)$ will be *increasing* in the initial periods.

Nevertheless, after sufficiently many public signals the uncertainty about dimension α vanishes and, as a result, the expected value of $\Delta(\delta_b, \delta_a^1, \delta_a^0)$ converges to 0.

Proposition 4. *Let public signals be of both types. Then,*

$$\lim_{t_a, t_b \to \infty} E\left\{\Delta(\delta_b, \delta_a^1, \delta_a^0)|t_a, t_b, \alpha, \beta\right\} = 0.$$

Proof. The result follows from Theorem 1 in Freedman [17]. □

4 Experimental Design

The experiment was conducted using undergraduate subjects in 8 sessions involving 6 subjects each. Each session involved three sets. In each set, we randomly selected one of the four urns, A, B, C, and D, corresponding respectively to the four states of the world $(0,0)$, $(1,1)$, $(0,1)$, and $(1,0)$ in our model. Figure 1 illustrates how the urns were presented to the subjects.

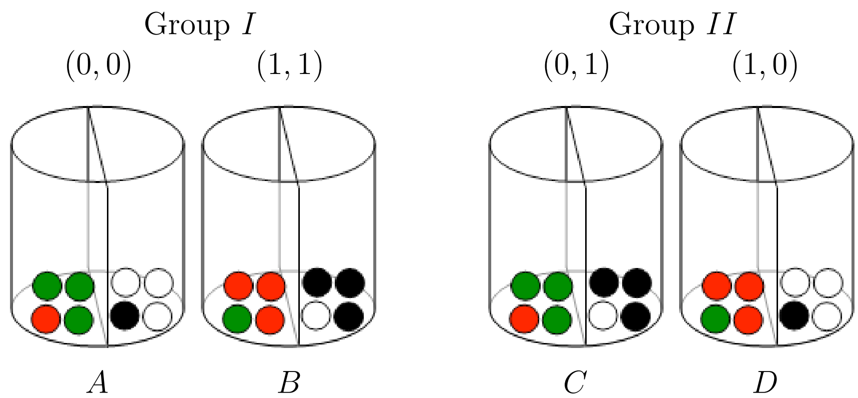

Figure 1: Content of the urns

There were two compartments in each urn. One compartment had red and green balls. The urns A and C (states $(0,0)$ and $(0,1)$) had three green and one red balls. The urns B and D (states $(1,1)$ and $(1,0)$) had one green and three red balls. A random draw from this compartment is equivalent to a signal $\tilde{a}$ with the support $\{Green, Red\}$, whose distribution is given by $p_\alpha = \Pr(Red|\alpha = 1) = \Pr(Green|\alpha = 0) = 3/4$.

The other compartment contained white and black balls. The urns A and D (states $(0,0)$ and $(1,0)$) had three white and one black balls. The urns B and C (states $(0,1)$ and $(1,1)$) had one white and three black balls. A random draw from this compartment is equivalent to a signal $\tilde{b}$ with the support $\{White, Black\}$, whose distribution is given by $p_\beta = \Pr(White|\beta = 1) = \Pr(Black|\beta = 0) = 3/4$.

In every set, each subject observed a total of 15 draws with replacement from the selected urn. First, each subject observed one private draw from the compartment containing red and green balls (signal $\tilde{a}$). After that, subjects commonly observed 14 public draws. There were two types of sets: joint and separate. In the joint set, public draws were equally likely to be made from either of the compartments (signals $\tilde{a}$ and $\tilde{b}$).[12] In the separate set, public draws were made only from the compartment containing white and black balls (signal $\tilde{b}$).

The urns were divided in two groups. Group 1 consisted of urns A and B (action *Even*) and Group 2 consisted of urns C and D (action *Odd*). To infer subjects' beliefs, subjects placed bets on which Group they thought the urn was in. There were 16 rounds of bets for each set. First, subjects could place bets after each one of the 15

[12] In the 8 sessions of the experiment there were a total of 12 joint sets. In the three of these sets, the separator between the compartments was removed and the public draws were made from the common pool containing balls of all four colors. In the remaining 9 sets, the draws alternated between the two compartments.

draws. In addition, after the bets on the 15th draw, the total cumulative numbers of bets on each group by all the participants were revealed, and the subjects could make bets one more time. The purpose of the 16th round of bets was to allow subjects to update their beliefs based on the information contained in the aggregate of all bets. If subjects are all perfect risk-neutral Bayesians and this is common knowledge, then there should be no disagreement on the round 16 bets.

In each round of bets, subjects could place from 0 to 9 bets on *each* of the groups of urns. That is, after each draw they could simultaneously make up to 18 bets, 9 or fewer bets on Group 1 and 9 or fewer bets on Group 2.

In the end of each set, one of the 16 rounds of bets was selected at random to determine the earnings of the subjects in that set. The subjects were given 10 points for every successful bet in this round, that is, the bet on the group which contained the urn used in this round. The bets made in this round also entailed costs. The first bet made on Group 1 in this round cost one point. The cost of each additional bet on Group 1 was one point more than the cost of the previous bet on Group 1. That is, the nth bet cost n points. Similarly, the 1st bet made on Group 2 in this round cost 1 point and the nth bet cost n points. If individuals are risk neutral Bayesian payoff maximizers, then bets should be revealing of their beliefs about the probabilities, that is, this is a proper scoring rule. For instance, a subject who thinks the likelihood is 0.35 that Group 1 is the true state, and 0.65 that it is Group 2, should place 3 bets on Group 1 and 6 bets on Group 2. Total bets across the two groups should always be 9 or 10.[13]

There are eight possible permutations of the sequence of three sets, each of which could be either joint or separate: (separate, separate, separate), (joint, separate, separate), etc. We conducted eight sessions with three sets each, one session for every possible permutation of sets. Hence, there were 12 separate sets and 12 joint sets. In one set, each of the six subjects made 16 rounds of bets on two groups. We obtained a total of 4,608 observations (16 rounds $\times$ 3 sets $\times$ 6 subjects $\times$ 8 sessions $\times$ 2 groups).

To guarantee that earnings were non-negative, subjects were endowed with 45 points in each set. Subjects kept track of the draws, their bets, and their earnings in each round using a computer interface. The information about the rules of the experiment and the content of the urns was known to the participants.

The experiment lasted about one hour. We paid US $1 for each 10 points earned in the experiment. The subjects were anonymously paid their cumulative earnings in

[13]To our knowledge, this paper's scoring rule is unique. Its advantage over a quadratic scoring rule is that it is easy to explain and understand.

cash at the end of the experiment. The earnings averaged \$20.52 (standard deviation \$3.98), ranging from \$9 to \$27.

5 Experimental Results

In the next two subsections, we present our evidence for Propositions 1 and 3 on the frequency of disagreement and for Propositions 2 and 4 on the expected value of disagreement. Our model predicts the overall pattern in the data well, with several important exceptions. Subsection 5.3 will examine individual data to address these exceptions. Then subjection 5.4 will examine our predictions about choices in the first round of each set, and subjections 5.5 will discuss our predictions for round 16, when we reveal the sum of all prior bets by all players. Finally, we discuss learning over the course of the experiment.

5.1 Probability of disagreement

Figure 2 depicts the frequency of the disagreement, for both theoretical and observed and both joint and separate sets, over the first 15 rounds.[14] The observed frequency has an increasing trend in separate sets and a decreasing trend in joint sets, in accordance with the theoretical predictions. At the same time, the observed frequency of disagreement in the first round is significantly larger than the theoretical value of zero. We will consider the bets made by the subjects in the first period in detail in Section 5.4.

