Information Aggregation and Group Decisions*

Joel Sobel
University of California, San Diego

January 17, 2006

Abstract

Individuals with identical preferences each receive a signal about the unknown state of the world and separately decide upon a utility-maximizing recommendation on the basis of that signal. The group's decision maximizes the common utility function based on perfect pooling of individual information. With no restrictions on the information structure, the individual recommendations place no constraints on the group's decision. In a monotone environment in which individuals receive conditionally independent signals, the paper presents conditions under which polarization does and does not arise. *Journal of Economic Literature* Classification Numbers: A12, D01; Keywords: statistical decision problem; group polarization; behavioral economics; psychology.

*I am happy to acknowledge useful comments from seminar audiences at Michigan, Stanford, and UCSD and from Doug Bernheim, Antonio Cabrales, Juan Carrillo, Vince Crawford, Todd Davies, Rick Harbaugh, Navin Kartik, David Laibson, Debraj Ray, Ricardo Serrano, Ross Starr, Joel Watson, and, especially, Mark Machina. Individually, these people each dislike the paper a bit. Collectively, they hate it.

1 Introduction

Group polarization refers to the tendency of groups to make decisions that are more extreme than some average of the individual positions of the members of the group. The phenomenon, first observed in experiments reported by Stoner (1968), has been widely replicated.[1] This paper is an attempt to understand group polarization as a natural outcome of information sharing.

The literature implicitly assumes that rationality imposes restrictions on the relationship between decisions based on individual information and the collective decision and that polarization is evidence that group decision making is prone to systematic errors. My first goal is to present a framework in which one can evaluate whether a particular group decision is not rational merely by comparing it to recommendations made independently by the members of the group.

I study the problem assuming that groups aggregate information. Individuals have common preferences but different information.[2] Individuals observe a signal and independently recommend an action. The group then pools the individual signals and makes a decision based on the pooled information. I assume that there is no conflict of interest between group members and that the group perfectly aggregates the information of its members. In this setting, I investigate whether there are conditions under which the group's optimal decision is constrained by the recommendations of the individuals. One would be able to use these results to conclude whether experimental evidence necessarily was inconsistent with rationality.

I model differences in information by assuming that there is an underlying state of the world and individuals receive private signals that convey information about the state. The information structure describes the relationship between states of the world and signals. By asking individuals to make recommendations separately and then as a group, one observes a profile of actions (consisting of a recommendation from each individual and a separate group decision). I characterize the set of action profiles that arise with

[1]Brown (1986) devotes 50 pages of a textbook to the topic of group polarization. Isenberg (1986), Myers and Lamm (1976), and Turner (1991) also review the psychological literature.

[2]In experiments, individuals receive standardized information. The variation in their reaction to the information suggests that they interpret the information differently. I assume heterogeneous information, which can be viewed as a reduced form of a more elaborate model.

positive probability for some information structure. The results in Section 2 demonstrate that with no restrictions on the information structure, any action profile that avoids dominated actions is possible. In other words, there is an information structure in which individuals receive signals that induce any given profile of actions with positive probability. This result means that without further assumptions it is impossible to conclude that a group decision is irrational by observing individual recommendations. In fact, I show a stronger result. A belief profile is a collection of distributions over states of the world, one for each member of the group and one for the group taken together. For any finite collection of belief profiles there is an information structure such that each belief profile in the collection can be induced by some signals (one for each individual). Consequently, when beliefs determine recommendations, any action profile is consistent with rationality.

The constructions in Section 2 rely strongly on the ability to fine tune individual and collective information using correlated signals. I next make strong restrictions on the information structure. First, I assume that individual signals are independent conditional on the state of the world. With this restriction, individual beliefs completely determine the group's beliefs. On the other hand, when many different beliefs can induce the same recommendation, it is still typically the case that individual recommendations place no constraints on the group's decision.

In Section 3, I assume that individuals receive independent and identically distributed signals about the state of the world and, further, that the decision problem is monotone. In a monotone decision problem, states, actions, and signals are all linearly ordered so that higher signals are associated with higher actions and the optimal action is an increasing function of the signal. This environment is common in economic applications. The monotone structure does place restrictions on the set of action profiles that can be observed. For example, if one observes two action profiles and, in the first, each individual recommendation is at least as great as the corresponding recommendation in the second profile, then the group decision must also be greater in the first profile. Nevertheless, I find that even with these strong assumptions on the information structure, group decisions are generally not constrained by individual recommendations. Given any profile of actions, there exists an information structure that gives rise to signals inducing these actions with positive probability.

The general analysis suggests strongly that polarization need not be irrational. My second goal is to identify useful conditions under which one should expect to find group decisions that are moderate (bounded by individual recommendations) and alternative conditions under which polarization arises. For this purpose, it is useful to contrast problems of information aggregation with problems of preference aggregation. When aggregating preferences, it is standard (and usually not controversial) to assume a variation on Arrow's (1963) Pareto Principle. If every member of the group ranks choice X higher than choice Y, then the group should do so as well. In problems of information aggregation, this property is quite strong, and likely to be inappropriate in realistic settings. Separate individuals may, on the basis of limited private information, prefer a moderate recommendation to an extreme one. When these individuals pool their information, they may become more confident. Their confidence may shift their decision to an extreme.[3] Section 4 gives conditions under which the group's decision is bounded by individual recommendations. A necessary condition for this is a property that I call invariance. Assume that every member of the group receives a signal that, conditional on the true state, is independent and identically distributed. In an invariant decision problem, when the realization of the signal is the same for all members of the group, the group's decision is the same as the decision of any single individual. That is, the problem is invariant if when all agents individually prefer recommendation X, then so does the group.

Section 5 investigates conditions under which polarization should be expected. I distinguish between two classes of monotone decision problems. In single-crossing models, a perfectly informed decision maker would select between one of two decisions. Polarization is likely in this kind of model. On the basis of more precise information, groups make extreme choices. If all individuals favor recommendations that are close to one of the two actions, then the aggregate decision will be even closer to that action. An example of a problem of this kind is a simple allocation decision in which an investor must choose how to divide her wealth between a safe and risky investment. With limited information, it might be wise for a risk-averse investor to place only a small fraction of her wealth in the risky option, even when its mean return exceeds that of the safe option. On the other hand, if most members of the group obtain information that suggests that the mean return of the

[3] In a non-Bayesian framework, Baurmann and Brennan (2005) give examples that illustrate potential difficulties of the Pareto Principle for problems involving aggregation of beliefs.

risky investment is high, then it may be optimal for the group to concentrate investments in the risky asset.

In the alternative class of models, individuals have supermodular utility functions. In these problems the optimal decision can take on many different values even under complete information. An example of a problem of this sort is trying to guess the number of balls in an urn. I identify conditions on the distribution of signals that characterize when groups make recommendations that are less extreme than individual guesses. These conditions are restrictive and polarization arises under a variety of natural conditions. If, for example, it is more costly to overestimate the number of balls than to overestimate, and the signals indicate that the urn contains more than the (ex ante) expected number of balls, then the group's guess will be higher than the average individual guess. The group's decision is shaded upward because it has superior information. On one hand, this causes the group to place less weight on the prior estimate of the mean. On the other hand, it causes the group to be more confident of its guess and so less inclined to reduce its guess to avoid losses.

There is an alternative interpretation of the model.[4] Imagine that a single decision maker requests the opinion of I informed experts prior to making a decision. Each expert observes a signal and recommends an action to the decision maker. The decision maker, knowing the information structure and the experts' signals, selects an action. The experts are assumed to report truthfully (or make recommendations consistent with the decision maker's preferences). Under these assumptions my results show that in general it may be rational for the decision maker to select an action different from the advice of the experts. In particular, I give conditions under which the decision maker takes a more extreme action than suggested by any of his advisors.

Eliaz, Ray, and Razin (2005) present the first, and to my knowledge only other, theoretical model of choice shifts. Groups must decide between a safe and a risky choice. The paper summarizes group decision making by a pair of probabilities: the probability that an individual's choice will be pivotal (determine the group's decision) and the probability distribution over outcomes in the event that the individual is not pivotal. In this framework, choice shifts arise if an individual would select a different recommendation alone than as part of a group. If individual preferences could be represented

[4]Douglas Bernheim suggested this interpretation.

by von Neumann-Morgenstern utility functions, then choice shifts do not arise. Eliaz, Ray, and Razin (2005) prove that systematic choice shifts do arise if individuals have rank-dependent preferences consistent with observed violations of the Allais paradox. Moreover, the choice shifts they identify are consistent with experimental results.[5] Assuming that an individual is indifferent between the safe and risky actions in isolation, as a pivotal member of the group, she will choose the safe action if and only if the probability that the group would otherwise choose the safe action is sufficiently high. Unlike my approach, this model does not rely on information aggregation. Eliaz, Ray, and Razin (2005) concentrates on how preferences revealed within groups may differ from preferences revealed individually, but it is not designed to study how deliberations may influence individual recommendations. An appealing aspect of the Eliaz, Ray, and Razin (2005) approach is the connection it makes between systematic shifts in group decisions and systematic violations of the expected utility hypothesis.

Section 6 discusses other related research and contains concluding comments.

2 A Benchmark Model

There are $I > 1$ individuals. Individual i has a utility function that depends on an action[6] $a \in A$ and the state of the world, $\theta \in \Theta$. Denote the utility function by $u_i(\cdot)$. Individuals receive a private signal $s \in S$ about the state of the world. I assume in this section that Θ, A, and S are finite sets. Let $\pi(\theta)$ be the prior probability of state θ. Assume that $\pi(\theta) > 0$ for all $\theta \in \Theta$. Let $P(\theta; \mathbf{s})$ be the joint probability that the state is θ and the profile of signals $\mathbf{s} = (s_1, \ldots, s_I)$; and $p(\theta|\mathcal{I})$ the conditional probability that the state is θ given the information $\mathcal{I}$.[7] It is straightforward to represent $p(\cdot)$ in terms of $P(\cdot)$ and $\pi(\cdot)$. I refer to (Θ, S, π, p, P) as the information structure.

I compare two situations. When individuals act privately, they each select $a_i^*(s_i)$ to maximize $\sum_{\theta \in \Theta} u_i(a, \theta) p(\theta|s_i)$. When individuals act collectively, they select $a_0^*(\mathbf{s})$. In general, $a_0^*(\mathbf{s})$ will depend on the institution by which

[5] Because the set of actions is binary, Eliaz, Ray, and Razin cannot explain situations in which group actions are strictly more extreme than individual actions.

