On the relationship between individual and group decisions *

Joel Sobel
University of California, San Diego

July 3, 2012

Abstract

Individuals with identical preferences each receive a signal about the unknown state of the world and decide upon a utility-maximizing recommendation on the basis of that signal. The group makes a decision that maximizes a common utility function assuming perfect pooling of the information in individual signals. An action profile is a group action and a recommendation from each individual. A collection of action profiles is rational if there exists an information structure under which all elements in the collection arise with positive probability. With no restrictions on the information structure, essentially all action profiles are rational. In fact, given any distribution over action profiles it is possible to find an information structure that approximates the distribution. In a monotone environment in which individuals receive conditionally independent signals, essentially any single action profile is rational, although some collections of action profiles are not. *Journal of Economic Literature* Classification Numbers: A12, D01; Keywords: statistical decision problem; group polarization; behavioral economics; psychology; forecasting.

*I am happy to acknowledge useful comments from several seminar audiences and from Doug Bernheim, Antonio Cabrales, Juan Carrillo, Vince Crawford, Todd Davies, Eddie Dekel, Rick Harbaugh, Navin Kartik, David Laibson, Mark Machina, Philip Neary, Debraj Ray, Michael Rothschild, Nicolas Roux, David Schkade, Ricardo Serrano, Ross Starr, Joel Watson, three referees, and Co-editor Martin Osborne. I am grateful to the Guggenheim Foundation, NSF, and the Secretaría de Estado de Universidades e Investigacíon del Ministerio de Educación y Ciencia (Spain) for financial support. I worked on the project while visiting the Departament d'Economia i d'Història Econòmica, Institut d'Anàlisi Econòmica of the Universitat Autònoma de Barcelona, and the Institute for Advanced Studies at the Hebrew University of Jerusalem. I thank these institutions for hospitality and administrative support.

1 Introduction

A decision maker asks several informed experts for advice prior to making a decision. Each expert recommends an action to the decision maker. The recommended action maximizes the decision maker's utility given the information available to the expert. The decision maker selects an action. Suppose that the decision maker appears to ignore the recommendations of the experts – say by taking an action that is not in the convex hull of these recommendations. Can we conclude that the decision maker is irrational? More generally, can an observer examine the experts' recommendations and form a useful estimate of the decision maker's rational action?

A scientist has access to several independent forecasts of the weather. Is there any relationship between these forecasts and the optimal aggregate forecast based on all available information? Is the optimal aggregate forecast bounded by the individual forecasts?

Individuals receive information about the facts in a lawsuit. They are asked in isolation to recommend punitive damage awards. Later, these individuals meet together as a jury, deliberate, and make a collective decision on punitive damages. How does a rational jury's decision depend on the expressed recommendations of individual jury members prior to deliberation?

The paper studies a simple model of information aggregation designed to give answers to these questions. Individuals have common preferences but different information. Each individual observes a signal and, using the signal, makes a recommendation. The recommendation maximizes the individual's utility given the signal. The group then pools all of the individual signals and makes a decision based on the pooled information. There is no conflict of interest between group members and the group perfectly aggregates the information of its members. In this setting, I investigate to what extent the group's optimal decision is constrained by the recommendations of the individuals.

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. Asking individuals to make recommendations separately and then as a group generates an **action profile** (consisting of a recommendation from each individual and a separate group decision). A collection of action profiles is **rational** if there exists an information structure under which all elements in the collection arise with positive probability.

Section 2 demonstrates in different ways that individual recommendations do not constrain the group action. In particular, I show that for fixed preferences essentially any finite collection of action profiles is rational. Hence essentially no finite data set consisting of individual recommendations and group actions is inconsistent with rational decision making. When preferences are not fixed, a much stronger result is possible. Here for any given probability distribution over action profiles, there exists a specification of preferences and information that generates a probability distribution over action profiles arbitrarily close to the given distribution. These results say that there is no connection between individual recommendations and the group's optimal decision.

In Section 3 I assume that the information structure is monotone. The state of nature, signals, and action are real numbers. Higher signals are stochastically associated with higher states of the world. I further assume that preferences are restricted so that higher beliefs about the state induce higher actions. In this environment, individual recommendations do constrain group actions. If all individuals recommend a strictly higher action, then the corresponding group action must be higher. This cross-profile restriction need not hold in non-monotone problems. On the other hand, even in monotone models the group's optimal action is not restricted by the individual recommendations.

Suppose (as in Section 3) actions are elements of the real line. Here there is a natural, weak notion of moderation. The group's decision is moderate if it is bounded by the lowest and highest recommendations. Group decisions would be moderate if they were weighted averages of the individual recommendations. The results in Section 2 and 3 demonstrate that moderate problems are special. In Section 4, while still assuming a monotone structure, I identify a (restrictive) condition that is necessary and sufficient for moderation.

An example illustrates why group decisions are unlikely to be moderate and provides intuition for other results in the paper.[1] Suppose that the state of the world is normally distributed with a known mean and variance and that individuals receive a signal that is equal to the true state plus noise. The noise is normally distributed with mean zero and individual signals are conditionally independent given the state. The common objective is to minimize the distance between the recommendation and the state of the world, so recommendations equal to the expected value of the state given available information. Suppose an individual's signal is higher than the prior mean. Her recommendation will be a weighted average of the signal and the prior mean. If every member of the group makes the same recommendation, then the group is more confident in the information content of the signal than any individual. The group's action is therefore greater than the individual recommendations. This follows because the aggregate signal as more informative than an individual signal and hence the group places relatively less weight on the prior mean than an individual. It is not surprising that a group of individuals with similar biases makes decisions that are less extreme than recommendations of individual group members.