How is Figure 2 generated? First, we excluded two sets, one joint set and one separate set, in which all subjects observed the same private information.[15] To determine the theoretical frequency of disagreement, for each subject we calculate Bayesian beliefs about whether the urn belongs to Group I given the subject's information. The theoretical frequency of disagreement is the frequency with which Bayesian beliefs of the subjects with different private information disagree in a given round in a given type of set.

[14] We exclude the 16th round because in this round the information observed by the subjects is the cumulative number of bets made by all subjects in the previous rounds. The Bayesian model alone cannot determine the combination of bets that maximize the expected payoff; this combination depends on the beliefs of the subject about how the other subjects make their bets.

[15] We do *not* exclude any subjects from the analysis, even the subjects who, regardless of the realized information, have placed the same combination of bets in all rounds in all sets.

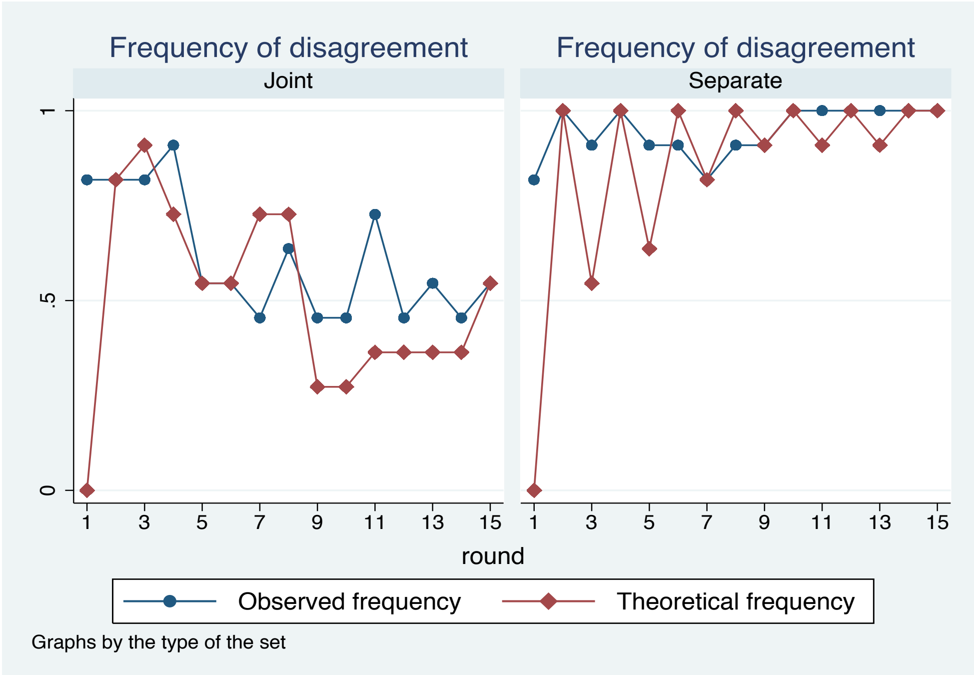

Figure 2: The expected and observed frequency of disagreement. In *joint* sets public signals are of both types, in *separate* sets public signals are *Black* and *White*.

There are multiple ways to define the observed frequency of disagreement. We have chosen the following definition. First, for each subject we calculate the difference between the bets made on Group I and on Group II. If the difference between the bets is larger than one, we say that the subject prefers the group on which she places more bets. Otherwise we say that the subject is indifferent.[16]

Next, we determine the group preferred by the majority of subjects who observe the same private information as follows: we exclude the subjects who are indifferent between the groups, and find the group preferred by the majority of the remaining subjects with strict preference. We say that the subjects are, on average, indifferent about which group to bet on if equal numbers of subjects strictly prefer different groups or if all subjects are indifferent. The observed frequency of disagreement is, then, the frequency with which the subjects who observe different private information

[16]One alternative we have considered is to define a subject to be indifferent between groups if and only if she places the same number of bets on both groups. The disadvantage of this definition is that it may incorrectly classify subjects as not indifferent. Imagine that a risk-neutral subject believes that both groups are equally likely. Then, she is willing to pay up to 5 points for a bet on each of the groups. Because, in our experiment, the 5^{th} bet costs 5 points, the subject is indifferent between placing 4 or 5 bets on each of the groups. Hence, the following combinations of bets are consistent with the belief that both groups are equally likely: (4,4), (5,5), (4,5), and (5,4).

prefer different groups: that is, the (majority of the) subjects with one private signal prefer one group and the subjects with the other private signal either prefer the other group or are, on average, indifferent about which group to bet on.[17]

5.2 Expected value of disagreement

Our first task is to describe how we will determine the theoretical and observed values of disagreement. First, we exclude the sets, one joint set and one separate set, in which all subjects observe the same private information. Second, to determine the theoretical value of disagreement, for each subject we calculate Bayesian beliefs about whether the urn belongs to Group I for each subject given the subject's information. The theoretical value of disagreement is the absolute value of the difference of Bayesian beliefs, multiplied by 10, about the event that the urn belongs to Group I for subjects with different private information in a given round, averaged over all sets of a given type. Finally, we determine the observed value of disagreement as follows. First, for each round in each set and for each group we calculate the absolute value of the difference between the average bets made by the subjects who observe different private information. The average of this value in a given round over both groups and all sets of a given type is the observed value of disagreement.

Figure 3 depicts the expected absolute value of disagreement, both theoretical and observed, for the first 15 rounds, averaged for each of the joint and separate sets. The observed value has an increasing trend in separate sets, in accordance with the theoretical predictions. In joint sets, both the observed and the theoretical value of disagreement do not increase. Similarly to the results for the frequency of disagreement, the observed value of disagreement in the first round is larger that the theoretical value of zero; this difference is less pronounced than in the case of the frequency of disagreement.

We can see in Figures 2 and 3 that our predictions are largely met, but with some important exceptions. First, both joint and separate sets badly miss the prediction on the first round for both the probability and expected value of disagreement. In addition, for the separate sets the change in disagreement typically is in the predicted

[17] In order to count, the direction of disagreement does not have to coincide with the theoretical prediction, although is does in almost every case. Deviations, therefore, work against hypothesis. Also, note that our definition of preference for a group applies regardless of whether the bets made by the subject in question are consistent with payoff-maximizing behavior. Hence, our conclusion that the observed frequency of disagreement is close to the theoretical frequency of disagreement *does not imply* that our theoretical model can explain the bets of the subjects.