[6] I refer to action choices of individuals as recommendations and action choices of groups as decisions.

[7] $\mathcal{I}$ will be one signal s or a profile $\mathbf{s}$.

agents share information. When preferences differ, it is not clear how the group should decide upon its collective decision. Even when preferences coincide, psychological or strategic considerations may prevent the group decision from being optimal given available information.

I focus on the benchmark case in which the interests of the individuals are the same, $u_i(\cdot) \equiv u(\cdot)$ for all i, and that $a_0^*(\mathbf{s})$ is chosen optimally so that $a_0^*(\mathbf{s})$ solves:

$$\max_{a \in A} \sum_{\theta \in \Theta} u(a, \theta) p(\theta | \mathbf{s}). \tag{1}$$

When $u_i(\cdot)$ is independent of i, $a_i^*(\cdot)$ is also independent of i for $i > 0$. $a_0^*(\cdot)$ is a distinct function because it is a function of signal profiles not of individual signals.

If group polarization is evidence of irrationality, then the optimal group choice must be constrained by the vector of individual choices. Assume that individual recommendations are chosen optimally. An observer knows the actions taken at the group and individual level (but not the information structure). Is it possible for the observer to conclude that a particular collective decision is not optimal? If so, then observing that decision is evidence of irrationality. If not, then the argument that polarization (or any other characteristic of the group decision) is irrational must be re-examined.

The first result describes a property of aggregate beliefs obtained from information aggregation. Proposition 1, stated formally below, says that essentially any finite set of belief profiles can be described by a single information structure. Suppose the observer managed to elicit the beliefs of the group before and after information aggregation in a finite number of situations. Further suppose that all of the beliefs elicited placed positive probability on all of the states. The proposition asserts that there is an information structure that is consistent with these observations in the sense that there are signal profiles that induce all of the observed beliefs. Hence individual beliefs place no constraints on group beliefs.

To state the proposition, define a **belief profile** to be a vector $\mathbf{q} = (q_0; q_1, \ldots, q_I)$ such that each q_i is a probability distribution on Θ. The belief profile $\mathbf{q}$ is **interior** if $q_i(\theta) > 0$ for all i and θ.

Proposition 1 *Let Q be a finite set of interior belief profiles. There exists a signal set S and a joint probability distribution $P(\theta; s_1, \ldots, s_I)$ such that for $\mathbf{q} = (q_0; q_1, \ldots, q_I) \in Q$ there exist a signal profile $\mathbf{s} = (s_1, \ldots, s_I)$ with $P(\theta; \mathbf{s}) > 0$ such that $q_0(\theta) = p(\theta \mid \mathbf{s})$ and $q_i(\theta) = p(\theta \mid s_i)$ for all $i = 1, \ldots, I$.*

The existence of a signal profile **s** satisfying the conclusion of the proposition is, mathematically, the statement that there exists an information structure for which a family of linear inequalities has a solution. The proof of Proposition 1 constructs an information structure with the appropriate characteristics.[8] There is a signal for each belief profile $\mathbf{q^k} \in Q$ and one other residual signal. The signal associated with $\mathbf{q^k}$ is sent to all agents with probabilities that make q_0^k the posterior or to exactly one agent (while the others receive the residual signal) with the probability necessary to guarantee that when the individual i receives signal s^k she updates her beliefs to q_i^k. Such a signaling technology satisfies the conditions of the proposition and is not difficult to construct.

A consequence of Proposition 1 is that individual recommendations place no constraints on the group's recommendation. Let $\mathbf{a} = (a_0; a_1, \ldots, a_I) \in A^{I+1}$ denote an action profile. Interpret a_0 as the joint action and each a_i, $i = 1, \ldots, I$ as an action of individual i. Call a decision $a \in A$ undominated if there exists $q_i \in \text{int}(\Delta)$ such that a solves: $\max_{a \in A} \sum_{\theta \in \Theta} u(a, \theta) q_i(\theta)$.[9] The signal profile $\mathbf{s} = (s_1, \ldots, s_I)$ **induces a** if $a_0 = a_0^*(\mathbf{s})$ and $a_i = a_i^*(s_i)$ for all $i = 1, \ldots, I$. **a** is **possible** if there exists a signal vector **s** that induces **a**.

Proposition 2 *There exists a signal set S and a joint probability distribution $P(\theta; s_1, \ldots, s_I)$ such that for all profiles of undominated actions $\mathbf{a} = (a_0; a_1, \ldots, a_I)$ there exists a signal profile $\mathbf{s} = (s_1, \ldots, s_I)$ with $P(\theta; \mathbf{s}) > 0$ such that* **s** *induces* **a**.

Proposition 2 states that any undominated action profile is possible. Consequently, there need not be any connection between individually optimal and collectively optimal actions. The proposition implies that group decisions that are "extreme" relative to individual choices need not be a sign of irrationality. In particular, if A is ordered, then nothing prevents a_0 from being greater than all of the other components of **a**. Therefore it is premature to assume that the group decision is not optimal even when collective decisions differ systematically from individual recommendations.

Proposition 2 is an immediate consequence of Proposition 1. Since A is finite, only a finite number of distinct action profiles exist. If **a** is one of these

[8] The Appendix contains the proof of Proposition 1 and all subsequent results requiring proof.

[9] The definition rules out degenerate cases in which action a maximizes the expected payoff only if one or more states is assigned probability zero.

profiles, then there exists a belief profile $\mathbf{q}$ such that a_i is a best response to q_i for each $i = 0, 1, \ldots, I$.

The conclusion that no group decision is inconsistent with individual recommendations does not depend on the assumption that agents select a recommendation that maximizes expected utility. The result continues to hold provided that beliefs determine actions (so the preferences can be described by a non-expected utility functional or a behavioral rule of thumb).

Proposition 2 indicates that for general information structures, individual choices place no constraints on the optimal decision of the group. It is possible that these results rely on "strange" information structures. Propositions 1 and 2 depend on the assumption that signals can be correlated. A polar opposite assumption is that individuals receive signals that are conditionally independent. Henceforth, I assume that the information structure can be described by functions $\alpha_i : S \times \Theta \to [0, 1]$, where $\alpha_i(s \mid \theta)$ is the probability that individual i receives signal s given that the state is θ (so that $\sum_s \alpha_i(s \mid \theta) = 1$ for all θ and i).

This environment is considerably more restrictive than the general framework. Proposition 1 asserted that essentially any collection of individual and group posteriors is consistent with some information structure. On the contrary, if individuals receive conditionally independent signals, then the group posterior is determined by the individual posteriors.

Proposition 3 *If the individual signals are conditionally independent, then the group posterior distribution is completely determined given individual conditional beliefs. In particular, if individual i has beliefs q_i, then the group's beliefs are q, where*

$$q(\theta) = \frac{\pi(\theta)\Pi_{i=1}^{I}\left(q_i(\theta)/\pi(\theta)\right)}{\sum_{\omega}\pi(\omega)\Pi_{i=1}^{I}\left(q_i(\omega)/\pi(\omega)\right)}.$$

Proposition 3 follows directly from the Bayes's Rule and the independence assumption.

Although Proposition 3 rules out the strong conclusions of Propositions 1 and 2, Example 1 demonstrates that it is still may be difficult to draw inferences about group decisions from individual recommendations.

Example 1 Suppose that $I > 1$; $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_I)$; individual i observes $s_i = \theta_i$ (that is, individual i observes the i^{th} component of $\boldsymbol{\theta}$ without error);

each component of $\boldsymbol{\theta}$ is independently and uniformly distributed on $\{-1, 1\}$; and $u_i(a, \boldsymbol{\theta}) = -(a - \prod_{i=1}^{I} \theta_i)^2$. An individual sets $a_i^*(s_i) = 0$. The group sets $a_0^*(s) = \prod_{i=1}^{I} s_i$.

Information obtained by an individual (or, in fact, any proper subset of the group) is useless – it conveys no information that improves making decisions while the entire group's information perfectly reveals the state. Individual recommendations therefore do not depend on private information while the group decision does. Knowing everything about individual recommendations provides no information about the group's preferred action. Also, the group's decision is both more variable and more extreme than individual recommendations. □

Unlike the example, the construction in Proposition 2 does permit an observer to draw inferences from individual recommendations. The example differs from the construction because it requires a particular specification of the utility function.

It is possible to generalize the logic of the example.

Proposition 4 *If $I > 1$, then given any finite A, there exists preferences and an information structure such that individual recommendations convey no information about the group decision and all group decisions are possible.*

Example 1 and Proposition 4 are perverse because information from any proper subset of the agents does not lead to better decisions. In the next section, I make further restrictions on the information structure and preferences. I revisit the basic question in a standard, but restrictive, economic environment.

3 Monotone Problems

Since my objective is to show that individual recommendations make weak restrictions on the group decision, I will make strong assumptions on the information structure. Without restrictions on preferences, individual recommendations still place few restrictions on the optimal group decision. Proposition 5 demonstrates that there are restrictions across problems: if the action profile **a** is possible, then some other action profiles are ruled out. Propositions 6 and 7 demonstrate that the optimal group decision is only weakly constrained by individual recommendations. Essentially all action profiles are possible. Subsequent sections present restrictions on preferences might organize findings on polarization.

I concentrate on monotone information structures, which satisfy the following conditions.[10] Assume that A is the unit interval. Assume that $\alpha_i(s \mid \theta)$ is independent of i so that signals are identically (as well as independently) distributed. Assume that S has more than one element and that signals are distinct in the sense that if $s' \neq s$, then $p(\theta \mid s) \neq p(\theta \mid s')$. Further assume that the information structure and the utility function have a monotone structure: the signals satisfy the monotone-likelihood ratio property, so that $\alpha(s \mid \theta)/\alpha(s' \mid \theta)$ is decreasing in θ for all $s' > s$;[11] and that for all $a' > a$, the function $v(\theta; a, a') = u(a', \theta) - u(a, \theta)$ is either increasing in θ (supermodular incremental utility) or there exists θ_0 such that $v(\theta) < 0$ for $\theta < \theta_0$ and $v(\theta) > 0$ for $\theta > \theta_0$ (single-crossing incremental utility). These conditions guarantee that optimal actions are increasing in signals, meaning that $a_i^*(s') \geq a_i^*(s)$ whenever $s' > s$. I will refer to these cases as the **supermodular** and **single-crossing** models, respectively.