Research on "expert resolution" (Winkler 1986) studies rules one might use to aggregate the opinions of experts. A typical rule is a function from individual opinions to an aggregate opinion. This literature contains examples that show linear aggregation rules need not be optimal. The literature on combining forecasts (starting from Bates and Granger 1969) also looks at ways in which to aggregate estimates from different sources. This literature is more in the tradition of classical (rather than Bayesian) statistics so it is not directly comparable to the approach of this paper, but like the research on expert resolution, focuses on identifying aggregation rules that are optimal within a parametric family. These models assume that individual expert opinions or forecasts are available to the modeler. The literature asks for optimal ways to aggregate the recommendations of experts. My results indicate that there is no reason to expect aggregation rules in a parametric class to perform well. In particular, it may be rational for a decision maker who receives identical recommendations from several reliable but partially informed experts to ignore the common recommendation and make another decision. Hence rules that generate an aggregate decision by taking a convex combination of individual recommendations may not perform well.

[1] Roux and Sobel 2012 discuss the example in more detail.

There is a large literature in social psychology on "group polarization."[2] Group polarization refers to the tendency of groups to make decisions that are more extreme than some central tendency of the individual positions of the members of the group. The phenomenon, first observed in experiments reported by Stoner 1968, has been widely replicated. Experiments typically elicit recommendations from individuals and then put individuals in groups and record a decision made by the group. Hence the experiments record individual recommendations and group decisions. The structure of these experiments justifies my modeling approach. The experiments observe individual recommendations and a group decision. I assume the modeler observes these data. The experiments also provide opportunities for individuals to arrive at the group decision. Neither the experiments nor I model the deliberation process explicitly. I assume that somehow individuals use all available information optimally.[3]

The experiments on group decision making performed by Schkade, Sunstein, and Kahneman 2000 provide a concrete example of the third motivating example. Individual subjects received information relevant to a hypothetical court case. They each recorded a punitive verdict and a damage verdict. Subjects then were randomly placed into groups of six; these groups deliberated and decided on punitive and damage verdicts. In this setting, the experimenter observes the individual recommendations about verdict and damage and the group opinion about these quantities. The experimenter does not observe detailed information about these quantities. My model assumes that the deliberation process aggregates this information. When pre-deliberation juror judgments favored a high punishment rating, deliberation tended to increase the rating of the group relative to the median individual rating. When pre-deliberation juror judgments favored a low punishment rating, deliberation tended to decrease the rating of the group relative to the median individual rating. Hence both group punishment and damage awards were more extreme than individual awards. These results suggest a systematic relationship between the group's action and the individual recommendations (group polarization). In particular, the group does not moderate individual recommendations. This paper presents a framework that explains why moderation is not a general property of information aggregation.

The psychology literature often interprets polarization as a sign that group interactions lead to non-optimal decisions and introduces behavioral explanations for the experimental results.[4] This paper demonstrates that polarization is consistent with rational decision making of both groups and individuals. The experimental literature finds polarization arises systematically in a wide range of settings. This observation leads to the question of whether there are environments in which polarization is not only possible, but likely. A companion paper, Roux and Sobel 2012, gives conditions under which one might expect the polarization found in the experiments of Schkade, Sunstein, and Kahneman 2000 and Stoner 1968.

Eliaz, Ray, and Razin 2006 present the first, and to my knowledge only other, decision-

[2] Brown's 1986 devotes a long chapter to the topic. Isenberg 1986 provides a review.

[3] In some examples, there is a one-to-one relationship between individual recommendations and individual information. In these cases, individual recommendations are sufficient statistics for individual information.

[4] Brown 1986 surveys the results and the psychological theories.

theoretical model of choice shifts.[5] 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 by von Neumann-Morgenstern utility functions, then choice shifts would not arise. Eliaz, Ray, and Razin 2006 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.[6] Assuming that an individual is indifferent between the safe and risky actions in isolation, she will choose the safe action when a pivotal member of the group 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 2006 concentrate 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 2006 approach is the connection it makes between systematic shifts in group decisions and systematic violations of the expected utility hypothesis.

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, but for reasons not captured in my model.

2 A Benchmark Model

There are $I > 1$ individuals. Individual i has a utility function that depends on an action[7] $a \in A$ and the state of the world, $\theta \in \Theta$. Denote the utility function by u. Each individual receives a private signal $s \in S$ about the state of the world. I assume in this section that Θ, A, and S are finite. 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 is $\mathbf{s} = (s_1, \ldots, s_I)$; and $p(\theta \mid \mathcal{I})$ the conditional probability that the state is θ given the information $\mathcal{I}$.[8] Note that $\pi(\theta) = \sum_{\mathbf{s}} P(\theta; \mathbf{s})$ and that 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

[5] Dixit and Weibull 2007 demonstrate that when individuals have heterogeneous priors, the arrival of new information may cause their posteriors to diverge. In this way, information may lead to polarization of beliefs. Dixit and Weibull do not compare group beliefs (or actions) to those of the individuals within the group.

[6] 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.

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

[8] $\mathcal{I}$ is one signal s or a profile $\mathbf{s}$.

structure and to (S, P, u) as the decision problem.

I compare two situations. When individuals act privately, they each select $a_i^*(s_i)$ to maximize $\sum_{\theta \in \Theta} u(a, \theta) p(\theta \mid s_i)$. When individuals act collectively, they select $a_0^*(\mathbf{s})$. In general, $a_0^*(\mathbf{s})$ depends on the institution by which 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 in which $a_0^*(\mathbf{s})$ is chosen optimally so that $a_0^*(\mathbf{s})$ solves

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

Since $u(\cdot)$ is independent of i, $a_i^*(\cdot)$ is also independent of i for $i > 0$. The group decision $a_0^*(\cdot)$ is a different function because it depends on signal profiles not of individual signals.

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 collective decision is not optimal? If so, then observing that action is evidence that the group decision was incorrect. If not, then the argument that polarization (or any other tendency of the group decision) is irrational must be re-examined.