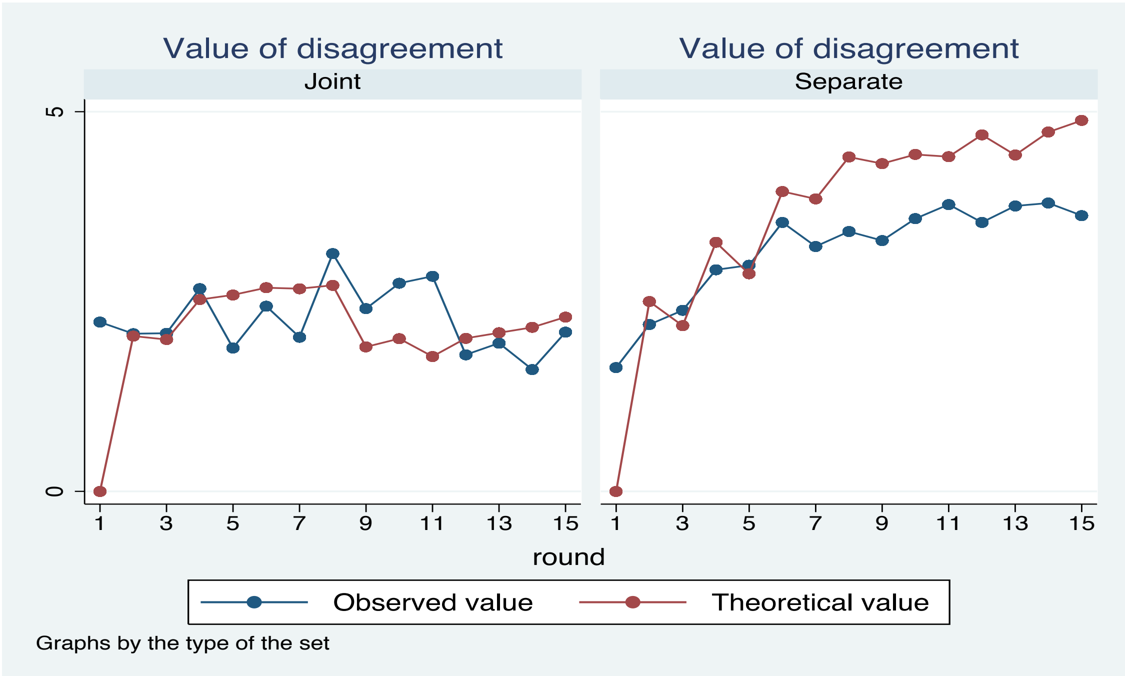

Figure 3: The expected and observed value of disagreement. In *joint* sets public signals are of both types, in *separate* sets public signals are *Black* and *White*.

direction, but the average tends to adjust more incrementally than predicted. Moreover, for the expected value of disagreement, Figure 3, the prediction does worse on average later in the session.

5.3 Median Split

In this section we want to focus on those subjects who are making the most costly errors. In particular, we would like to examine if deviations from Bayesianism are systematic, such as, for example, favoring a confirmatory bias, and have other unforeseen regularities. The alternative is that the errors are simply noise that is eliminated over time through learning.

Our first task is to identify those with the largest errors. We do this by dividing our subjects into two evenly sized groups based on how their earnings in the experiment compared with the theoretical prediction. For each subject we calculate the maximal expected payoff that the subject can achieve in each round given the information observed in the current and the previous rounds.[18] We then calculate the expected payoff of the subject for the bets that were made. The difference between

[18]We exclude the 16th round.

these payoffs divided by the maximal expected payoff is called the *lost share*. We use the lost share, averaged over the first 15 rounds and all sets for each subject, as our measure of this subject's consistency with the behavior of a Bayesian decision maker who maximizes expected payoff.

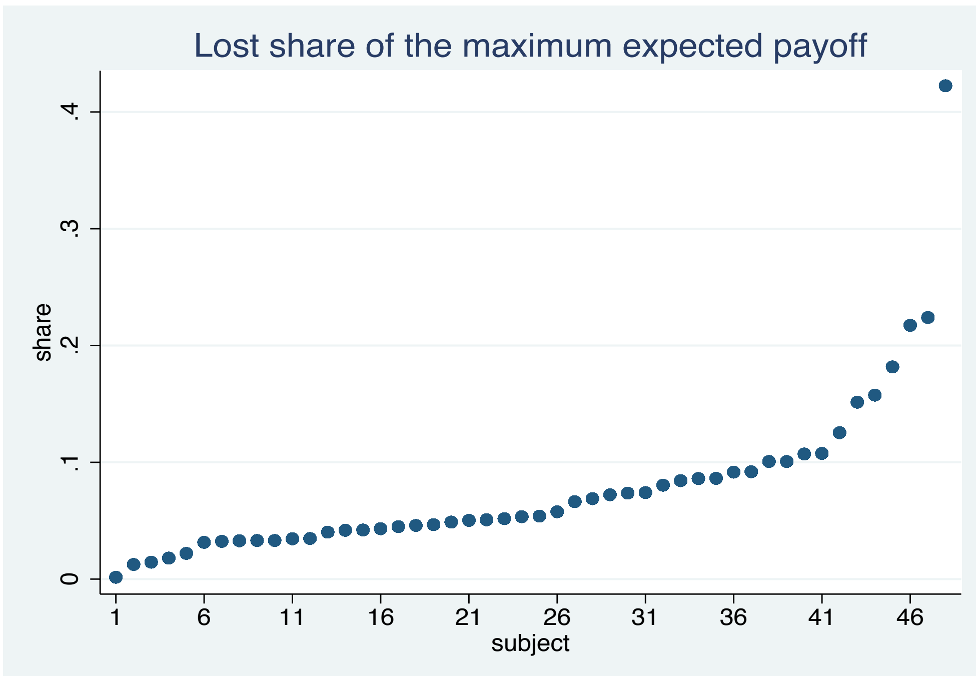

Figure 4: The share of the maximal expected payoff lost by the subjects in the experiment. There were 48 subjects total.

Figure 4 shows the distribution of the lost shares across subjects. Its average is equal to 0.078 (standard deviation 0.07), ranging from 0.0016 to 0.42. The subject with the smallest share made bets perfectly consistent with the Bayesian model in 35 out of 45 rounds. By contrast, the subject with the largest lost share put 9 bets on one group and 0 bets on the other group in all rounds in all sets, which was never optimal unless the subject was extremely risk loving. In fact, an extremely risk averse strategy of placing 5 (or 4) bets on each group every round would have been more profitable than this largest lost share strategy.[19]

Ranking our subjects by lost share, we split them at the median and consider separately the 24 subjects with a relatively small lost share less than 0.054, and the 24 subjects with a relatively large lost share above 0.054.[20]

[19]Note, we had one subject who actually chose this risk averse strategy. His lost share was 0.08,

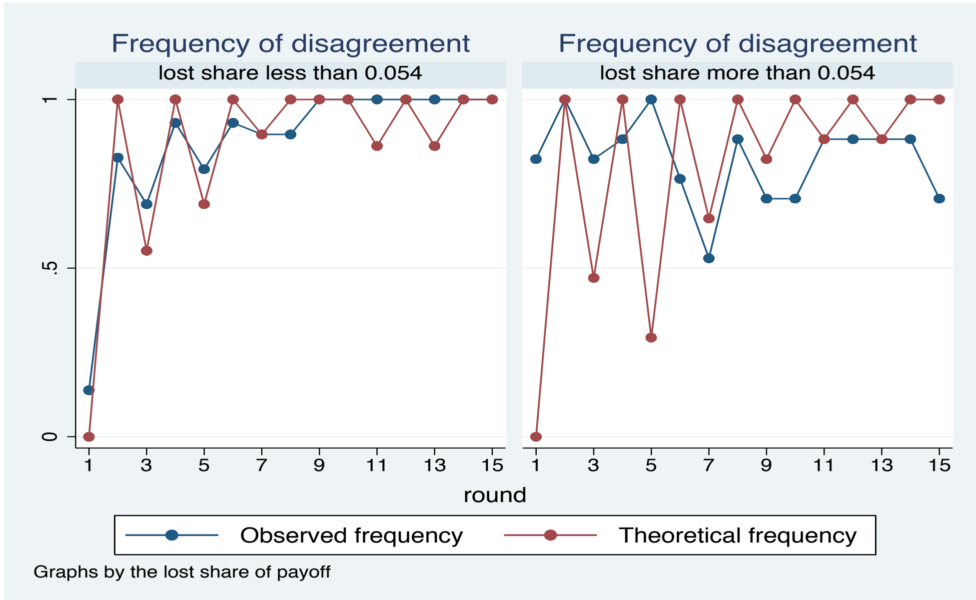

Figure 5: The expected and observed frequency of disagreement in *separate* sets, in which public signals are *Black* and *White*.