Proposition 2 cannot hold for this restricted class of problems because the monotonicity condition imposes a restriction across problems. If one observes two action profiles $\mathbf{a}$ and $\mathbf{a}'$ such that $\mathbf{a}'_{-\mathbf{0}} \geq \mathbf{a}_{-\mathbf{0}}$, then $a_0' \geq a_0$. Hence there does not exist a single monotonic information structure that makes all undominated action profiles possible.

Proposition 5 *For a fixed monotonic information structure, if* $\mathbf{a}$ *and* $\mathbf{a}'$ *are possible and* $\mathbf{a}'_{-\mathbf{0}} \geq \mathbf{a}_{-\mathbf{0}}$*, then* $a_0' \geq a_0$*.*

Proposition 5 is a special case of an observation in Milgrom and Weber (1982, Theorem 5). It is a straightforward implication of the the monotone information structure. If an individual makes a higher recommendation, then she must have received a higher signal. If all signals are higher, then the group decision must also be higher.

To make the subsequent discussion concrete, consider two leading special cases. A monotone model is an **urn model** if $u(a, \theta) = -(a - f(\theta))^2$ for some strictly increasing function $f(\cdot)$.[12] Here $v(\theta)$ is increasing so the urn

[10] Athey and Levin (2000) provide an analysis of a more general class of monotone decision problems.

[11] This condition implies that the posterior distribution of θ given s is increasing in s (in the sense of first-order stochastic dominance).

[12] At this point, strictly increasing transformations of the state space Θ do not change

model is a supermodular model. In the urn model θ represents the number of balls in an urn and $f(\theta)$ a target determined by the number of balls. The agents want to make the best estimate of the target $f(\theta)$.

A monotone model is a **portfolio model** if $u(a, \theta) = U(a\theta + (1 - a)\theta_0)$ where $U(\cdot)$ is a concave function defined over monetary outcomes. A portfolio model is a single-crossing model. The problem is to determine the share of wealth to allocate over a safe asset, which yields θ_0, and a risky one, which yields θ. Individuals must pick the fraction a of the portfolio to invest in the risky asset. Risk averse agents will typically select $a < 1$ even when their information suggests that the mean of θ exceeds θ_0. On the other hand, if sufficiently many agents receive independent information suggesting that the mean return of the risky asset is high, this will induce higher investments in the risky asset when information is pooled.

The next results demonstrate that even in monotone models it is difficult to draw inferences about the group decision merely by observing individual recommendations. In light of Proposition 3, such a result will not be possible if the utility function is completely arbitrary. To see this concretely, suppose that an individual had a utility function with the property that he would select a recommendation $a \leq \underline{a}$ if and only if the probability of state $\underline{\theta}$ was greater than .5. It follows from Proposition 3 that if all individuals made recommendations less than $\underline{a}$, then the group decision must also be less than $\underline{a}$. Consequently, in an environment with independently distributed messages, strong results that suggest that individual recommendations place no restrictions on the group decision cannot specify preferences completely.

Proposition 6 *For all* $\mathbf{a} = (a_0; a_1, \ldots, a_I)$ *with* $a_i \in [0, 1]$ *there exists both an urn model and a portfolio model such that there exists* $\mathbf{s}$ *such that* $\mathbf{s}$ *induces* $\mathbf{a}$*.*

Proposition 6 states that an observer who knows the recommendations of all of the individuals and who knows that a decision problem is either an urn problem or a portfolio problem (but not the specific form of the utility function) still cannot conclude that the group has made an irrational decision. This result is weaker than Proposition 2 for three reasons. First, Proposition 2 constructed one information structure that was compatible with any given (finite) set of recommendation profiles. Proposition 6 instead

the underlying decision problem, so including $f(\cdot)$ in the specification of $u(\cdot)$ adds no generality. The formulation is useful for subsequent applications.

constructs a different information structure for each profile. Proposition 5 explains why the stronger result is not possible in a monotone environment. Second, Proposition 2 holds even if the observer knows the utility function. In Proposition 6 the utility function is selected to support observed behavior. The utility function is not arbitrary, however. It is always possible to find a suitable utility function from the class of urn models and portfolio models that is consistent with the action profile.[13] Finally, the construction requires that there be more than one signal that leads to the same action in some circumstances. To understand and relax this restriction, it is useful to explain the proof of Proposition 6.

To prove Proposition 6, I construct an information structure with the property that if all but two agents receive the lowest possible signal and two others receive the next lower signal, the group posterior is higher than the posterior of individuals who received the second lowest signal. In order for this to be possible, the individual who receives the second lowest signal must place high probability that everyone else will receive the lowest signal. When she learns that this is not true, she (and hence the group) revises her prior strongly upward. Under the assumptions of Proposition 6 it is possible that many signals induce the lowest action. Therefore even if all individuals wish to take the lowest action, they may not have received the lowest signal, and the group may prefer a higher decision. If the optimal action is a strictly increasing function of the signal, however, the conclusion of the proposition must be weakened.

To make a precise statement, let $O_k(\mathbf{a_{-0}})$ be the k^{th} largest of the set $\mathbf{a_{-0}} = \{a_1, \ldots, a_I\}$ (so that $O_1(\mathbf{a_{-0}}) = \max_{i=1,\ldots,I} a_i$, $O_2(\mathbf{a_{-0}})$ is the second highest, and so on).

Proposition 7 *If* $\mathbf{a} = (a_0; a_1, \ldots, a_I)$ *with* $O_2(\mathbf{a_{-0}}) > 0$ *and* $O_{I-1}(\mathbf{a_{-0}}) < 1$, *then there exists both an urn model and a portfolio model with the property that* $a_i^*(s)$ *is strictly increasing such that* $\mathbf{a}$ *is possible.*

The assumptions in Proposition 7 rule out the possibility that individual information would lead all but one agent to make the same extreme recommendation (either the highest or the lowest). Provided these assumptions hold, it can be rational for the group to make any decision. For monotone

[13] In the proof of Proposition 6, $U(x)$ can be taken to be of the form $U(x) = x^\beta$ for $\beta \in (0, 1)$. That is, it is possible to satisfy the conclusion of Proposition 6 using a narrow class of utility functions.

problems in which optimal actions are strictly increasing in the signal, the conditions are necessary. If all but individual i wishes to make the lowest recommendation, learning that all other agents also wish to make the lowest recommendation must be "bad news," which makes the group's decision weakly lower than individual's i optimal recommendation. Hence an observer can place bounds on the possible group decision assuming that all but one individual wants to take the lowest recommendation. Proposition 7 demonstrates that no further restrictions are possible. In particular, the proposition states that it is possible for the group to want to make a more extreme decision than any individual in the group.

4 Invariance

The results in Section 3 imply that even in monotone problems it is premature to conclude that any group decision is irrational given the decisions of individual group members. While staying within the framework of monotone problems, I now identify conditions under which group decisions are well behaved in the sense that they are guaranteed to be bounded by the individual recommendations.

Intuition suggests that for a suitable range of information structures, the group guess in the urn model should be bounded by individual guesses. If everyone thinks that there are between 100 and 300 balls in the urn, then it would be surprising if the group's guess was outside that range.

This section makes the intuition rigorous. To motivate the basic idea, contrast the problem of information aggregation with the problem of preference aggregation. In social choice problems involving aggregation of preferences, it is typical to assume that if all individuals prefer decision X to decision Y, then so should the group. In information aggregation problems, this implication is not clear. Consider the portfolio problem. It could be the case that risk-averse individuals will prefer to invest a substantial fraction of their portfolio in the safe asset even when informed that the mean of θ is greater than θ_0. On the other hand, a large enough number of independent signals that $\theta > \theta_0$ would be sufficient to convince the group to take a more extreme position.

This observation suggests a critical difference between the urn and portfolio models and motivates the following definition.

Imagine a situation in which every member of the group receives the same signal. They would consequently make the same recommendation. Under what conditions would the group decision be the same as the common recommendation of each individual?

Call a monotone decision problem **invariant** if

$$a_i^*(s_i) = a_0^*(s_i, \ldots, s_i) \text{ for all } s_i. \tag{2}$$

That is, a decision problem is invariant if the optimal group decision when all members of the group independently observe the same signal realization is that same as the optimal individual recommendation given one observation of that realization. Proposition 9 provides conditions on the underlying data of the problem that guarantee that (2) holds.

Proposition 8 *In an invariant monotone decision problem, if* **a** *is possible, then*

$$a_0 \in [\min_{1 \le i \le I} a_i, \max_{1 \le i \le 1} a_i]. \tag{3}$$

Proposition 8 is a simple consequence of Proposition 5.

Proposition 8 states that invariant monotone problems are well behaved in the weak sense that the individual recommendations form a bound for the groups' decision. Invariance is a necessary condition for the Proposition 8. If invariance failed, then there would exist a s such that (3) would fail if everyone in the population received that signal.

The next example exhibits an invariant problem by describing a situation where (2) holds.