This section contains a series of results that suggest that individual recommendations do not constrain the group's action. Propositions 1 and 2 show that given essentially[9] any distribution of action profiles there is an information structure under which these action profiles arise with positive probability. These results make no restrictions on preferences. While they show that any combination of individual recommendations and group actions is possible, they do not rule out the possibility that one can make inferences about the action of the group from the individual recommendations. Proposition 3 demonstrates that the distribution of group beliefs does place restrictions on individual beliefs and characterizes these restrictions. Proposition 4 states that there are preferences and an information structure consistent with essentially any pattern of individual recommendations and group decisions. Finally, Proposition 6 shows that similar results are possible even when individuals receive conditionally independent signals.

Proposition 1 describes a property of aggregate beliefs obtained from information aggregation. 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 place 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.

[9] The formal statements of the propositions make the meaning of "essentially" precise.

Proposition 1. *Let Q be a finite set of interior belief profiles. There exist a signal set S and a joint probability distribution $P(\theta; s_1, \ldots, s_I)$ such that for every $\mathbf{q} = (q_0; q_1, \ldots, q_I) \in Q$ there exists 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 $\mathbf{s}$ 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.[10] There is a signal s^k for each belief profile $\mathbf{q^k} \in Q$ and one distinct residual signal. When an individual receives the signal s^k her updated belief is q_i^k. When all individuals receive the signal s^k the group's posterior is $\mathbf{q^k}$. Such a signaling technology satisfies the conditions of the proposition and is not difficult to construct.

A simple consequence of Proposition 1 is that individual recommendations place no constraints on the group's decision. 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 an action $a \in A$ undominated if there exists $q_i \in \text{int}(\Delta)$ such that a solves $\max_{a \in A} \sum_{\theta \in \Theta} u(a, \theta) q_i(\theta)$.[11] 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 $i = 1, \ldots, I$. The action profile $\mathbf{a}$ is **possible** if there exists a signal profile $\mathbf{s}$ that induces $\mathbf{a}$.

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 $\mathbf{s}$ induces $\mathbf{a}$.*

Proposition 2 states that any undominated action profile is possible. Dominated action profiles are not possible and so observing a dominated action is evidence that someone failed to optimize. The proposition demonstrates that there need not be any connection between individually optimal and collectively optimal actions. In particular, 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 $\mathbf{a}$. 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 $\mathbf{a}$ is one of these 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).

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

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

There is a possible extension to Proposition 1 and 2. One could ask: Does there exist an information structure that gives rise to any joint distribution over belief and action profiles? An affirmative answer to this question would be the strongest possible "anything goes" result. It would say that it is not only possible to rationalize any finite set of observations (as in Proposition 2), but it is also possible to rationalize any distribution over observations. The next two results investigate constraints placed on distributions over belief and action profiles.

The information structure (Θ, S, π, p, P) induces a probability distribution over belief profiles in a natural way. Each $\mathbf{s}$ determines a belief profile and P determines the probability of each $\mathbf{s}$. Proposition 1 demonstrates that for any given family of interior belief profiles, there is an information structure for which each member of the family arises with positive probability. The information structure may induce other belief profiles and Proposition 1 says nothing about the induced probability distribution over belief profiles. For example, the information structure constructed in Proposition 1 may induce the belief profiles in Q with arbitrarily small probability. Proposition 3 characterizes the set of distributions over belief profiles that can be generated by an information structure.

Proposition 3. *Given a positive integer K and $k \in \{1, \ldots, K\}$, fix a finite family of belief profiles $Q = \{\mathbf{q}^k = (q_0^k; q_1^k, \ldots, q_I^I)\}$ and positive numbers r_k such that $\sum_{k=1}^K r_k = 1$. If for all k the information structure (Θ, S, π, p, P) induces belief profile $\mathbf{q}^k = (q_0^k; q_1^k, \ldots, q_I^I)$ with probability r_k, then*

1. *$\sum_{k=1}^K q_i^k r_k = \sum_{k=1}^K q_0^k r_k$ for all $i = 1, \ldots, I$ and*
2. *there exists $\lambda_i^k(j)$, $i = 1, \ldots, I$ and $j, k = 1, \ldots, K$, such that for all i, j and k, $\lambda_i^k(j) \geq 0$, and, for all i and j, $\sum_{k=1}^K \lambda_i^k(j) = 1$ and $q_i^j = \sum_{k=1}^K \lambda_i^k(j) q_0^k$.*

Conversely, if (1) and (2) hold, then for any $\varepsilon > 0$, there exist a signal set S, a joint probability distribution $P(\theta; s_1, \ldots, s_I)$, and, for $k = 1, \ldots, K$, belief profiles $\tilde{\mathbf{q}}^k$ and positive numbers $\tilde{r}_k$ with $\sum_{k=1}^K \tilde{r}_k = 1$, such that the information structure (Θ, S, π, p, P) induces the belief profile $\tilde{\mathbf{q}}^k$ with probability $\tilde{r}_k$ for $k = 1, \ldots, K$, and

$$| \mathbf{q}^k - \tilde{\mathbf{q}}^k | < \varepsilon \text{ and } | r_k - \tilde{r}_k | < \varepsilon.$$

An implication of Proposition 3 is that not all distributions over belief profiles can be generated by an information structure. The necessary conditions in the proposition are intuitive and follow directly from Bayes's Rule. Condition 1 states that the average belief of individual i is equal to the average belief of the group. Condition 2 states that any belief that arises with positive probability for individual i is in the convex hull of group beliefs. Since individual i knows the information structure, she can compute $\{q_0^k\}$; $\lambda_i^k(j)$ is the condition probability that individual i believes that the group's belief will be q_0^k given that the belief she forms based on her own information is q_i^j. Given (2) it is straightforward to verify that (1) holds if and only if $\sum_{j=1}^K r_j \lambda_i^k(j) = r_k$.