Figure 5 shows the frequency of disagreement in the first 15 rounds in separate sets for the small and large lost share.[21] The frequency of disagreement has an increasing trend for the small lost share group, and tracks the predictions closely. Importantly, this is especially true in round 1, where the frequency of disagreement is only 0.138. By contrast, in the large lost share group there is no clear trend in disagreement, and disagreement in round 1 is extreme.

Similar patterns emerge for the probability of disagreement in the joint sets, as seen in Figure 6. As last time, the outcome in the small lost share group is much closer to the prediction.[22] Most satisfying is the first round where the small lost share group now meets the prediction of zero disagreements. Disagreement then has a decreasing trend afterwards. The large lost share group has a large frequency

among the lowest.

[20]Unfortunately, because of the smaller number of subjects, there are now 6 separate sets and 3 joint sets in which all subjects in the group with the larger shares observe the same private information and 3 separate sets and 6 joint sets in which all subjects in the group with the smaller shares observe the same private information. We must therefore exclude these sets from the analysis.

[21]The average frequency of disagreement over the first 15 rounds is 0.824 for small and 0.874 for large lost share group, which is not significantly different.

[22]Again, the average frequency of disagreement is approximately the same for both groups, 0.454 for small and 0.517 for the large lost share and not significantly different.

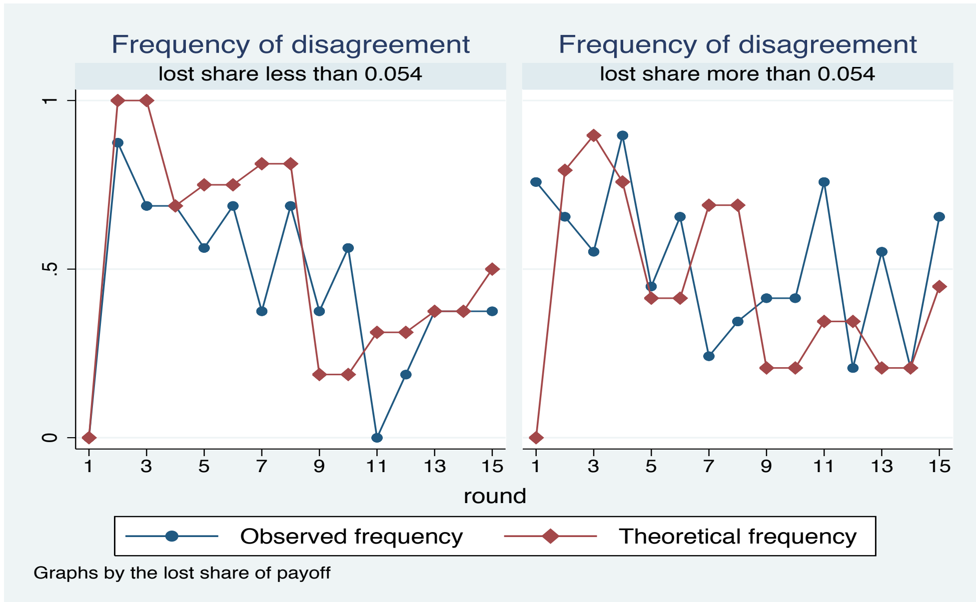

Figure 6: The expected and observed frequency of disagreement *joint* sets, in which public signals are of both types.

of disagreement in the first period and, perhaps, a less pronounced decreasing trend afterwards.

Turn next to the expected value of disagreement. Figure 7 shows this for the separate sets.[23] The value of disagreement has an increasing trend for both groups; this trend is more pronounced for the group with the smaller share.

Figure 8 shows the same information for the joint sets.[24] The value of disagreement is more consistent with the theoretical prediction for small lost share group. Interestingly, only in the large lost share group for joint sets do we see a propensity for subjects to disagree more than the model predicts. The excess disagreement seems to be present from the start of sessions, fluctuates greatly across rounds, and does not show any consistent pattern over time.

[23]The average value of disagreement is 3.29 bets for the small and 3.98 bets for large lost share group.

[24]The average value of disagreement is 1.9 bets for the small and 3.0 bets for the large lost share group.

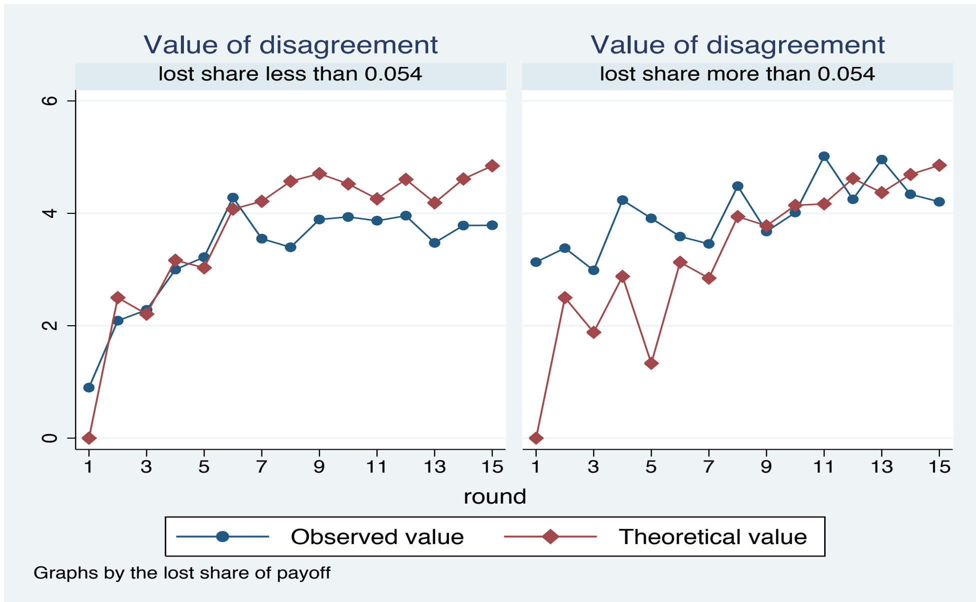

Figure 7: The expected and observed value of disagreement in *separate* sets, in which public signals are *Black* and *White*.