Example 2 Assume that $\Theta = S = \{0, 1/K, \ldots, K - 1/K, 1\}$; $\pi(\cdot)$ is the uniform distribution on Θ and that $u(a, \theta) = -(a - \theta)^2$ for $a \in A = [0, 1]$. Let $\gamma \in (1/2, 1)$ and

$$\alpha(s, \theta) = \begin{cases} 1 & \text{if } s = \theta \in \{0, 1\}, \\ \gamma & \text{if } s = \theta \in \{1/K, \ldots, (K-1)/K\}, \\ \frac{1-\gamma}{2} & \text{if } s = \theta \pm 1 \text{ and } \theta \in \{1/K, \ldots, (K-1)/K\}, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Individuals seek the best estimate of θ. The signal is the true state plus a symmetric error. Individual i recommends $a_i(s) = a$. If every member of the group receives the same signal, the recommendation is the same. □

One way to get a better understanding of invariance is to think about the condition when I is large. If all I members of the population receive the signal s, then the group's posterior distribution is given by

$$r(\theta \mid s; I) = \frac{\alpha^I(s \mid \theta)\pi(\theta)}{\sum_{\omega \in \Theta} \alpha^I(s \mid \omega)\pi(\omega)}. \tag{5}$$

Let $\Theta(s) = \text{argmax}_\theta \alpha(s \mid \theta)$. It follows that the limiting posterior distribution, $r^*(\theta; s) \equiv \lim_{I \to \infty} r(\theta \mid s; I)$, is given by

$$r^*(\theta; s) = \begin{cases} \frac{\pi(\theta)}{\sum_{\omega \in \Theta(s)} \pi(\omega)} & \text{if } \theta \in \Theta(s), \\ 0 & \text{if } \theta \notin \Theta(s) \end{cases} \tag{6}$$

In particular, if $\alpha(s \mid \theta)$ has a unique maximum $\theta^*(s)$, then $r^*(\cdot; s)$ is the point mass on $\theta^*(s)$. If a decision problem is invariant for all I, then the optimal response to signal s, $a_i^*(s)$, also maximizes $\sum_{\theta \in \Theta} u(a, \theta) r^*(\theta \mid s)$. I conclude that it requires strong assumptions to guarantee that individual recommendations place even mild restrictions on the group decision. Nevertheless, the conditions make sense in the urn model (provided that signals are symmetric estimates of the true state).

It is possible to generalize Example 2. First, assume that the information technology is **non-degenerate**: for each s, $\alpha(s \mid \theta)$ has a unique maximizer, denoted by $\theta^*(s)$. It follows that $r^*(s \mid \theta)$ is a point mass on $\theta^*(s)$ and equation (2) becomes

$$E\{u_a(a_i^*(s), \theta) \mid s\} = u_a(a^*(s), \theta^*(s)). \tag{7}$$

Second, assume that

$$\frac{\sum_{\theta \in \Theta} \alpha(s \mid \theta)\theta\pi(\theta)}{\sum_{\omega \in \Theta} \alpha(s \mid \omega)\pi(\omega)} = \theta^*(s) \text{ for all } s. \tag{8}$$

That is, the mean of θ given s is equal to $\theta^*(s)$ for all s. Call the information technology **uniformly neutral** if equation (8) holds. The first assumption is mild. The second assumption is restrictive. One would expect that the highest signal is "good news" so that receiving multiple independent draws will strictly increase the mean of the distribution. Indeed, while there exist uniformly neutral information technologies (see Example 2), equation (8) requires that the extreme signals completely reveal the extreme states.

The following result is immediate from the definitions.

Proposition 9 *If $u(a,\theta) = -(a-\theta)^2$, and the information structure is non-degenerate and uniformly neutral, then the decision problem is invariant.*

Propositions 8 and 9 combine to identify a class of decision problems in which group decisions are bounded by individual recommendations. Even invariant problems may give rise to apparently extreme behavior, however. There are natural situations under which the recommendation of the most extreme individual becomes the recommendation of the group.

Example 3 Recall that the Pareto distribution with strictly positive parameters θ_0 and β has the probability density function

$$f(\theta|\theta_0,\beta) = \begin{cases} \frac{\beta\theta_0^\beta}{\theta^{\beta+1}}, & \text{when } \theta > \theta_0 \\ 0, & \text{when } \theta \le \theta_0 \end{cases} .$$

The following is a standard property of conjugate distributions (see DeGroot (1970, page 172)).

Fact 1 *Suppose that each of the I agents receives a signal from a uniform distribution on $[0,\theta]$ where θ itself is unknown. Suppose that the prior distribution of θ is the Pareto distribution with parameters θ_0 and β, θ_0 and $\beta > 0$. The posterior distribution of θ given that individual i receives the signal s_i is a Pareto distribution with parameters $\tilde{\mathbf{s}}$ and $\beta + I$, where*

$$\tilde{\mathbf{s}} = \max\{\theta_0, s_1, \ldots, s_I\}.$$

Now assume that $u(a,\theta) = -(a-\theta)^2$. An individual who receives the signal s_i believes that θ has a Pareto distribution with parameters $\tilde{s}_i = \max\{\theta_0, s_i\}$ and consequently, because maximizing $u(\cdot)$ requires choosing a equal to the expected value of θ, selects $a_i^*(s_i) = (\beta+1)\tilde{\mathbf{s}}_i/\beta$ while the collectively optimal choice is $a_0^*(\mathbf{s}) = (\beta+I)\tilde{\mathbf{s}}/(\beta+I-1)$.

In this example, the highest signal provides a lower bound on θ and therefore is a sufficient statistic for all of the signals. When individuals pool their information two things happen: the variance of the distribution of θ decreases, because there is more information;[14] and the maximum signal determines the lower bound of the distribution. That is, when the individuals pool their information, only the signal of the most extreme individual determines the collective decision. Due to the first effect, the collective decision

[14] This follows because the exponent in the Pareto distribution increases.

will be less than the choice of the individual who received the greatest signal, but the ratio of the collectively rational decision to the maximum individual recommendation converges to one as β and I grow.

The specification is special, but could be appropriate for some contexts. For example, imagine that the signal an individual receives indicates the minimum amount of damage that could have been done to a plaintiff. When jurors pool their information, it is only the highest signal that is relevant for estimating damages. Hence, efficient information aggregation implies polarization. Notice that if the objective was to take an action equal to the lowest element of the support of θ, then $a_0^*(\mathbf{s}) = \tilde{\mathbf{s}}$. □

Some signal structures imply that the recommendations of "extreme" individuals dominate the group decision. For example, denote by $e(j)$ the j^{th} unit vector and suppose that there is one signal for each state. Let $\alpha_N = e(N), \alpha_j = \lambda_j e(j) + \mu_j \alpha_{j+1}$, for appropriate non-negative weights, λ_j and μ_j. In this information structure, higher signals rule out lower states and the highest signal is a sufficient statistic. Therefore, the group will always follow the largest individual recommendation. A symmetric construction creates an information structure in which the group follows the smallest individual recommendation. This kind of ordering of signals makes sense in certain applications. For example, if individual information provides a lower bound to damage, then it is reasonable to assume that the group's decision will be largely determined by the individual with the highest individual signal of damage.

5 Cautious and Risky Shifts

Proposition 8 provides conditions under which the group's decision is between the lowest and highest individual recommendations. Polarization is possible when the decision problem is not invariant. In this section, I present sufficient conditions under which polarization arises. The nature of the results differ depending on whether incremental utility is supermodular or satisfies the single-crossing property.

5.1 Single-Crossing Models

For models like the portfolio problem in which the incremental utility $v(\theta)$ is negative for $\theta < \theta_0$ and positive for $\theta > \theta_0$, there is a strong form of polarization. Recall that $r^*(\cdot; s)$ is the limiting posterior distribution assuming that all members of a large population receive the signal s. Let $S_0 = \{s : r^*(\theta; s) = 0 \text{ for } \theta \geq \theta_0\}$ and $S_1 = \{s : r^*(\theta; s) = 0 \text{ for } \theta \leq \theta_0\}$.

Proposition 10 *Assume that the decision problem is monotone and the utility function has single-crossing increments. Given any $\epsilon > 0$, there exists N such that:*

1. Cautious Shift. *If individual signals are all in S_0, then if $I > N$ the group decision is less than ϵ.*
2. Risky Shift. *If individual signals are all in S_1, then if $I > N$ the group decision greater than $1 - \epsilon$.*

If the group is fairly certain that the state is less than θ_0, then the optimal decision is 0. If the group is large and enough agents receive signals in S_0, then the group will be confident that the state is less than θ_0. In the limit, the group decision is weakly more extreme than any individual recommendation. In many interesting situations, individuals will not be confident enough to make extreme recommendations on the basis of one signal. In those cases, the group decision is strictly more extreme than any individual recommendation.

When individuals with similar biases make a recommendation as a group, they treat the tendencies of others as independent evidence in support of their position. Consequently, the group will be more confident in the state of the world's value (relative to θ_0) and, in single-crossing models, more likely to make an extreme recommendation. Proposition 10 demonstrates that models with single-crossing incremental utility exhibit both cautious and risky shifts as identified in the psychology literature. Shifts to extreme positions arise if the group consists of individuals with similar, moderate, tendencies. The form of preferences needed for the result is consistent with the assessments of damages of expermental juries (Schkade, Sunstein, and Kahneman (2000)), portfolio choice,[15] or election polling.

[15]Barber and Odean (2000) report that investment clubs tend to make riskier investments than individuals, which is consistent with the proposition. They also show that returns of investment clubs are lower than those of individual investors (which in turn under performed the market). This result runs counter to my approach, which assumes that groups share information honestly and optimize perfectly. Rather than evidence directly in favor or against my model, I interpret these results more generally as reasons to be skeptical about using simple optimizing models as descriptions of behavior in financial markets.

Proposition 10 has two large limitations. First, it guarantees choice shifts only if the group is arbitrarily large. In fact, the convergence is exponential. For example, assuming a uniform prior, if a single message in S_0 implies that the probability of $\theta < \theta_0$ is .6, then when all members of a five-person group receives messages in S_0, the probability that $\theta < \theta_0$ increases to more than .98. Second, the proposition requires that *every* member of the group receives a signal in S_0. For large groups, the next result demonstrates that polarization is the norm.

Proposition 11 *Assume that the decision problem is monotone and the utility function has single-crossing increments. For any $\epsilon > 0$, there exists N such that if $I > N$, then $\{\mathbf{s} : a_0^*(\mathbf{s}) \in [\epsilon, 1-\epsilon]\}$ has probability less than $\epsilon + \pi(\theta_0)$.*

Proposition 11 states that if the group is large enough, then the group's decision will be polarized "most" of the time. The proposition follows from the law of large numbers. With high probability, the group's posterior will converge to the true state of the world. Provided that the true state of the world is not θ_0, the group will want to take an extreme decision. Hence (adjusting for the prior probability that the true state is θ_0) large groups will make extreme decisions with high probability. Unlike Proposition 10, Proposition 11 provides no information about the direction of polarization.

5.2 Supermodular Models

When members of the group independently receive the same signal, the group information differs from individual information in two ways. The group is likely to have a better estimate of the state of nature and the group may have a different estimate of the mean of θ. In the previous subsection, only the change in precision was relevant. As the group's information improves, the group's recommendation will almost always converge to a boundary recommendation. The direction of polarization depends on whether θ is expected to be less than or greater than θ_0.