Conditions (1) and (2) are approximately sufficient. Given a probability distribution over belief profiles and a positive ε, there is an information structure that leads to a distribution over (approximately) these belief profiles within ε of the given distribution.

The next result describes which distributions over action profiles can be induced by some decision problem. Proposition 3 suggests that it is not possible to generate all distributions over action profiles because individual actions contradict information in the joint distribution. To see this clearly, imagine a distribution over action profiles in which the group action is constant. In this case, all individuals must recommend the group's action with probability one since an individual can infer that the group – acting with better information – will always take this action. Hence certain distributions of action profiles are not consistent with my basic model. It turns out that this concern is only a technicality. The next proposition demonstrates that there is a decision problem that gives rise to a distribution on action profiles that is arbitrarily close to any given joint distribution.

Recall that an action profile is an element $\mathbf{a} = (a_0; a_1, \ldots, a_I) \in A^{I+1}$. A distribution on action profiles is a probability distribution on A^{I+1}. A decision problem (S, P, u) induces a distribution on action profiles. The probability of $\mathbf{a} = (a_0; a_1, \ldots, a_I) \in A^{I+1}$ is the probability of

$$\{\mathbf{s} \in S^I : a_i^*(s_i) = a_i, i = 1, \ldots, I; a_0^*(\mathbf{s}) = a_0\}.$$

Proposition 4. *Given any $\varepsilon > 0$ and any joint distribution on action profiles ρ there exists a decision problem (S, P, u) such that distribution of action profiles induced by (S, P, u) is within ε of ρ.*

Proposition 4 states that it is not possible to refute the hypothesis that the group is rational without making a priori restrictions on the information structure or preferences. The difference between Proposition 2 and Proposition 4 is that the first result fixes preferences and then shows that any action profile is possible, while the second result provides a stronger conclusion (rationalizing any distribution over actions) but requires a possibly different specification of preferences for every distribution.

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 1-4 depend on the assumption that signals can be correlated. A more restrictive 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 asserts 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 5. *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*

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

Proposition 5 follows directly from the Bayes's Rule and the independence assumption.

Although Proposition 5 rules out the strong conclusions of Propositions 1 and 2, Example 1 demonstrates that it still may be difficult to draw inferences about group decisions from individual recommendations.

Example 1. Suppose that $\theta = (\theta_1, \ldots, \theta_I)$; individual i observes $s_i = \theta_i$ (that is, individual i observes the i^{th} component of θ without error); each component of θ is independently and uniformly distributed on $\{-1, 1\}$; and $u_i(a, \theta) = -(a - \prod_{i=1}^{I} \theta_i)^2$. An individual sets $a_i^*(s_i) = 0$ for all s_i. The group sets $a_0^*(\mathbf{s}) = \prod_{i=1}^{I} s_i$ for all $\mathbf{s}$.

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. $\square$

Unlike the construction in 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 6. *There exists a decision problem (S, P, u) such that all action profiles are possible and the distribution of the group action $a_0^*(\mathbf{s})$ is independent of s_i for each i.*

Proposition 6 provides conditions under which individual recommendations convey no information about the group decision. To prove Proposition 6, I generalize the example by exhibiting preferences and an information structure under which no individual signal conveys information about the optimal action.[12] The information structure exhibits a strong form of complementarity in that no useful inferences can be drawn from any proper subset of the signals. If the prior is such that individuals are indifferent over all actions ex ante, then they continue to be indifferent after they receive their private signals. Hence observing their individual recommendations conveys no information about the optimal group action. In this setting, not only are arbitrary action profiles possible, there need be no relationship between the distribution of individual recommendations and the group recommendations.

Example 1 and Propositions 6 are perverse because information from any proper subset of the agents does not lead to better decisions than the decision a single individual would make. In the next section, I make further restrictions on the information structure and preferences. I then revisit the basic question in a standard, but restrictive, economic environment.

[12] See Börgers, Hernando-Veciana, and Krähmer 2011 for a useful analysis of complementarity of signals.

3 Monotone Problems

The previous section shows that with minimal restrictions on the information structure there need not be any connection between the group's decision and individual recommendations. In this section I make strong restrictions on the information structure and investigate the extent to which the results in Section 2 continue to hold. Proposition 7 demonstrates that there are restrictions across problems: if the action profile $\mathbf{a}$ is possible, then some other action profiles are ruled out. This result contrasts with those of Section 2 and indicates that, within the class of decision problems that I consider in this section, individual recommendations place testable restrictions on the behavior of the group when there are several observations. On the other hand, Propositions 8 and 9 demonstrate that when one observes only one action profile, the optimal group decision is only weakly constrained by individual recommendations.

This section studies monotone information structures, which satisfy the following conditions. The set A is the unit interval. The conditional probability of a signal s given the state θ, $\alpha_i(s \mid \theta)$, is independent of i so that signals are identically (as well as independently) distributed. To avoid trivial cases, I assume that S has more than one element and that signals are distinct in the sense that if $s' \neq s$, then $p(\cdot \mid s) \neq p(\cdot \mid s')$. The information structure and the utility function have a monotone structure: I assume that both Θ and S are linearly ordered and the signals satisfy the monotone-likelihood ratio property, so that $\alpha(s \mid \theta)/\alpha(s' \mid \theta)$ is decreasing in θ for all $s' > s$;[13] 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 $s' > s$.[14] 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 $\mathbf{a}'_{-0} \geq \mathbf{a}_{-0}$, then $a_0' \geq a_0$. Hence there does not exist a single monotonic information structure that makes all undominated action profiles possible.