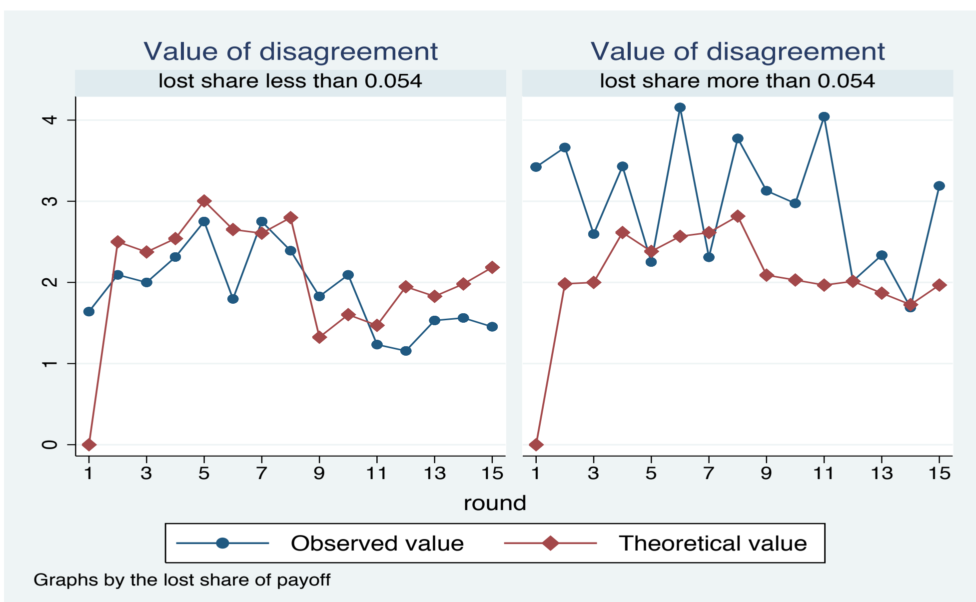

Figure 8: The expected and observed value of disagreement in *joint* sets, in which public signals are of both types.

5.4 First round

In the first round a decision maker has information on only one dimension of the information, therefore a Bayesian decision maker should believe that both groups are equally likely. As a result, subjects with different private information should not disagree. Since this contradicts what we see for some subjects, this section examines first round bets in more detail.[25]

Table 1: Preference over groups in the first round

Preference	Subjects with the lost share less than 0.054	Subjects with the lost share more than 0.054
Indifferent	70	48
Group I	2	13
Group II	-	11
Total	72	72

Table 1 describes the preferences of the subjects in the first round.[26] Consistent with the model, subjects with the smaller lost share of payoffs are indifferent between the groups in all except two cases. By contrast, the large lost share subjects have a strict preference over groups in one third of cases. Over all 144 observations, however, 81.9% of our subjects met our prediction of indifference in round 1.

Table 2 lists the combinations of bets made in the first period by the subjects with the small lost share. In total, 47 out of 72 pairs of bets (two thirds) are consistent with maximizing the expected payoff: (4,5), (5,4), (5,5), and (4,4). Furthermore, while most subjects in this group make optimal bets, those that do err tend to make too few or too many bets, but but do not favor one group.

Table 3 provides the combinations of first round bets for those with the large lost share. Only 16 out of 72 observations (22 percent) make bets consistent with

[25]Recall that to calculate the frequency of disagreement, we say that a subject is indifferent between the groups if her bets on the groups differ by at most one. If the difference between the bets on different groups is more than or equal to two, we say that the subject prefers the group on which she places majority of her bets.

[26]There is a total of 144 observations (48 subjects $\times$ 3 sets).

Table 2: Combinations of bets in the first round for the subjects with the lost share less than 0.054. The combinations of bets that are not classified as indifference are in bold. The combinations of bets consistent with payoff maximizing Bayesian behavior are in italics.

Group I \ Group II	0	1	2	3	4	5	6	7	8	9
0	-	-	-	-	-	-	-	-	-	-
1	-	2	-	-	-	-	-	-	-	-
2	-	-	2	-	-	-	-	-	-	-
3	-	-	-	3	-	-	-	-	-	-
4	-	-	-	-	*23*	*1*	-	-	-	-
5	-	-	-	**1**	*3*	*20*	1	-	-	-
6	-	-	-	-	-	-	7	-	-	-
7	-	-	-	-	-	-	1	4	-	-
8	-	-	-	-	-	-	**1**	-	2	-
9	-	-	-	-	-	-	-	-	-	1

Table 3: Combinations of bets in the first round for the subjects with the lost share more than 0.054. The combinations of bets that are not classified as indifference are in bold. The combinations of bets consistent with payoff maximizing Bayesian behavior are in italics.

Group I \ Group II	0	1	2	3	4	5	6	7	8	9
0	16	-	-	**2**	-	-	-	-	-	**7**
1	-	4	1	-	-	-	-	-	-	-
2	**1**	-	3	-	-	-	-	-	-	**1**
3	**2**	-	-	-	-	-	-	-	-	-
4	**1**	-	-	-	*10*	-	**1**	-	-	-
5	-	-	-	-	*2*	*4*	-	-	-	-
6	**1**	-	-	-	**1**	-	2	-	-	-
7	-	-	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-	-	-
9	**4**	**1**	-	-	**1**	-	-	**1**	1	5

maximizing expected payoff. While 23 pairs make bets that favor one group, 21 of these pairs put either 0 or 9 bets on one of the groups, which can only be rationalized with extreme risk loving preferences. Another 16 subjects erred by placing zero bets, and 11 by placing more than 10 total bets. The subjects in this group, therefore, make many more mistakes, both in the distribution and in the total number of bets.

This subsection reinforces the findings of the last subsection. Those in the small lost share group are largely consistent with our predictions, while the large lost share group make sundry errors in first-round betting with no evident pattern or regularity.[27]

5.5 Round Sixteen

The data from round 16 will allow us to test a fundamental prediction of the model: If there is common knowledge of rationality, then when people can aggregate others' information they should no longer disagree. On the other hand, if aggregating others' information is not enough to bring agreement, it suggests deeper issues with being able to convey information between individuals who disagree.

Recall that in round 16, the last round in a set, no balls were drawn. Instead, the subjects were informed about the total number of bets on each of the groups made in the previous rounds by all subjects in the current set. Then, the subjects placed their bets once again. Of course, if all subjects were risk-neutral Bayesian decision makers, then it should be impossible to have common knowledge of disagreement.[28]

Using the same measure of disagreement that we used in prior rounds, there was disagreement in 9 out of 24 sets. More importantly, we can also ask how frequently in the last round subjects prefer the group on which there was a majority of bets in all previous rounds. We have 144 individual observations, of which there were 39 disagreements with the majority. The subjects whose behavior is more consistent with payoff maximization, the low lost share group, follow the majority of bets a bit more often, 55 out of 72 times (76 percent), whereas the other subjects follow the majority 50 out of 72 times (69 percent).

There are multiple reasons why subjects may not place most of their bets on the group that received the majority of bets in the previous periods. One possibility is that the subjects expect the information about the majority of bets to be a very noisy signal and therefore might place a higher weight on their own opinion. Nevertheless,

[27]Note also that the pattern of the errors is inconsistent with simply a differential degree of risk aversion between low and high lost share subjects.

[28]More precisely, this follows from Theorem 3 in Nielsen et al [37].

in our experiment the majority of bets indicated the winning group in 21 out of 24 sets.[29] It might be useful to point out that a disagreement with the majority of bets was more frequent when it indicated the wrong urn: There was disagreement in 8 of 18 observations (44 percent) when the majority of bets was on the losing group and in 31 of 126 observations (24 percent) when the majority of bets was on the winning group. This indicates that some subjects were able to correctly conjecture when the other participants were not betting rationally.

A more precise question is whether subjects were able to switch their opinions from round 15 when they found their bets at odds with the group overall in round 16.[30] In our experiment, subjects should have changed their preferences in round 16 in 39 of 144 cases. They did change in 26 of these 39 cases (67%). Of the 13 cases where subjects continued to disagree in round 16, 12 of these cases correspond to the subjects in the large lost share half of the sample.[31]

These observations suggest that the subjects whose behavior is consistent with payoff maximization are less likely to disagree with majority in their 15th round and more likely to reconsider their bets if, in fact, they disagree. This is a rational strategy given that the information about the majority of bets is a rather precise signal about the winning group. On the other hand, the subjects who are not payoff-maximizers tend to disagree with the majority and continue to disagree after they learn about it more frequently in our experiment.