The supermodular case is different because the decision is likely to vary continuously with the estimated mean of θ even as the precision of the estimate becomes perfect. There is, however, a simple variation of Proposition 11 that applies to supermodular models provided that perfectly informed groups make extreme recommendations.

Consider the case in which $\Theta = \{\theta_L, \theta_H\}$ has two elements, with $\theta_L < \theta_H$. In a slight abuse of notation, let $a^*(\theta_k)$ be the optimal action given that the state is θ_k, for $k = 1$ and 2. When the problem is monotone, $a_i^*(s) \in [a^*(\theta_L), a^*(\theta_H)]$ for all signals.

Proposition 12 *Assume that Θ has cardinality two and the decision problem is monotone. For any $\epsilon > 0$, there exists a group size N such that if $I > N$, then $\{\mathbf{s} : a_0^*(\mathbf{s}) \in [a^*(\theta_L) + \epsilon, a^*(\theta_H) - \epsilon]\}$ has probability less than ϵ.*

When there are only two possible states of the world, poorly informed individuals will make intermediate recommendations, but well informed groups will make decisions close to the optimal decision for one of the two states.

When there are many states, changes in both precision and mean will influence whether group decisions shift systematically relative to individual ones. Example 4 illustrates how these two effects interact and provides a motivation for the main result of this subsection.

Example 4 Suppose that $\theta \in \mathbb{R}$, $\pi(\cdot)$ is normal with mean μ and precision[16] $\tau > 0$, and that given θ, s is a normal distribution with mean θ and precision $r > 0$. The posterior distribution of θ given K independent signals $\mathbf{s} = (s_1, \ldots, s_K)$ is a normal distribution with mean $\mu^*(\mathbf{s})$ and precision $\tau + Kr$, where

$$\mu^*(\mathbf{s}) = \frac{\tau\mu + r\sum_{i=1}^{K} s_i}{\tau + Kr}. \tag{9}$$

In this example, the posterior distribution depends on the average signal. Furthermore, if $u(a, \theta) = U(a - \theta)$, then $a^*(s) = C + \mu^*(s)$, for some constant C.[17] It follows that the average individual recommendation, $\bar{a}^*(\mathbf{s}) = C + \mu^*(\bar{\mathbf{s}})$ where $\bar{\mathbf{s}}$ is the average signal.

Consider first the case where $U(x) = -x^2$. The optimal recommendation is simply the conditional mean of θ. The conditional mean is an average of the prior mean and the signal for an individual and the average of the prior mean and the average signal for the group. Further, the group places lower weight on the prior mean. Consequently, the group decision will be more

[16]The precision of a normal random variable is the inverse of its variance.

[17]$a_i^*(s)$ solves: $\max \int U(a - \theta) dF(\theta \mid s)$, where $F(\theta \mid s)$ is the distribution of θ given s. By assumption, we can write $F(\theta \mid s) = G(\theta - \mu^*(s))$, where $G(\cdot)$ is a normal distribution with mean 0 and precision $\tau + r$. Hence $\int U(a - \theta) dF(\theta \mid s) = \int U(a - t - \mu^*(s)) dG(t)$, which establishes the claim.

extreme than the average decision because increasing the size of the group that receives a given average signal leads to a systematic **mean shift**: The estimate of the mean (for a fixed average signal) is further from the prior μ the larger the size of the group. Recall that Proposition 9 gave conditions under which problems with $u(a, \theta) = -(a - \theta)^2$ would be invariant. The assumption of uniform neutrality fails in this example because receiving two copies of the same signal leads the group to revise the estimate of the mean.

Proposition 13, stated below, describes the nature of choice shifts when $U(\cdot)$ is not necessarily quadratic. In this case, the losses associated with incorrect guesses are asymmetric. If $U'(\cdot)$ is concave, then overestimates are relatively more costly than underestimates, and individuals (with less precise information) tend to make lower recommendations than groups. When $\bar{\mathbf{s}} \geq \mu$, this effect goes in the same direction as the mean-shift described above. As a result, the group's recommendation is greater than the average recommendation. If $U'(\cdot)$ is convex and $\bar{\mathbf{s}} \leq \mu$, then the group's recommendation is less than the average recommendation.

In this example, it is also possible to show the existence of choice shifts when the population is finite. The recommendation of an individual if the signal is s, $a_i^*(s)$, is the solution to the equation

$$\int U'(a - \theta) p(\theta \mid s) d\theta = 0. \tag{10}$$

On the other hand, given the signal profile $\mathbf{s} = (s_1, \ldots, s_I)$, the group's decision is the solution to the equation

$$\int U'(a - \theta) p(\theta \mid \mathbf{s}) d\theta = 0. \tag{11}$$

By equation (9), the distribution of θ given one signal s has a higher (lower) mean than the distribution of θ given $I > 1$ signals with average value s if $s < \mu$ ($s > \mu$). It follows that the left-hand sides of equations (10) and (11) can be ranked using second-order stochastic dominance. Specifically,

$$\int U'(a - \theta) p(\theta \mid s) d\theta > \int U'(a - \theta) p(\theta \mid \mathbf{s}) \text{ if } U'(\cdot) \text{ is convex and } s < \mu \tag{12}$$

while

$$\int U'(a - \theta) p(\theta \mid s) d\theta < \int U'(a - \theta) p(\theta \mid \mathbf{s}) \text{ if } U'(\cdot) \text{ is concave and } s > \mu. \tag{13}$$

Since the posterior distribution of θ given $\mathbf{s}$ depends on only the mean of $\mathbf{s}$, $a_0^*(\mathbf{s})$ is characterized by equation (10) with $s = \bar{\mathbf{s}}$. It follows that if $U'(\cdot)$ is convex and $\bar{\mathbf{s}} < \mu$, then $a_0^*(\mathbf{s}) < \sum_{i=1}^{I} a_i^*(s_i)/I$ and if $U'(\cdot)$ is concave and $\bar{\mathbf{s}} > \mu$, then $a_0^*(\mathbf{s}) > \sum_{i=1}^{I} a_i^*(s_i)/I$. That is, the group's recommendation is higher than the average individual recommendation when the average individual recommendation is high and $U'(\cdot)$ is concave and the group's recommendation is lower than the average individual recommendation when the average individual recommendation is low and the $U'(\cdot)$ function is convex.

In the example, agents want to find a good estimate for a function of θ. Two factors generate choice shifts. First, the group relies less on its prior information. As a result, its estimate of the mean is more extreme than the average individual estimate. Second, when $U'(\cdot)$ has the specified second-order properties, the group's decision will increase the bias in one direction. Specifically, if $U'(\cdot)$ is concave, recommendations that are too high are more costly than recommendations that are too low. This induces individuals to make lower recommendations than groups (given the same estimate of the mean). Combined with the first effect, this means that risky shifts arise if the group's information favors risky choices. □

The remainder of this section states and discusses a generalization of Example 4. The example compares the group decision to the average recommendation made by individuals in the group. In order to make this kind of comparison, I must impose structure on the state space Θ and the signal space S. Until now, these sets had to be ordered (so that higher states were associated with higher signals and higher signals induced higher actions), but they were otherwise arbitrary. For the rest of the section, assume that the action set is an interval; Θ and S are connected subsets of the real line; the joint distribution admits a smooth density; the distribution and utility function are sufficiently regular so that conditional expected utilities are well defined; the optimal recommendation function $a_i^*(\cdot)$ is a smooth function; and that it is permissible to interchange integration and differentiation.[18] In order to relate signals and states, assume that signals are equal to the true state plus an error, so that $s = \theta + \eta$. Assume that the error η is independent of θ and has a symmetric density $g(\cdot)$. Further assume that there is an affine function $L(s) = ls + m$ and a function $h(\cdot)$ such that the posterior probability

[18]It is sufficient to assume that the third derivative of $u(\cdot)$ is continuous and integrable.

of θ given s, $p(\theta \mid s)$ can be written

$$p(\theta \mid s) = h(\theta - L(s)). \tag{14}$$

I call information structures that satisfy these conditions **additive**. The additivity assumption, which holds in Example 4, implies that the expected signal given θ is equal to θ. Furthermore, it is possible to compare the average recommendation to a recommendation made in response to an average signal.

Lemma 1 *Assume that the information structure is additive and $u(a, \theta) = U(a - f(\theta))$ for U three times differentiable and $U''(\cdot) < 0$ and $f(\cdot)$ strictly increasing. Individual i's optimal recommendation function, $a_i^*(\cdot)$ is a concave [resp. convex] function of the signal if $U'(\cdot)$ and $f(\cdot)$ are concave [resp. convex].*

To get some intuition for Lemma 1, note that a simple change of variables argument (like the one used in Example 4) establishes that $a_i^*(\cdot)$ is linear if $f(\cdot)$ is linear.

It follows from Lemma 1 that (14) combined with concavity assumptions implies that $a_j^*(\bar{\mathbf{s}}) \geq \sum_{i=1}^{I} a_i^*(s_i)/I$. That is, the regularity assumption makes it possible to compare the average response of individuals with an individual's response to an average signal.

Lemma 1 makes strong assumptions on preferences. If I only made ordinal assumptions on the state space, concavity or convexity assumptions in $f(\cdot)$ can be made without loss of generality. However, assumption (14) is invariant only with respect to positive affine transformations of θ. Concavity and convexity assumptions on $U'(\cdot)$ translate into restrictions on the third derivative of $U(\cdot)$. These restrictions have sensible interpretations in the context of the model. Assume that $u(a, \theta) = U(a - \theta)$ for an increasing, concave function $U(\cdot)$. The urn model is a special case in which $U(\cdot)$ is quadratic. In the urn model the costs associated with making an incorrect recommendation are symmetric. Mathematically, this property follows because the second derivative of $U(\cdot)$ is constant. One way to describe situations in which losses are not symmetric is by making assumptions on the second derivative of $U(\cdot)$. Straightforward calculus arguments (provided in the appendix, Lemma 3) guarantee that

- If $U''(\cdot)$ is constant, then for all $x > 0$, $U(0) - U(x) = U(0) - U(-x)$.

- If $U''(\cdot)$ is strictly increasing, then for all $x > 0$, $U(0) - U(x) < U(0) - U(-x)$.
- If $U''(\cdot)$ is strictly decreasing, then for all $x > 0$, $U(0) - U(x) > U(0) - U(-x)$.