Proposition 7. *For a fixed monotonic information structure, if* $\mathbf{a}$ *and* $\mathbf{a}'$ *are possible and* $a_i' > a_i$ *for all* $i = 1, \ldots, I$*, then* $a_0' \geq a_0$*.*

Proposition 7 is a special case of Theorem 5 in Milgrom and Weber 1982. It is a straightforward implication of the the monotone information structure. If an individual makes a strictly higher recommendation, then she must have received a strictly 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

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

[14] When incremental utility is supermodular, optimal actions are increasing in signals whenever posteriors are ordered by first-order stochastic dominance. When incremental utility is single crossing, optimal actions are increasing in signals when signals satisfy the monotone-likelihood ratio property. See Athey and Levine 2001.

function $f(\cdot)$.[15] Here $v(\theta; a', a)$ is increasing in θ whenever $a' > a$ so the urn 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 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 induces 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 5, such result are not be possible if the utility function is completely arbitrary. To see this concretely, suppose that the prior is uniform and an individual has a utility function with the property that he selects a recommendation $a \leq \underline{a}$ if and only if the probability of state $\underline{\theta}$ is greater than .5. It follows from Proposition 5 that if all individuals make recommendations less than $\underline{a}$, then the group posterior places probability of more than .5 on the event that $\theta = \underline{\theta}$. Consequently, the group's decision is also be less that $\underline{a}$. It follows that restrictions on preferences when combined with a monotone information structure may cause individual recommendations to constrain the group's decision. Proposition 8 shows that the link between individual recommendations and group decisions does not exist without restrictions on preferences.

Proposition 8. *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 8 states that an observer who knows the recommendations of all of the individuals and who knows that a monotone 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 constructs one information structure that is compatible with any given (finite) set of recommendation profiles. Proposition 8 instead constructs a different information structure for each profile. Proposition 7 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 8 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.[16] Finally, the construction

[15] Strictly increasing transformations of the state space Θ do not change the underlying decision problem, so including $f(\cdot)$ in the specification of $u(\cdot)$ is just a relabeling of the states.

[16] In the proof of Proposition 8, $U(x)$ 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 8 using a narrow class of utility functions.

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 8.

To prove Proposition 8, 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 receive 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 8 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 member 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 9. *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 for all* i *such that* $\mathbf{a}$ *is possible.*

The assumptions in Proposition 9 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 problems in which optimal actions are strictly increasing in the signal, the conditions are necessary. To see this, suppose that all but individual i wishes to make the lowest recommendation, so $O_2(\mathbf{a}_{-0}) = 0$. In this case, the optimal group decision a_0 must be no larger than a_i, since learning that all other agents 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 9 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 were 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. 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. Consider the portfolio problem. It could be the case that risk-averse individuals 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 $\theta > \theta_0$ is sufficient to convince the group to take a more extreme position.[17]

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. While (2) is easy to understand, it is a statement about endogenously determined quantities. Proposition 11 provides conditions on the underlying data of the problem (the information structure and the utility function) that guarantee that (2) holds.

Proposition 10. *1. In an invariant monotone decision problem, if* **a** *is possible, then*

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

2. Any monotone decision problem in which (3) holds whenever **a** *is possible is invariant.*

Proposition 10 (1) is a simple consequence of Proposition 7. It states that invariant monotone problems are well behaved in the weak sense that the individual recommendations form a bound for the group's decision. To see Proposition 10 (2), notice that if invariance failed, then there would exist an s such that (3) would fail if everyone in the population received that signal.

When the sets of actions and states are small, it is not hard to construct invariant problems. For example, if there are only two actions, $\{h, l\}$, and two states, $\{H, L\}$, and it is uniquely optimal to take h (resp. l) if and only if the probability of H (resp. L) is greater than one half, then any monotonic information structure in which, for every signal, the posterior never gives probability one half to both states is invariant.

[17] 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.

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 $u(a, \theta) = -(a-\theta)^2$ for $a \in A = [0, 1]$. Let $\gamma \in (0, 1/2)$ and

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

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 (and the posterior places more weight on $s = \theta$). □

Clemen and Winkler 1990 say that a decision maker satisfies the **unanimity principle** if she accepts a forecast if both of her two advisors agree. Invariant problems with binary actions necessarily satisfy the unanimity principle. Clemen and Winkler show that the principle fails in general when there are two states of the world, but provide an example under which the unanimity principle holds. They also discuss the **compromise principle** which is equivalent to invariance in their context.

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 $\Theta(s) = \text{argmax}_\theta \alpha(s \mid \theta)$. In 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 $\sum_{\theta \in \Theta} u(a, \theta) r^*(\theta \mid s)$.

It is unlikely for a decision problem to be invariant for all I. When $\alpha(\cdot)$ must be a continuous function on $[0, 1] \times [0, 1]$, invariance fails for a given utility function on a set of information structures that is open and dense with respect to the sup norm.

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 (2) becomes

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

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 (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 strictly increases the mean of the distribution. Indeed, while there exist uniformly neutral information technologies (see Example 2), (8) requires that the extreme signals completely reveal the extreme states.

The following result is immediate from the definitions.

Proposition 11. *If $u(a, \theta) = -(a - \theta)^2$, and the information structure is non-degenerate and uniformly neutral, then the decision problem is invariant.*

Propositions 10 and 11 combine to identify a class of decision problems in which group decisions are bounded by individual recommendations.[18]

The next example describes a natural situation 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

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

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*

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

[18]Chambers and Healy 2010 study a related problem in which they characterize information structures with the property that the posterior mean given a signal lies between the prior mean and the signal. When preferences take the form $u(a, \theta) = -(a - \theta)^2$ (so that recommendations are equal to posterior means) and there exists an uninformative signal, invariant information structures must satisfy updating towards the mean.

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;[19] 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 is 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. This problem is not invariant. While the group decision is bounded above by the maximum individual recommendation, it is possible for the group decision to be lower than the recommendation of every member of the group.