5.6 Are the subjects learning?

If, as appears from the prior subsection, errors are fairly asystematic, we can ask whether subjects behave more like Bayesians with experience. Figure 9 shows that the subjects who perform well overall, the low lost share subjects, do well throughout the experiment. The high lost share subjects, by contrast, make measurable improve-

[29] In the two of the remaining sets, the majority of bets was on the losing group because of the behavior of the subject with the largest lost share of payoff, 0.42. In each round, this subject put 9 bets on the losing group and 0 bets on the winning group. The other subjects, by contrast, were relatively conservative and placed similar numbers of bets on both groups. As a result, the larger total number of bets was placed on the losing group.

[30] Unlike in the previous analysis, here we say that there is indifference between both groups if and only if the number of bets is the same on both groups.

[31] There is one observation in our data in which the subject whose bets in the 15th round agreed with the majority of the bets changed her preference over bets in the 16th round. Hence, she switched the group on which she placed the majority of bets even though she should not have. The lost share of payoff for this subject is 0.18; this is the fourth largest share among the subjects. We have not been able to identify a specific pattern in the behavior of this subject.

ments. This suggests that the performance of our subjects would likely be even closer to our model with more experience.

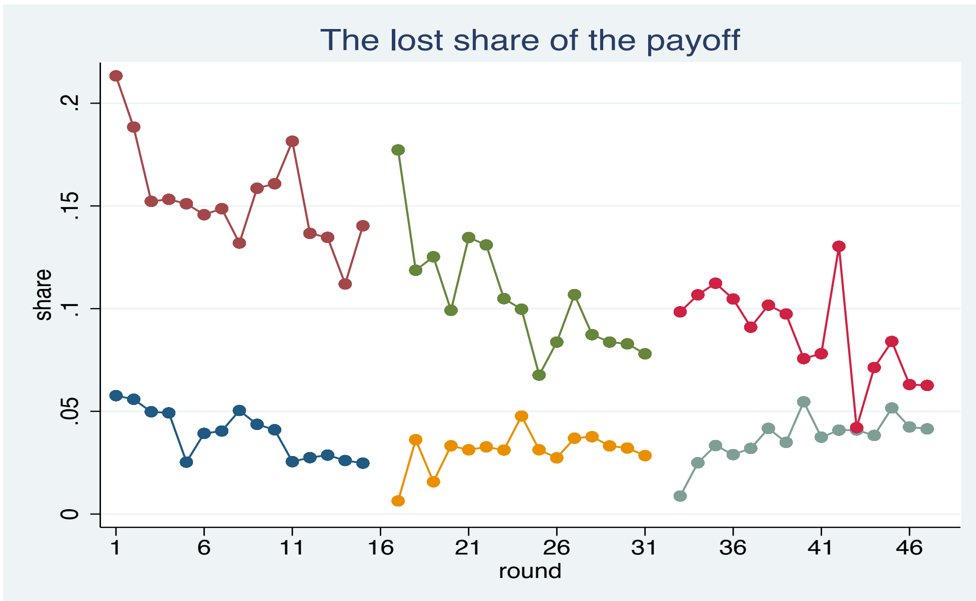

Figure 9: The share of the maximal expected payoff lost by the subjects in each round of the experiment. The subjects are split in two groups. The lower curve represent the subjects with the lost share of the payoff less than 0.054, whereas the upper curve represents the subjects with the lost share of the payoff more than 0.054.

6 Conclusions

It is intuitive that for any disagreement, more evidence on the question should tend to bring people together. Yet we see disagreements frequently in life, and sometimes these disagreements become even more polarized as additional information becomes available. How could such behavior be consistent with rational choice?

We present a model in which disagreement may arise and continuously increase as more commonly shared information becomes available, even under the assumption of Bayesian decision makers. The intuition for our model is that individuals may have private information, perhaps gained through personal experience or introspection, that may rationally color how they interpret later information.

Consider, for example, the question of whether the government should be given freedom to wiretap expected terrorists without court supervision. Compare an opponent, who believes it is more likely that government will abuse its power to conduct wiretaps, to a proponent, who believes it is more likely court supervision will slow the government's urgent need to capture terrorists. When news breaks that the government has conducted warrantless wiretaps, one side sees this stealth activity as evidence of abuse while the other sees this as evidence of a dire need for secrecy to stop terrorists. Both sides, according to our model, may be able to claim that the evidence rationally supports their view.

Formalizing this intuition, we present a model in which disagreement arrises because information about the true state has more than one dimension. Information on one dimension informs people about the probable "model" of the world, which later affects how they process commonly available evidence on another dimension. That is, if public information arrives on only some dimensions, it can lead to opposite inferences about the true state. Disagreement increases because information fails to identify the state of nature precisely. The model is consistent with results like those in the famous motivating works of Lord, Ross, and Lepper [33] and Zaller [49]. Moreover, the model provides testable implications about how the degree of disagreement depends on the structure of information.

We test this model with an experiment. The experiment mirrors the theoretical assumptions. We find that, as predicted, disagreement is both more likely and more polarized when information arrives on only one of the dimensions.

We also find great heterogeneity in the behavior of subjects. To illustrate this, we ranked subjects by their deviation from the maximum expected payoff they could have made, given the information they saw, and split the sample along the median. We found that both of the groups have a larger degree of disagreement when they received information on only one dimension. However, the group whose behavior is more consistent with payoff maximization tended to disagree less among themselves in round one and perhaps most importantly, virtually all agree after they were provided with information in the final round about the prior bets of the other subjects in their session. Their choices matched our predictions quite closely.

The group that was less consistent with profit maximization, while still largely supportive of our predictions. Most importantly, when we informed subjects of the aggregate bets in their session, subjects in this lower-performing group were more likely to maintain disagreements than the other half of the subjects. This indicates that even in circumstances in which common knowledge of disagreement should be impossible, beliefs fail to converge.

Since many disagreements are public, this raises the question, how can people continue to disagree even when they can communicate? While our model and data do not answer this question, they do offer new avenues for addressing it that don't require assumptions of non-Bayesian behavior or non-common priors. For people who disagree in our model to find common ground, they would need to share the private information that formed their core "models" for filtering the public information. If this information is gained, for instance, through personal experience, trusted advice, religious tradition, family heritage, or personal identity, then this information will be difficult to convey. If one is skeptical about their opponent's ulterior motives, such as in the wire-tapping example, then even two sincere rivals would not be able to agree. Finally, as in the psychological studies of implicit learning, the opponents may simply be unaware of why and how they came to the "values," "philosophies," and "world views" that form their models, and are thus unable to articulate their intuitive knowledge to others—their only choice is to agree to disagree.