In the first case, losses are symmetric and choice shifts may not arise. In the second case, choosing a too small is more costly than choosing a too large. Proposition 13 demonstrates that the group's decision will be less than the average individual recommendation in this case. In the third case, choosing a too large is more costly than choosing a too small. Group decisions will tend to be larger than the average individual recommendation in this case.

Example 4 illustrated the possibility that choice shifts would arise because an individual estimate of the mean of the distribution might differ systematically from the estimate of a group that receives the same average signal. Let $\hat{\theta}$ denote the prior mean of θ and $\hat{\mu}(s)$ denote the conditional mean of θ given s. I assume that the information structure exhibits **mean shift** if $\hat{\mu}(s)$ is between $\hat{\theta}$ and s.[19] The mean shift condition holds in Example 4. More generally, it is straightforward to show that it holds in additive models when the error distribution $g(\cdot)$ is unimodal and symmetric about 0 and the prior distribution of θ ($\pi(\cdot)$) is unimodal and symmetric about $\hat{\theta}$.

Proposition 13 *Assume that the decision problem is monotone, $u(a, \theta) = U(a - f(\theta))$ for U three times differentiable, $U''(\cdot) < 0$, and $f(\cdot)$ increasing, and the information structure is additive and exhibits mean shifts. Let θ^* denote the true state of the world.*

1. *Assume $U'(\cdot)$ and $f(\cdot)$ are convex and $\theta^* < \hat{\theta}$. Given $\epsilon > 0$ there exists N such that if $I > N$, then the probability that $a_0^*(\mathbf{s}) \le \sum_{i=1}^{I} a_i^*(s_i)/I$ is greater than $1 - \epsilon$.*
2. *Assume $U'(\cdot)$ and $f(\cdot)$ are concave and $\theta^* > \hat{\theta}$. Given $\epsilon > 0$ there exists N such that if $I > N$, then $a_0^*(\mathbf{s}) \ge \sum_{i=1}^{I} a_i^*(s_i)/I$ is greater than $1 - \epsilon$.*

The proposition states conditions under which the group's decision is systematically biased relative to the average recommendation of the individuals. Situations in which the average decision of the individuals is larger than the

[19] More precisely, I require that if $s \ge \hat{\theta}$, then $\hat{\mu}(s) \in [\hat{\theta}, s]$ and if $s \le \hat{\theta}$, then $\hat{\mu}(s) \in [s, \hat{\theta}]$.

group's decision arise when $U'(\cdot)$ and $f(\cdot)$ are convex; shifts towards higher group decisions arise when these functions are concave.[20]

The direction of the choice shift depends on whether the state of the world is greater than or less than the prior mean. When the group is large enough, the group's decision will be approximately optimal given θ^*. Hence high group decisions will be higher than the average individual recommendation when overestimates more costly than underestimates (convexity) and low group decisions will be lower than the average individual recommendation when it is more costly to underestimate the optimal decision.

The first part of the proof of Proposition 13 uses Lemma 1 to relate the average of the individual recommendations to the individual best response to the average signal. In the second step of the proof of Proposition 13, I show that when $u_a(a, \cdot)$ is concave (convex),[21] the recommendation of an individual who receives the signal s is less (greater) than the recommendation if the state were known to be equal to the expected state given s. The law of large numbers implies that the average signal approximates the conditionally expected signal given the true state of the world. An individual recommendation based on this signal will be closer to the prior mean than the recommendation of a group that knew the true state. Hence, the second step always me to compare the group decision to how an individual would respond to an average signal, while the first step relates the individual recommendation to an average signal to the average individual recommendation.

When the error distribution $g(\cdot)$ is unimodal, it is possible to prove a finite-sample version of Proposition 13. When the average signal is high (low) and $U'(\cdot)$ and $f(\cdot)$ are convex (concave) the group's decision is higher (lower) than the average individual recommendation.

The conclusions of the proposition are qualitatively different from the conclusions of Propositions 10 and 11. When utility is single-crossing, choice shifts in the same decision problem can go in different directions. If information suggests that the state is less than θ_0, then there is a downward shift; if information suggests that the state is greater than θ_0, then there is an up-

[20]Motivated by the portfolio model, in Proposition 10 it was natural to describe situations in which the group made higher decisions than individuals as risky shifts and ones in which the group selected lower decisions as cautious shifts. Context will determine whether the choice shifts identified in Proposition 13 could be considered cautious or risky.

[21]If $u(a, \theta) = U(a - f(\theta))$ and $U(\cdot)$ is concave, then concavity (convexity) of $U'(\cdot)$ and $f(\cdot)$ imply concavity (convexity) of $u_a(a, \theta)$ so the sufficient conditions on preferences used in Lemma 1 will imply the conditions on preferences imposed in Proposition 13.

ward shift. In contrast, Proposition 13 gives sufficient conditions for a shift in any one direction. The direction of the shift depends on properties of the decision problem.

6 Conclusion

This section discusses some related research, summarizes my contribution, and suggests further research.

When individuals have the same preferences, there is always a mechanism that leads to efficient information aggregation. The mechanism designer asks each individual to report her signal and commits to taking the decision that maximizes expected utility conditional on the signals reported. Honest reporting is an equilibrium. On the other hand, there is no guarantee that institutions found in the world will work as well. Austen-Smith and Banks (1996) observe that sincere voting need not be an equilibrium when majority voting is used to aggregate individual information even when individuals have identical preferences. Feddersen and Pesendorfer (1997) prove that voting does aggregate information efficiently when the population is large and preferences are homogeneous. The literature on informational cascades and observational learning (for example, Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992)) identifies situations in which groups will fail to make use of information efficiently. Studying whether particular institutions for information aggregation lead to polarization would be an interesting extension of my approach.

There is an experimental literature on group decision making that focuses on topics traditionally studied by economists. A fundamental question is whether groups make better decisions than individuals. My model assumes perfect information aggregation, common interests, and optimization. Consequently, the group's recommendation must be better (ex ante) than any individual recommendation and at least as good as any function of individual recommendations. In practice, groups may not perform better than individuals. Blinder and Morgan (2005) ask individuals to make decisions separately and then as groups in a setting where decisions makers choose how much information to obtain. They find that groups tend to make better decisions than individuals whether the group's decision is determined by majority rule or consensus, but that the demand for information did not depend on whether the decision was made by an individual or a group. In strategic contexts, Cooper and Kagel (2005) study a limit pricing game and report that groups learn equilibrium strategies faster than individuals. Kocher and Sutter (2005) make a similar finding in an experimental beauty contest.

Several other papers try to identify differences in group and individual preferences in strategic settings. Bornstein and Yaniv (1998) study individual versus group behavior in a standard, one-shot ultimatum game, where a fixed amount of money c is split between a proposer and a responder. Bornstein and Yaniv compare two treatments, one with individuals playing against individuals and one with groups (of three subjects each) playing against groups. Their main result is that groups are more (game-theoretically) rational players than individuals by demanding more than individuals in the role of proposer and accepting relatively lower offers in the role of responder.[22] Cox (2002) finds that groups behave more like payoff maximizers than do individuals when they play the role of the second mover (trustee) in the trust game of Berg, Dickhaut, and McCabe (1995). In a sharp contrast to these results, Cason and Mui (1997) find that in dictator games groups are more trusting than individuals.[23]

Bone, Hey, and Suckling (1999) demonstrate that groups are prone to the same kinds of systematic violations of the expected utility hypothesis as individuals. I assume that both individuals and groups maximize expected utility (although Propositions 2 and 12 do not require this assumption).

This paper compares the decisions of individuals and groups for information aggregation problems. I show that generally there is no systematic relationship between recommendations taken by individuals in isolation and the decision that the individuals would take as a group. I then provide conditions under which polarization will and will not arise.

I establish my results in a narrow setting. I assume that groups have no problems aggregating information and reaching a joint decision. Anyone who has even served on a committee will know that these assumptions are unrealistic.

Experiments have identified choice shifts in a wide range of settings. This paper provides a theoretical foundation for the existence of choice shifts in a model of information aggregation. In future work, I hope to demonstrate

[22]Bornstein, Kugler and Ziegelmeyer (2004) reach a similar conclusion in a study of the centipede game.

[23]Cason and Mui (1997) interpret their results as evidence in favor of the social comparison theory since their groups shift in the direction of generous behavior, which is perceived to be socially desirable.

that the model of the paper organizes experimental evidence on polarization.

Appendix

Proof of Proposition 1. Given q^k and $C_k > 0$, define $\lambda_i^k(\theta)$ by

$$\lambda_0^k(\theta) = C_k q_0^k(\theta) \tag{15}$$

and for $i = 1, \ldots, I$ by

$$\lambda_i^k(\theta) = C_k \left(q_i^k(\theta) - q_0^k(\theta)\right) + d_k q_i^k(\theta), \tag{16}$$

where $d_k = C_k q_0^k / (\min_i q_i^k(\theta))$. The choice of d_k guarantees that $\lambda_k(\cdot) > 0$.

It follows from equation (15) that

$$q_0^k(\theta) = \frac{\lambda_0^k(\theta)}{\sum_{\omega\in\Theta} \lambda_0^k(\omega)}. \tag{17}$$

Further, it follows from equations (15) and (16) that

$$(d_k + C_k) q_i^k(\theta) = \lambda_0^k(\theta) + \lambda_i^k(\theta) \tag{18}$$

and so

$$q_i^k(\theta) = \frac{\lambda_0^k(\theta) + \lambda_i^k(\theta)}{\sum_{\omega\in\Theta}(\lambda_0^k(\omega) + \lambda_i^k(\omega))}. \tag{19}$$

Now consider a signaling technology in which there is a signal s^k for each k and an additional signal $\tilde{s}$. Let

$$P(\theta; \mathbf{s}) = \begin{cases} \lambda_0^k(\theta) & \text{if } s_j = s^k \text{ for all } j, \\ \lambda_i^k(\theta) & \text{if } s_i = s^k \text{ and } s_j = \tilde{s} \text{ for all } j \neq i, \\ \pi(\theta) - \left(\sum_k \lambda_0^k(\theta) + \sum_{i,k} \lambda_i^k(\theta)\right) & \text{if } s_j = \tilde{s} \text{ for all } j, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

By taking C_k sufficiently small, it is possible to make $P(\cdot) > 0$.