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 that the recommendations of the most extreme member of the group determines the group decision. $\square$

5 Conclusion

This paper compares the decisions of individuals and groups for information aggregation problems. I show that generally there is no systematic relationship between recommendations individuals make in isolation and the decision that the individuals make as a group. I then identify restrictive situations in which individual recommendations bound the decision of the group.

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 knows that these assumptions are unrealistic.

The weight of academic and popular evidence convinces me that groups can often make bad decisions for systematic reasons, that the reasons can be evaluated, and that institutions can be created to ameliorate the problems. The decision-making environment at NASA has been blamed for several tragedies in the U.S. space program. Janis's 1983 discussion of groupthink among President Kennedy's national security advisors foreshadows the recent failures of United States intelligence agencies. My model does not refute the existence flawed group decision making, but it does point out that apparent anomalies in group behavior are consistent with a simple, rational model of information aggregation.

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

Appendix

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

$$\lambda_0^k(\theta) = C_k q_0^k(\theta) \tag{9}$$

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

$$\lambda_i^k(\theta) = C_k \left(q_i^k(\theta) - q_0^k(\theta)\right) + d_k q_i^k(\theta), \tag{10}$$

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

Because $q_i^k(\cdot)$ is a probability distribution, it follows from these definitions that

$$\sum_{\theta \in \Theta} \lambda_0^k(\theta) = C_k \tag{11}$$

and, for $i > 0$,

$$\sum_{\theta \in \Theta} \lambda_i^k(\theta) = d_k. \tag{12}$$

It follows from (9) and (11) that

$$q_0^k(\theta) = \frac{\lambda_0^k(\theta)}{\sum_{\omega \in \Theta} \lambda_0^k(\omega)}. \tag{13}$$

Further, it follows from (9) and (10) that

$$(d_k + C_k) q_i^k(\theta) = \lambda_0^k(\theta) + \lambda_i^k(\theta) \tag{14}$$

and so, by (11) and (12), that

$$q_i^k(\theta) = \frac{\lambda_0^k(\theta) + \lambda_i^k(\theta)}{\sum_{\omega \in \Theta}(\lambda_0^k(\omega) + \lambda_i^k(\omega))}. \tag{15}$$

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

$$P(\theta; \mathbf{s}) = \begin{cases} \lambda_0^k(\theta) & \text{if } s_j = s^k \text{ for all } j, \\ \lambda_i^k(\theta) & \text{if } s_i = s^k \text{ and } s_j = \tilde{s} \text{ for all } j \neq i, \\ \pi(\theta) - \sum_k \lambda_0^k(\theta) + \sum_{i,k} \lambda_i^k(\theta) & \text{if } s_j = \tilde{s} \text{ for all } j, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

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 (13) that if the group receives the signal profile $\mathbf{s} = (s^k, \ldots, s^k)$ for some k, then the group posterior is $\mathbf{q^k}(\cdot)$, while equation (15) implies that if individual i receives s^k, then her posterior is $q_i^k(\cdot)$. ■

Proof of Proposition 3. Replace q_0^k by a rational distribution $\tilde{q}_0^k$, where $\tilde{q}_0^k(\theta) = m_k(\theta)/M$ (and $\sum_\theta m_k(\theta) = M$) such that $\mid q_0^k(\theta) - \tilde{q}_0^k(\theta) \mid < \varepsilon$ for all θ and k.

Create a set Λ by taking N copies of Θ. A typical element of Λ is (θ, m) for $\theta \in \Theta$ and $m = 1, \ldots, M$. Assume that there is a uniform distribution on Λ. Consider a partition of Λ into sets Λ_k,

$$\Lambda_k = \{(\theta, M_{k-1}(\theta) + 1), \ldots, (\theta, M_k(\theta)) : \theta \in \Theta\},$$

where $M_k(\theta) = \sum_{j=1}^k m_j(\theta)$ (and $M_0(\theta) = 0$ by convention).

Assuming a uniform distribution on Λ, learning the partition element Λ_k induces the belief $\tilde{q}_0^k$ because for each θ, Λ_k has $m_k(\theta)$ elements associated with θ.

Now form Γ. An element of Γ is of the form $\mathbf{g} = (g_1, \ldots, g_I)$, where $g_i \in \{1, \ldots, G\}$ and G is sufficiently large so that there is a rational approximation to $r = (r_1, \ldots, r_K)$, $\tilde{r}$, with $\tilde{r}_k = n_k/G$ and a rational approximation $\tilde{\lambda}_i^k(j)$ to $\lambda_i^k(j)$ such that $\tilde{\lambda}_i^k(j) = I l_i^k(j)/G^I$. The approximation partitions Γ into sets Γ_k such that the cardinality of Γ_k is n_k. Associate every element in Γ_k with Λ_k.

Note that a subset H of Γ corresponds to a probability distribution over the Λ_k (the fraction of elements of H in Γ_k is the probability of Λ_k), which in turn can be viewed as a probability distribution over Θ (because each Λ_k corresponds to the distribution q_0^k).

I claim that one can induce the distribution over belief profiles in which $\tilde{q}^k$ arises with probability $\tilde{r}_k$. Associate with each $\mathbf{g}$ in Γ_j subsets $H_i(\mathbf{g}) \subset \Gamma$ for each i. $H_i(\mathbf{g}) = \mathbf{g} \cup G_i(j)$, where $G_i(j)$ consists of $l_i^k(j)$ elements from Γ_k for $k \neq j$ and $l_i^j(j) - 1$ elements from Γ_j and $G_i(j) \cap G_{i'}(j) = \emptyset$ for $i \neq i'$. It follows that $\cap_{i=1}^I H_i(\mathbf{g}) = \mathbf{g}$. On the other hand, with the information in $H_i(\mathbf{g})$ i believes that the probability that $\mathbf{g}$ is in Γ_k is $\tilde{\lambda}_i^k(j)$. Using the natural association of elements in Γ to elements in Λ and to Λ_k and beliefs, this means that $H_i(\mathbf{g})$ induces beliefs $\tilde{q}_i^j$, which completes the construction. ■

Proof of Proposition 4. Given a distribution over action profiles ρ and $\varepsilon > 0$, let $\hat{\rho}$ be a distribution over action profiles that assigns positive, rational probability to all action profiles and is within ε of ρ. Denote the N elements of A by the integers $1, \ldots, N$. Without loss of generality, assume that the marginal probability of a_0 is a uniform distribution over $\{1, \ldots, N\}$ (if the probability that $a_0 = j$ is k_j/K under $\hat{\rho}$ create a new action set in with $\sum_{j=1}^N k_j$ elements, replacing each state j by k_j identical copies).