A Proofs omitted in the text

Proof of Proposition 2. Assume that the realized state is $(\alpha, \beta) = (1, 1)$. The proof for the other cases is analogous. For a given (α, β), define

$$\gamma(t_b) = \frac{E\left\{\Delta(\delta_b)|t_b, \alpha, \beta\right\}}{2p-1}.$$

We have

$$\begin{aligned}\gamma(2n+1) &= \sum_{i=0}^{2n+1} \frac{(2n+1)!}{(2n+1-i)!i!} p^{2n+1-i}(1-p)^i \left| 2\frac{p^{2n+1-i}(1-p)^i}{p^{2n+1-i}(1-p)^i + p^i(1-p)^{2n+1-i}} - 1 \right| \\ &= \sum_{i=n+1}^{2n+1} \frac{(2n+1)!}{(2n+1-i)!i!} v(i, 2n+1), \ (n \geq 0)\end{aligned}$$

where

$$v(i, l) = p^i(1-p)^{l-i} - p^{l-i}(1-p)^i, \ (i \geq 0, l \geq 1).$$

Similarly,

$$\gamma(2n) = \sum_{i=n}^{2n} \frac{(2n)!}{(2n-i)!i!} v(i, 2n), \ (n > 0).$$

Note that for all $n \geq 1$,

$$v(n, 2n) = 0, \tag{A1}$$

$$v(i, 2n-1) = v(i, 2n) + v(i+1, 2n). \tag{A2}$$

Before we proceed with the proof, recall the following useful fact

$$\frac{N!}{(N-i)!i!} = \frac{(N-1)!}{(N-i)!(i-1)!} + \frac{(N-1)!}{(N-1-i)!i!}, \ (0 < i < N, N \geq 2). \tag{A3}$$

We now prove (2). First, it is straightforward to check that $\gamma(2) = \gamma(1)$. Now, let $n \geq 1$. Using (A3), we can write

$$\begin{aligned}
\gamma(2n) &= \sum_{i=n}^{2n-1} \left[\frac{(2n-1)!}{(2n-1-i)!i!} + \frac{(2n-1)!}{(2n-i)!(i-1)!} \right] v(i,2n) + v(2n,2n) \\
&= \sum_{i=n}^{2n-1} \frac{(2n-1)!}{(2n-1-i)!i!} v(i,2n) + \sum_{i=n-1}^{2n-2} \frac{(2n-1)!}{(2n-1-i)!i!} v(i+1,2n) + v(2n,2n) \\
&= \sum_{i=n}^{2n-1} \frac{(2n-1)!}{(2n-1-i)!i!} (v(i,2n) + v(i+1,2n)) + \frac{(2n-1)!}{n!(n-1)!} v(n,2n).
\end{aligned}$$

Then, it follows from (A1) and (A2) that

$$\gamma(2n) - \gamma(2n-1) = 0.$$

Next, we prove (3). Using (A3),

$$\begin{aligned}
\gamma(2n+1) &= \sum_{i=n+1}^{2n} \left[\frac{(2n)!}{(2n-i)!i!} + \frac{(2n)!}{(2n+1-i)!(i-1)!} \right] v(i,2n+1) + v(2n+1,2n+1) \\
&= \sum_{i=n+1}^{2n} \frac{(2n)!}{(2n-i)!i!} v(i,2n+1) + \sum_{i=n}^{2n-1} \frac{(2n)!}{(2n-i)!i!} v(i+1,2n+1) \\
&+ v(2n+1,2n+1) \\
&= \sum_{i=n+1}^{2n} \frac{(2n)!}{(2n-i)!i!} (v(i,2n+1) + v(i+1,2n+1)) + \frac{(2n)!}{n!n!} v(n+1,2n+1).
\end{aligned}$$

Then, from (A1) and (A2), we get

$$\gamma(2n+1) - \gamma(2n) = \frac{(2n)!}{n!n!} v(n+1,2n+1) > 0.$$

□

Proof of Proposition 3. We provide a proof of the first part of the proposition for the case of $t_a = t_b = 2n$, $n \geq 1$. (The argument for the remaining cases is analogous.) Because t_a and t_b are even, the probability of disagreement is given by

$$z(t_a, t_b) = z(n) = \Pr(\delta_a^1 = 1 | t_a)(1 - \Pr(\delta_b = 0 | t_b)). \tag{A4}$$

Set $q_k = p_k(1-p_k)$ where $k = \alpha, \beta$, and note that $q_k \in [0, 1/4)$. Then, (A4) can be rewritten as

$$z(n) = \frac{(2n)!}{(n!)^2} q_\alpha^n \left(1 - \frac{(2n)!}{(n!)^2} q_\beta^n\right).$$

It follows that

$$\begin{aligned} z(n+1) - z(n) &= \frac{(2n+2)!}{(n+1!)^2} q_\alpha^{n+1} \left(1 - \frac{(2n+2)!}{(n+1!)^2} q_\beta^{n+1}\right) - \frac{(2n)!}{(n!)^2} q_\alpha^n \left(1 - \frac{(2n)!}{(n!)^2} q_\beta^n\right) \\ &= q_\alpha^n \frac{(2n)!}{(n!)^2} g(n), \end{aligned}$$

where

$$g(n) = \left(\frac{(2n+2)(2n+1)}{(n+1)^2} q_\alpha \left(1 - \frac{(2n+2)!}{(n+1!)^2} q_\beta^{n+1}\right) - \left(1 - \frac{(2n)!}{(n!)^2} q_\beta^n\right)\right).$$

Before we proceed with the proof, note that for any $q_\beta \in [0, 1/4)$ and any $n \geq 1$

$$\left(1 - \frac{(2n+1)^2}{(n+1)^2} q_\beta\right) q_\beta^n \leq \frac{n^n (n+1)^{n-1}}{(2n+1)^{2n}} \tag{A5}$$

Furthermore,

$$\frac{n^n (n+1)^{n-1}}{(2n+1)^{2n}} \leq \frac{1}{4^n (n+1)} \tag{A6}$$

Now, we have

$$\begin{aligned} g(n) &\overset{q_\alpha < 1/4}{<} \frac{2n+1}{2(n+1)} \left(1 - \frac{(2n+2)!}{(n+1!)^2} q_\beta^{n+1}\right) - \left(1 - \frac{(2n)!}{(n!)^2} q_\beta^n\right) \\ &= -\frac{1}{2(n+1)} + \frac{(2n)!}{(n!)^2} \left(1 - \frac{(2n+1)^2}{(n+1)^2} q_\beta\right) q_\beta^n \\ &\overset{(A5),(A6)}{\leq} \frac{1}{2(n+1)} \left(\frac{2(2n)!}{4^n (n!)^2} - 1\right) \leq 0. \end{aligned}$$

We now turn to the limit result. It follows from (A4) that if $t_a = 2n$, then the probability of disagreement is bounded from above by $\Pr(\delta_a^1 = 1 | t_a) = \frac{(2n)!}{(n!)^2} q_\alpha^n$. The bound converges to 0 as $n \to \infty$. Similarly, if $t_a = 2n+1$, the probability of disagreement is bounded from above by $\Pr(\delta_a^1 = 2 | t_a) + \Pr(\delta_a^1 = 0 | t_a) = \frac{(2n+1)!}{n!(n+1)!} q_\alpha^n$, which also converges to 0 as $n \to \infty$. $\square$

References

[1] Daron Acemoglu, Victor Chernozhukov, and Muhamet Yildiz, *Fragility of asymptotic agreement under Bayesian learning*, 2009.

[2] Robert J. Aumann, *Agreeing to disagree*, The Annals of Statistics **4** (1976), no. 6, 1236–1239.

[3] Nicholas Barberis and Richard Thaler, *A survey of behavioral finance*, Handbook of the Economics of Finance (G.M. Constantinides, M. Harris, and R. M. Stulz, eds.), Handbook of the Economics of Finance, vol. 1, Elsevier, 2003, pp. 1053–1128.