If the joint distribution of θ and s is given by $P(\cdot)$, then it follows from equation (17) that if the group receives the signal profile $\mathbf{s} = (s^k, \ldots, s^k)$ for some k, then the group posterior is $q^k(\cdot)$, while equation (19) implies that if individual i receives s^k, then her posterior is $q_i^k(\cdot)$. ■

Proof of Proposition 3. If individual i has beliefs $q_i(\cdot)$ given the signal s_i, then it follows from Bayes's Rule that the probability individual i receives signal s_i given θ, $\alpha_i(s_i \mid \theta)$, satifies:

$$\alpha_i(s_i \mid \theta) = \mu_i(s_i)\frac{q_i(\theta)}{\pi(\theta)}, \tag{21}$$

where $\mu_i(s_i) = \sum_\omega \alpha_i(s_i \mid \omega)\pi(\omega)$ is the probability that individual i receives s_i. Consequently, after any vector of signals $\mathbf{s} = (s_1, \ldots, s_I)$ that gives rise to the belief profile $\{q_1, \ldots, q_I\}$, the group's posterior is

$$\frac{\left(\Pi_{i=1}^I \alpha_i(s_i \mid \theta)\right)\pi(\theta)}{\sum_\omega \left(\Pi_{i=1}^I \alpha_i(s_i \mid \omega)\right)\pi(\omega)} = \frac{\pi^{-(I-1)}(\theta)\Pi_{i=1}^I q_i(\theta)}{\sum_\omega \pi^{-(I-1)}(\omega)\Pi_{i=1}^I q_i(\omega)}, \tag{22}$$

where the equation follows from equation (21) (the normalization factors $\mu_i(\cdot)$ cancel out). This completes the proof. ■

Proof of Proposition 4. If A has N elements, let A and S be the integers $1, \ldots, N$. Let $\Theta = A \times I$ and $\pi(\cdot)$ be the uniform distribution on Θ. For $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_I) \in \Theta$, let $f(\boldsymbol{\theta}) = 1 + \sum_{i=1}^I \theta_i \pmod N$ and $u_i(a, \boldsymbol{\theta}) = -(a - f(\boldsymbol{\theta}))^2$. An individual is indifferent over all $a \in A$ while the group sets $a_0^*(\mathbf{s}) = f(\mathbf{s})$. ■

Lemma 2 *There exists a monotone information structure with the property that if all but two of the individuals receive the lowest signal and two others receive the next lowest signal, then the posterior distribution given the group's information is greater (in the sense of first-order stochastic dominance and the monotone likelihood ratio property) than any individual posterior.*

Proof of Lemma 2. Assume that there are only two signals, s_{-1} and s_{I+1} and the technology satisfies $\alpha(s_{-1} \mid \theta_k) = b^{k-1}$ for $b \in (0, 1)$ and $k = 1, \ldots, N$ and $\alpha(s_{I+1} \mid \theta_k) = 1 - \alpha(s_{-1} \mid \theta_k)$. When the group receives a profile of individual signals $\mathbf{s}$ containing $I - 2$ copies of the s_{-1} and two copies of the signal s_{I+1}, then the posterior distribution satisfies, for $k = 1, \ldots, N$,

$$P(\theta_k \mid \mathbf{s}) = \lambda\pi(\theta_k)b^{(k-1)(I-2)}\left(1 - b^{k-1}\right)^2, \tag{23}$$

where λ is a normalization constant selected so that $\sum_{k=1}^{N} P(\theta_k \mid s) = 1$.

On the other hand, the posterior probability of state θ_k given s_{I+1} (as an individual signal) is $\mu(1 - b^{k-1})\pi(\theta_k)$, where μ is another normalization constant. I claim that if b is sufficiently close to one, then

$$\frac{P(\theta_k \mid \mathbf{s})}{P(\theta_{k+1} \mid \mathbf{s})} = \left(\frac{1}{b}\right)^{I-2} \left(\frac{1 - b^{k-1}}{1 - b^k}\right)^2 \frac{\pi(\theta_k)}{\pi(\theta_{k+1})} < \frac{1 - b^{k-1}}{1 - b^k} \frac{\pi(\theta_k)}{\pi(\theta_{k+1})}. \tag{24}$$

Inequality (24) is equivalent to

$$b^{I-2} \frac{1 - b^k}{1 - b^{k-1}} > 1. \tag{25}$$

Since the left-hand side of inequality (25) converges to $k/(k-1)$ as b approaches one, the claim follows. It follows from inequality (24) that the profile of signals $\mathbf{s}$ leads to a posterior that dominates the individual signal s_{I+1}. ■

Proof of Proposition 6. Add additional signals to the technology constructed in Lemma 2 that are mixtures of the signals s_{-1} and s_{I+1}. For $j = 0, \ldots, I$ define s_i so that $s_{-1} < s_0 < \cdots < s_{I+1}$ and

$$\alpha^*(s_j \mid \theta) = \lambda(\theta) \left(c_j \alpha(s_{-1} \mid \theta) + (1 - c_j)\alpha(s_{I+1} \mid \theta)\right) \text{ for } j = -1, \ldots, I+1, \tag{26}$$

where $\sum_{j=-1}^{I+1} \alpha^*(s_j \mid \theta) = 1$ for all θ, $0 \leq c_I < \cdots < c_0$, and c_1 is close enough to zero so that if the group receives I copies of the signal s_1, then the posterior will still dominate the posterior given only the signal s_{I+1}. This is possible by continuity.

First, I show that given a profile of actions $(a_1, \ldots, a_I)$ ordered so that $a_1 \leq a_2 \leq \cdots \leq a_I$, it is possible to pick c_i and a utility function so that $a_i^*(s_j) = a_j$ for $j = 1, \ldots, I$ and $a_0^*(s_1, \ldots, s_I) = 1$. Set $a_0 = 0$ and let $a_{I+1} \in [a_I, 1]$. Let $u(a, \theta) = -(a - f(\theta))^2$. I claim that it is possible to find $\lambda(\theta)$, c_i, and a strictly increasing $f(\cdot)$ such that for $j = 0, \ldots, I+1$,

$$\sum_\theta \frac{\alpha^*(s_j \mid \theta)\pi(\theta)}{\sum_\omega \alpha^*(s_j \mid \omega)\pi(\omega)} f(\theta) = a_j. \tag{27}$$

To establish the claim, define A_i and B_i for $i = -1$ and $I+1$ as:

$$A_i = \sum_\theta \alpha_i^*(\theta)\pi(\theta)f(\theta) \tag{28}$$

and

$$B_i = \sum_{\theta} \alpha_i^*(\theta)\pi(\theta). \tag{29}$$

Using (26), equation (27) can be written:

$$c_j A_{-1} + (1 - c_j)A_{I+1} = a_j \left(c_j B_{-1} + (1 - c_j)B_{I+1}\right). \tag{30}$$

It follows that for $j = 0, \ldots, I$

$$c_j = \frac{A_{I+1} - a_j B_{I+1}}{A_{I+1} - A_{-1} - a_j(B_{I+1} - B_{-1})}. \tag{31}$$

The fact that the posteriors are ranked by the monotone likelihood ratio property and the monotonicity of $\{a_j\}$ guarantee that it is possible to find $f(\cdot)$ such that the values of c_j defined in (31) are non-negative and decreasing. This establishes the claim.

When $u(a_i, \theta) = -(a_i - f(\theta))^2$, equation (27) guarantees that $a_i^*(s_j) = a_j$, for $j = 1, \ldots, I+1$. Since the posterior distribution given $(s_1, \ldots, s_I)$ dominates the posterior given s_{I+1}, a suitable choice of $a_{I+1} \in (a_I, 1)$ guarantees that $a_0^*((s_1, \ldots, s_I)) = 1$.

This construction therefore guarantees that it is possible to create an information structure in which the group's decision is 1 no matter what the individual recommendations are. The same type of construction can be used to create an information structure in which the group's decision is 0 given any individual recommendation. It is straightforward to modify the argument to information structures that induce group decisions that are inside the range of individual recommendations.

The construction demonstrated the proposition was true when the utility function took the form: $u(a_i, \theta) =-(a_i - f(\theta))^2$ for an increasing function $f(\cdot)$. One can modify the argument to show that it is possible to take $u(a_i, \theta) = (a_i\theta + (1 - a_i)\theta_0)^\beta$ for appropriate choices of $\theta_0 > 0$ and $\beta < 1$. Specifically, let

$$\theta_0 = \sum_{\theta} \frac{\alpha^*(s_0 \mid \theta)\pi(\theta)\theta}{\sum_{\omega} \alpha^*(s_0 \mid \omega)\pi(\omega)}. \tag{32}$$

The definition of θ_0 in (32) and the fact that the distribution generated by the signal s_{I+1} dominates that of s_0 guarantees that there exists $\beta \in (0, 1)$ such that

$$\sum_{\theta} \frac{\alpha^*(s_{I+1} \mid \theta)\pi(\theta)\theta}{\sum_{\omega} \alpha^*(s_{I+1} \mid \omega)\pi(\omega)} \theta^{\beta-1}(\theta - \theta_0) = 0. \tag{33}$$

If $u(a_i, \theta) = (a_i\theta + (1 - a_i)\theta_0)^\beta$, then equation (32) guarantees that the best response to s_0 is the action 0 and equation (33) guarantees that the best response to s_{I+1} is the action 1. Having constructed the utility function, given $(a_1, \ldots, a_I)$ it is routine to find appropriate values of c_i so that $a_i^*(s_i) = a_i$ for $i = 1, \ldots, I$. ■

Proof of Proposition 7. Suppose that $\mathbf{a} = (a_1, \ldots, a_I)$, with $0 \leq a_1 \leq \cdots \leq a_I$ and $a_{I-1} > 0$. Proposition 6 implies the result unless $a_1 = 0$. If $a_1 = 0$, then set $s_{-1} = s_0 = s_1$, but construct the information structure as in the proof of Proposition 6 so that $a_i^*(s_i) = a_i$ and $a^*(s) = 1$ (which, provided $a_{I-1} > 0$, is still possible, since the posterior given s will dominate the posterior given the signal associated with a_I). ■