Let $\Theta = A^I \times A^I$. Denote an element of Θ by $(\boldsymbol{\theta}, \mathbf{b})$ where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_I) \in A^I$ and $\mathbf{b} = (b_1, \ldots, b_I) \in A^I$. $S = A \times A$. States and signals are selected as follows. A uniform distribution selects $\boldsymbol{\theta}$. The remainder of the state is selected so that the distribution of $(b_0; b_1, \ldots, b_I)$ is given by $\hat{\rho}$, where $b_0 = \sigma(\boldsymbol{\theta}) = 1 + \sum_{i=1}^I \theta_i \pmod N$. In state $(\boldsymbol{\theta}, \mathbf{b})$ individual i receives the signal (θ_i, b_i). Finally, define $u(\cdot)$ so that

$$u(a, \boldsymbol{\theta}, \mathbf{b}) = \begin{cases} 1 & \text{if } a = \sigma(\boldsymbol{\theta}), \\ -c(a, \sigma(\boldsymbol{\theta}), \mathbf{b}) & \text{if } a \neq \sigma(\boldsymbol{\theta}) \end{cases}$$

and

$$c(a, \sigma(\boldsymbol{\theta}), \mathbf{b}) = \begin{cases} C & \text{if } \mathbf{b} = (b, \ldots, b) \text{ and } a \neq b, \\ 0 & \text{otherwise.} \end{cases}$$

The optimal decision for the group is to set $a_0^* = \sigma(\boldsymbol{\theta})$. This earns payoff 1. All other actions receive non-positive payoffs. For sufficiently large C, the optimal action for individual i is to set $a_i^* = b_i$. This action earns positive payoff (equal to the probability that $a_i^* = \sigma(\boldsymbol{\theta})$). Any other decision earns negative payoff for sufficiently large C (this requires that $\hat{\rho}$ assign strictly positive probability to all action profiles). It follows that the induced distribution of action profiles is equal to $\hat{\rho}$. ■

It may be useful to describe the construction in the proof of Proposition 4 in a bit more detail. The objective of the group is to guess the target $\sigma(\boldsymbol{\theta})$. If it does so, then its payoff is 1. Otherwise, the group may pay a cost. When it pools its information, the group has enough information to compute the target exactly. Individuals do not. Interpret $\mathbf{b}$ as a recommended action profile. If an individual follows the recommendation, then it does not pay a cost. If an individual ignores the recommendation, then with positive probability it pays a cost. (In order to maintain symmetry across individuals, the cost is paid only when each individual receives the same recommendation.) Individuals therefore have a choice between following the recommendation and earning a positive expected payoff or ignoring the recommendation and paying a cost with positive probability. By assumption, no individual has perfect information about the group's best action. Consequently there is a positive probability that failure to follow the recommended action will trigger the cost. For sufficiently large C it will be strictly optimal for individuals to follow the recommendation.

Proof of Proposition 5. If individual i has belief $q_i(\cdot)$ 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:

$$\alpha_i(s_i \mid \theta) = \mu_i(s_i)\frac{q_i(\theta)}{\pi(\theta)}, \tag{17}$$

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

$$\frac{\pi(\theta)\Pi_{i=1}^I \alpha_i(s_i \mid \theta)}{\sum_\omega \pi(\omega)\Pi_{i=1}^I \alpha_i(s_i \mid \omega)} = \frac{\pi^{-(I-1)}(\theta)\Pi_{i=1}^I q_i(\theta)}{\sum_\omega \pi^{-(I-1)}(\omega)\Pi_{i=1}^I q_i(\omega)}, \tag{18}$$

where the equation follows from (17) (the normalization factors $\mu_i(\cdot)$ cancel out). This completes the proof. ■

Proof of Proposition 6. Denote the N elements of A by the integers $1, \ldots, N$. Let $S = A$, $\Theta = A^I$ and $\pi(\cdot)$ be the uniform distribution on Θ. For $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_I) \in \Theta$, let $\sigma(\boldsymbol{\theta}) = 1 + \sum_{i=1}^I \theta_i \pmod N$, and

$$u(a, \boldsymbol{\theta}) = \begin{cases} 0 & \text{if } a = \sigma(\boldsymbol{\theta}), \\ -1 & \text{if } a \neq \sigma(\boldsymbol{\theta}) \end{cases}$$

An individual is indifferent over all $a \in A$ while the group sets $a_0^*(\mathbf{s}) = \sigma(\mathbf{s})$. ■

Lemma 1. *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 1. 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 signal s_{-1} and two copies of the signal s_{I+1}, then the posterior distribution satisfies, for $k = 1, \ldots, N$,

$$P(\theta_k \mid \mathbf{s}) = \lambda \pi(\theta_k) b^{(k-1)(I-2)} \left(1 - b^{k-1}\right)^2, \tag{19}$$