[4] David Blackwell and Lester Dubins, *Merging of opinions with increasing information*, The Annals of Mathematical Statistics **33** (1962), no. 3, 882–886.

[5] Tilman Boergers, Angel Hernando-Veciana, and Daniel Krähmer, *When are signals complements or substitutes?*, 2009.

[6] Markus K. Brunnermeier and Jonathan A. Parker, *Optimal expectations*, American Economic Review **95** (2005), no. 4, 1092–1118.

[7] Christopher C. Byrne and Jeffrey A. Kurland, *Self-deception in an evolutionary game*, Journal of Theoretical Biology **212** (2001), no. 4, 457–480.

[8] Axel Cleeremans, *Models of implicit learning*, Encyclopedia of Cognitive Sciences (L. Nadel, ed.), vol. 2, London: Macmillan Publishers, 2003, pp. 491–499.

[9] Olivier Compte and Andrew Postlewaite, *Confidence-enhanced performance*, American Economic Review **94** (2004), no. 5, 1536–1557.

[10] Martin W. Cripps, Jeffrey C. Ely, George J. Mailath, and Larry Samuelson, *Common learning*, Econometrica **76** (2008), no. 4, 909–933.

[11] Avinash K. Dixit and Jörgen W. Weibull, *Political polarization*, Proceedings of the National Academy of Science of the United States of America **104** (2007), no. 18, 7351–7356.

[12] David Eli and Justin Rao, *The good news - bad news effect: Asymmetric processing of objective information about yourself*, 2010.

[13] Larry G. Epstein, *An axiomatic model of Non-Bayesian updating*, Review of Economic Studies **73** (2006), no. 2, 413–436.

[14] Larry G. Epstein, Jawwad Noor, and Alvaro Sandroni, *Non-Bayesian updating: A theoretical framework*, 2008.

[15] Jonathan St. B. T. Evans, *Bias in human reasoning: Causes and consequences*, Psychology Press, 1990.

[16] Leon Festinger, *A theory of cognitive dissonance*, Stanford University Press, 1962.

[17] David A. Freedman, *On the asymptotic behavior of bayes' estimates in the discrete case*, The Annals of Mathematical Statistics **34** (1963), no. 4, 1386–1403.

[18] James Friedrich, *Primary error detection and minimization (pedmin) strategies in social cognition: A reinterpretation of the confirmation bias phenomena*, Psychological Review **100** (1993), no. 2, 298–319.

[19] John D. Geanakoplos and Heraklis M. Polemarchakis, *We can't disagree forever*, Journal of Economic Theory **28** (1982), no. 1, 192–200.

[20] Alan Gerber and Donald Green, *Misperceptions about perceptual bias*, Annual Review of Political Science **2** (1999), 189–210.

[21] Milton Harris and Artur Raviv, *Differences of opinion make a horse race*, Review of Financial Studies **6** (1993), no. 3, 473–506.

[22] Martin E. Hellman and Thomas M. Cover, *Learning with finite memory*, The Annals of Mathematical Statistics **41** (1970), no. 3, 765–782.

[23] David Hirshleifer, *Investor psychology and asset pricing*, Journal of Finance **56** (2001), no. 4, 1533–1597.

[24] Ed Hopkin, *Adaptive learning models of consumer behavior*, Journal of Economic Behavior and Organization (forthcoming).

[25] David A Houston and Russell H. Fazio, *Biased processing as a function of attitude accessibility: Making objective judgments subjectively*, Social Cognition **7** (1989), no. 1, 51–66.

[26] Eugene Kandel and Neil D. Pearson, *Differential interpretation of public signals and trade in speculative markets*, Journal of Political Economy **103** (1995), no. 4, 831–872.

[27] Elihu Katz and Jacob J. Feldman, *The debates in light of research*, The Great Debates (Sidney Kraus, ed.), Bloomington: Indiana University Press, 1962, pp. 173–223.

[28] Oliver Kim and Robert E. Verrecchia, *Pre-announcement and event-period private information*, Journal of Accounting and Economics **24** (1997), 395–419.

[29] Donald R. Kinder and Walter R. Mebane, Jr., *Politics and economics in everyday life*, The Political Process and Economic Change (Bruno S Frey and Kristen R Monroe, eds.), Algora Publishing, 1983, pp. 141–180.

[30] Derek J Koehler, *Explanation, imagination, and confidence in judgment*, Psychological Bulletin **110** (1991), no. 3, 499–519.

[31] Peter Kondor, *The more we know, the less we agree: Higher-order expectations, public announcements and rational inattention*, 2009.

[32] Ziva Kunda, *The case for motivated reasoning*, Psychological Bulletin **108** (1990), no. 3, 480–498.

[33] Charles G. Lord, Lee Ross, and Mark R. Lepper, *Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence*, Journal of Personality and Social Psychology **37** (1979), no. 11, 2098–2109.

[34] Daniel Monte, *Reputation and bounded memory: A theory of inertia and skepticism*, December 2007.

[35] Chakravarthi Narasimhan, Chuan He, Eric Anderson, Lyle Brenner, Preyas Desai, Dmitri Kuksov, Paul Messinger, Sridhar Moorthy, Joseph Nunes, Yuval Rottenstreich, Richard Staelin, and George Wu and, *Incorporating behavioral anomalies in strategic models*, Marketing Letters **16** (2005), no. 3, 361–373.

[36] Raymond S. Nickerson, *Confirmatory bias: A ubiquitous phenomenon in many guises*, Review of General Psychology **2** (1998), no. 2, 175–220.

[37] et al Nielsen, Lars Tyge, *Common knowledge of an aggregate of expectations*, Econometrica **58** (1990), no. 5, 1235–39.

[38] Richard E. Nisbett and Lee Ross, *Human inference*, Prentice Hall, 1980.

[39] Matthew Rabin, *Psychology and economics*, Journal of Economic Literature **36** (1998), no. 1, 11–46.

[40] Matthew Rabin and Joel L. Schrag, *First impressions matter: A model of confirmatory bias*, The Quarterly Journal of Economics **114** (1999), no. 1, 37–82.

[41] Luis Santos-Pinto and Joel Sobel, *A model of positive self-image in subjective assessments*, American Economic Review **95** (2005), no. 5, 1386–1402.

[42] Robert A. Schuette and Russell H. Fazio, *Attitude accessibility and motivation as determinants of biased processing: A test of the MODE model*, Personality and Social Psychology Bulletin **21** (1995), no. 7, 704–710.

[43] David O. Sears, *Political behavior*, The Handbook of Social Psychology (Gardner Lindzey and Elliot Aronson, eds.), Reading, MA: Addison-Wesley, 2 ed., 1968, pp. 315–458.

[44] Rajiv Sethi and Muhamet Yildiz, *Public disagreement*, 2009.

[45] Lee Sigelman and Carol K. Sigelman, *Judgments of the Carter-Reagan debate: The eyes of the beholders*, The Public Opinion Quarterly **48** (1984), no. 3, 624–628.

[46] Drew Westen, Pavel S. Blagov, Keith Harenski, Clint Kilts, and Stephan Hamann, *Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 u.s. presidential election*, Journal of Cognitive Neuroscience **18** (2006), no. 11, 1947–1958.

[47] Andrea Wilson, *Bounded memory and biases in information processing*, 2005.

[48] Leeat Yariv, *I'll see it when i believe it: A simple model of cognitive consistency*, 2005.

[49] John R. Zaller, *The nature and origins of mass opinion*, Cambridge University Press, 1992.