Proof of Proposition 8. Without loss of generality, let $a_1 = \min_{1 \leq i \leq I} a_i$ and $a_I = \max_{1 \leq i \leq I} a_i$. Let s_i satisfy $a_i^*(s_i) = a_i$. By invariance, $a_1 = a_1^*(s_1) = a_0(s_1, \ldots, s_1)$ and $a_I = a_I^*(s_I) = a_0(s_I, \ldots, s_I)$. By monotonicity, $s_I \geq s_1$ and $s_i \in [s_1, s_I]$ for all i. It follows from Proposition 5 that $a_0^*(s_1, \ldots, s_I) \in [a_1, a_I]$. ■

Proof of Proposition 10. I prove the first part of the proposition. The second part follows from a symmetric argument. Let $\hat{s}$ be the largest element of S_0. By monotonicity, the group's decision given that all signals are in S_0 is less than or equal to the decision given that all signals are $\hat{s}$. There exists $\delta > 0$ such that if the group's posterior places probability of at least $1 - \delta$ on S_0, then the group's decision is no larger than ϵ. If all I signals are $\hat{s}$, then the group's posterior probability that the state is θ, $q^I(\theta)$, is

$$q^I(\theta) = \frac{\alpha^I(\hat{s} \mid \theta)\pi(\theta)}{\sum_{\omega \in \Theta} \alpha^I(\hat{s} \mid \omega)\pi(\omega)}. \tag{34}$$

By the definition of $\hat{s}$, there exists $\theta' < \theta_0$ such that $\alpha(\hat{s} \mid \theta') > \alpha(\hat{s} \mid \theta)$ for all $\theta \geq \theta_0$. It follows from equation (34) that for all $\theta \geq \theta_0$,

$$\lim_{I \to \infty} \frac{q^I(\theta)}{q^I(\theta')} = 0. \tag{35}$$

The proposition follows from equation (35). ■

Proof of Proposition 11. Given any $\epsilon > 0$ there exists $\delta > 0$ such that if the posterior probability of $\theta > \theta_0$ given $\mathbf{s}$ is less than δ then $a_0^*(\mathbf{s}) < \epsilon$ and if the posterior probability of $\theta < \theta_0$ given $\mathbf{s}$ is less than δ then $a_0^*(\mathbf{s}) > 1 - \epsilon$. The proposition now follows because, by the law of large numbers, with probability one the posterior probability converges to a degenerate distribution on the true state.[24] ■

Proof of Proposition 12. By the law of large numbers, with probability one the posterior probability converges to a degenerate distribution on the true state. The result follows from continuity. ■

Lemma 3 *Suppose that $U : \mathbb{R} \to \mathbb{R}$ is a twice continuously differentiable function such that $U'(0) = 0$. If $U''(\cdot)$ is strictly decreasing [resp. strictly increasing], then for all $a > 0$, $U(0) - U(a) >$ [resp. $<$] $U(0) - U(-a)$.*

Proof of Lemma 1. $a_i^*(s)$ is the solution to the equation:

$$\int u_a(a, \theta) p(\theta \mid s) d\theta = 0. \tag{36}$$

Since $p(\theta \mid s) = h(\theta - L(s))$ equation (36) can be written (using the charge of variables $\rho = \theta - L(s)$)

$$\int u_a(a, \rho + L(s)) h(\rho) d\rho = 0. \tag{37}$$

Differentiation of equation (37) shows that if $u(a, \theta) = U(a - f(\theta))$, then

$$a_i^{*''}(s) = l^2 \int f''(\rho + L(s)) h(\rho) d\rho - \left(a_i^{*'}(s) - l f'(\theta)\right)^2 \frac{\int U'''(a, \rho + L(s)) h(\rho) d\rho}{\int U''(a, \rho + L(s)) h(\rho) d\rho}, \tag{38}$$

where $L'(s) \equiv l$. It follows from equation (38) and the concavity of $U(\cdot)$ that $a_i^*(\cdot)$ is concave when $U'(\cdot)$ and $f(\cdot)$ are concave and convex when $U'(\cdot)$ and $f(\cdot)$ are convex.[25] ■

[24] DeGroot (1970, pages 202-204) proves precisely this variation of the strong law of large numbers.

[25] If $f(\cdot)$ is linear, then linearity of $a_i^*(\cdot)$ follows with no further assumptions on $U(\cdot)$.

Proof of Lemma 3. By Taylor's Theorem and $U'(0) = 0$, it follows that

$$U(x) - U(0) = \frac{x^2 U''(\eta)}{2} \tag{39}$$

for η between 0 and x. Equation (39) implies that for $a > 0$, $U(0) - U(a) >$ [resp. $<$] U(0) - U(-a) provided that for all $y, z > 0$, $U''(y) <$ [resp. $>$] $U''(-z)$. ■

Proof of Proposition 13. Assume that $u_a(\cdot)$ is strictly concave in θ and $\theta^* \geq \hat{\theta}$. The argument for $u_a(\cdot)$ strictly convex is analogous. It follows from Jensen's inequality that $E\{u_a(a, \theta) \mid s\} < u_a(a, \hat{\mu}(s))$ (where $\hat{\mu}(s)$ is the conditional mean of θ given s) and, in particular,

$$0 = E\{u_a(a_j^*(s), \theta) \mid s\} < u_a(a_j^*(s), \hat{\mu}(s)), \tag{40}$$

where the equation follows from the definition of $a_j^*(s)$.

Because the information structure exhibits mean shifts and $\theta^* > \hat{\theta}$ it follows that $\hat{\mu}(\theta^*) \leq \theta^*$. It follows from supermodularity of $u(\cdot)$ that

$$u_a(a, \theta^*) > u_a(a, \hat{\mu}(\theta)). \tag{41}$$

The law of large numbers implies that $\bar{\mathbf{s}}$ converges to the conditional mean of s given $\theta = \theta^*$, which, by additivity, is equal to θ^* and that the group's posterior converges to a point distribution on θ^*. Therefore, inequalities (40) and (41) imply that given any $\epsilon > 0$ there exists N such that if $I > N$, then $u_a(a_j^*(\bar{\mathbf{s}}), \theta^*) > 0$. From concavity of $u(\cdot, \theta)$ in its first argument,

$$a_0^*(\mathbf{s}) > a_j^*(\bar{\mathbf{s}}) \tag{42}$$

for a set of signal profiles that has probability at least $1 - \epsilon$.

Assume now the stronger concavity conditions in Lemma 1 hold. It follows from Lemma 1 that $a_j^*(\bar{\mathbf{s}}) \geq \sum_{i=1}^{I} a_i^*(s_i)/I$. The proposition now follows from inequality (42). ■

References

ARROW, K. J. (1963): *Social Choice and Individual Values.* Yale University Press.

ATHEY, S., AND J. LEVIN (2000): "The Value of Information in Monotone Decision Problems," Discussion paper, Stanford University.

AUSTEN-SMITH, D., AND J. S. BANKS (1996): "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review*, 90, 34–45.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, pp. 797–817.

BARBER, B. M., AND T. ODEAN (2000): "Too Many Cooks Spoil the Profits: The Performance of Investment Clubs," *Financial Analyst Journal*, pp. 17–25.

BAURMANN, M., AND G. BRENNAN (2005): "Majoritarian Inconsistency, Arrow Impossibility and the Comparative Interpretation: A Context-Based View," Discussion paper, Australian National University.

BERG, J., J. DICKHAUT, AND K. MCCABE (1995): "Trust, Reciprocity and Social History," *Games and Economic Behavior*, 10, 122–144.

BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades," *Journal of Political Economy*, 100(5), 992–1026.

BLINDER, A. S., AND J. MORGAN (2005): "Are Two Heads Better Than One?: Monetary Policy by Committee," *Journal of Money, Credit, and Banking*, 37(5), 789–811.

BONE, J., J. HEY, AND J. SUCKLING (1999): "Are Groups More (or Less) Consistent than Individuals?," *Journal of Risk and Uncertainty*, 18(1), 63–81.

BORNSTEIN, G., T. KUGLER, AND A. ZIEGELMEYER (2004): "Individual and Group Decisions in the Centipede Game: Are Groups More 'Rational' Players?," *Journal of Experimental Social Psychology*, 40, 599–605.

Bornstein, G., and I. Yaniv (1998): "Individual and group behavior in the ultimatum game: Are groups more 'rational' players?," *Experimental Economics*, 1, 101–108.

Brown, R. (1986): *Social Psychology.* New York : Free Press, second edn.

Cason, T., and V. Mui (1997): "A Laboratory Study of Group Polarisation in Team Dictator Game," *Economic Journal*, 107, 1465–1483.

Cooper, D. J., and J. H. Kagel (2005): "Are Two Heads Better Than One? Team versus Individual Play in Signaling Games," *American Economic Review*, 95(3), 477–509.

Cox, J. C. (2002): "Trust, reciprocity, and other-regarding preferences: groups vs. individuals and males vs. females," in *Advances in Experimental Business Research*, ed. by R. Zwick, and A. Rapoport, pp. 331–350. Kluwer Academic Publishers.

DeGroot, M. (1970): *Optimal Statistical Decisions*. McGraw-Hill, New York.

Eliaz, K., D. Ray, and R. Razin (2005): "A Decision-Theoretic Basis for Choice Shifts in Groups," Discussion paper, New York University.

Fedderson, T., and W. Pesendorfer (1997): "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 65(5), 1029–1058.

Isenberg, D. J. (1986): "Group Polarization: A Critical Review and Meta-Analysis," *Journal of Personality and Social Psychology*, 50(6), 1141–51.

Kocher, M. G., and M. Sutter (2005): "The Decision Maker Matters: Individual Versus Group Behaviour in Experimental Beauty-Contest Games," *The Economic Journal*, 115, 200–223.

Milgrom, P. R., and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50(5), 1089–1122.

Myers, D. G., and H. Lamm (1976): "The Group Polarization Phenomenon," *Psychological Bulletin*, 83(4), 602–627.

Schkade, D., C. R. Sunstein, and D. Kahneman (2000): "Deliberating About Dollars: The Severity Shift," *Columbia Law Review*, 100, 1139–1175.

Stoner, J. A. F. (1968): "Risky and Cautious Shifts in Group Decisions: The Influence of Widely Held Values," *Journal of Experimental Social Psychology*, 4, 442–459.

Turner, J. C. (1991): *Social Influence*. Open University Press, Milton Keynes.