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

$$\frac{P(\theta_k \mid \mathbf{s})}{P(\theta_{k+1} \mid \mathbf{s})} = \left(\frac{1}{b}\right)^{I-2} \left(\frac{1-b^{k-1}}{1-b^k}\right)^2 \frac{\pi(\theta_k)}{\pi(\theta_{k+1})} < \frac{1-b^{k-1}}{1-b^k} \frac{\pi(\theta_k)}{\pi(\theta_{k+1})} = \frac{P(\theta_k \mid s_{I+1})}{P(\theta_{k+1} \mid s_{I+1})}. \tag{20}$$

To establish the claim, observe that the inequality in expression (20) is equivalent to

$$b^{I-2} \frac{1-b^k}{1-b^{k-1}} > 1. \tag{21}$$

Since the left-hand side of inequality (21) converges to $k/(k-1)$ as b approaches one, the claim follows. It follows from inequality (20) that the profile of signals $\mathbf{s}$ leads to a posterior that dominates the individual signal s_{I+1}. ■

Proof of Proposition 8. Add additional signals to the technology constructed in Lemma 1 that are mixtures of the signals s_{-1} and s_{I+1}. Specifically, for $j = 0, \ldots, I$ define s_i so that $s_{-1} < s_0 < \cdots < s_{I+1}$ and

$$\alpha^*(s_j \mid \theta) = \lambda(\theta) \left(c_j \alpha(s_{-1} \mid \theta) + (1 - c_j)\alpha(s_{I+1} \mid \theta)\right) \text{ for } j = -1, \ldots, I+1, \tag{22}$$

where $\sum_{j=-1}^{I+1} \alpha^*(s_j \mid \theta) = 1$ for all θ, $0 \le c_I < \cdots < c_0$, 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 Lemma 1 and continuity.

First, I show that given a profile of actions $(a_1, \ldots, a_I)$ ordered so that $a_1 \le a_2 \le \cdots \le a_I$, 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 $j = 0, \ldots, I+1$,

$$\sum_{\theta} \frac{\alpha^*(s_j \mid \theta)\pi(\theta)}{\sum_{\omega} \alpha^*(s_j \mid \omega)\pi(\omega)} f(\theta) = a_j. \tag{23}$$

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

$$A_i = \sum_\theta \alpha^*(s_i \mid \theta)\pi(\theta)f(\theta) \tag{24}$$

and

$$B_i = \sum_\theta \alpha^*(s_i \mid \theta)\pi(\theta). \tag{25}$$

Using (22), (23) can be written:

$$c_j A_{-1} + (1 - c_j)A_{I+1} = a_j \left(c_j B_{-1} + (1 - c_j)B_{I+1}\right). \tag{26}$$

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

$$c_j = \frac{A_{I+1} - a_j B_{I+1}}{A_{I+1} - A_{-1} - a_j(B_{I+1} - B_{-1})}. \tag{27}$$

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 (27) are non-negative and decreasing. This establishes the claim.

When $u(a_i, \theta) = -(a_i - f(\theta))^2$, (23) 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 used a utility function of 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 do so for $u(a_i, \theta) = (a_i\theta + (1 - a_i)\theta_0)^\beta$ for appropriate choices of $\theta_0 > 0$ and $\beta < 1$. Specifically, let

$$\theta_0 = \sum_\theta \frac{\alpha^*(s_0 \mid \theta)\pi(\theta)\theta}{\sum_\omega \alpha^*(s_0 \mid \omega)\pi(\omega)}. \tag{28}$$

The definition of θ_0 in (28) 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

$$\sum_\theta \frac{\alpha^*(s_{I+1} \mid \theta)\pi(\theta)\theta}{\sum_\omega \alpha^*(s_{I+1} \mid \omega)\pi(\omega)} \theta^{\beta-1}(\theta - \theta_0) = 0. \tag{29}$$

If $u(a_i, \theta) = (a_i\theta + (1 - a_i)\theta_0)^\beta$, then equation (28) guarantees that the best response to s_0 is the action 0 and (29) 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 9. Suppose that $\mathbf{a}_0 = (a_1, \ldots, a_I)$, with $0 \leq a_1 \leq \cdots \leq a_I$ and $a_{I-1} > 0$. Proposition 8 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 8 so that $a_i^*(s_i) = a_i$ and $a_0^*(\mathbf{s}) = 1$ (which, provided $a_{I-1} > 0$, is still possible, since the posterior given $\mathbf{s}$ will dominate the posterior given the signal associated with a_I). ■

Proof of Proposition 10. 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 7 that $a_0^*(s_1, \ldots, s_I) \in [a_1, a_I]$. ■

References

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

Athey, S., and J. Levin (2001): "The Value of Information in Monotone Decision Problems," Discussion paper, Stanford University.

Bates, J. M., and C. W. J. Granger (1969): "The Combination of Forecasts," *Operational Research Quarterly*, 20(4), 451–468.

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

Börgers, T., A. Hernando-Veciana, and D. Krähmer (2011): "When Are Signals Complements or Substitutes?," Discussion paper, University of Michigan.

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

Chambers, C. P., and P. J. Healy (2010): "Updating Toward the Signal," *Economic Theory*, pp. 1–22.

Clemen, R. T., and R. L. Winkler (1990): "Unanimity and Compromise Among Probability Forecasters," *Management Science*, 36(7), 767–779.

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

Dixit, A. K., and J. W. Weibull (2007): "Political polarization," *Proceedings of the National Academy of Sciences*, 104(18), 7351.

Eliaz, K., D. Ray, and R. Razin (2006): "Choice Shifts in Groups: A Decision-Theoretic Basis," *American Economic Review*, 96(4), 1321–1332.

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

Janis, I. L. (1983): *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Houghton Mifflin Company, Boston.

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

Roux, N., and J. Sobel (2012): "Group Polarization in a Model of Information Aggregation," Discussion paper, UCSD.

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.

Winkler, R. L. (1986): "Expert Resolution," *Management Science*, 32(3), 298–303.