Group Polarization in a Model of Information Aggregation*

Nicolas Roux[†] and Joel Sobel[‡]

September 17, 2013

Abstract

Experiments identify the empirical regularity that groups tend to make decisions that are more extreme, but in the same direction as the tendency of individual members of the group. We present a model of information aggregation consistent with these findings. We assume individuals and groups are rational decision makers facing monotone statistical decision problems where groups and individuals have common preferences but groups have superior information. We provide conditions under which the distribution of the optimal actions of the group is more variable than the distribution of actions taken by individuals. *Journal of Economic Literature* Classification Numbers: A12, D01; Keywords: statistical decision problem; group polarization; behavioral economics; psychology.

*We acknowledge useful comments from many seminar audiences and from Sophie Bade, Doug Bernheim, Richard Brady, Antonio Cabrales, Juan Carrillo, Vince Crawford, Todd Davies, Pierre Fleckinger, Thibault Gajdos, Juan José Ganuza, Rick Harbaugh, Navin Kartik, David Laibson, Mark Machina, José Penalva, Debraj Ray, David Schkade, Ricardo Serrano, Ross Starr, and Joel Watson. Sobel is 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. Sobel 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 and he thanks these institutions for their support.

[†]Max Planck Institute for Research on Collective Goods

[‡]University of California, San Diego

1 Introduction

Groups make decisions that are more extreme than some central tendency of the individual positions of the members of the group. Stoner [25], first observed this phenomenon.[1] Other researchers have replicated and refined Stoner's insight, which the literature calls group polarization. This paper argues that a simple model of rational decision making can organize the experimental results on group polarization. The psychological literature implicitly attributes the existence of polarization as evidence of a failure of rationality. Our results suggest that this conclusion is premature.

As an example of the phenomenon, consider the experiments on group decision making performed by Schkade, Sunstein, and Kahneman [22]. Individual subjects received information (written documents and a videotape presentation) relevant to a series of hypothetical court cases. Individually, they recorded a punitive verdict on a nine-point scale and a damage verdict (a non-negative number) for each case. Subjects then were randomly assigned to groups of six; these groups deliberated and decided on punitive and damage verdicts. In those cases where individual ratings are severe, the group rating tends to be more severe than the median individual rating. In those cases where individual ratings are lenient, the group ratings tend to be more lenient than the median individual rating. As a result group ratings are more polarized across cases than individual ratings.[2]

We study a statistical decision problem where individuals have common preferences but different information. 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. We concentrate on monotone decision problems, in which 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. Decision makers (groups and individuals) select actions to maximize expected utility given their information. We assume that individuals have common preferences and that groups perfectly aggregate information available to members of the group. Consequently,

[1] Brown [5] provided a extensive review of the psychological literature through the mid 1980's.

[2] In many experiments, subjects are asked to supply individual recommendations after deliberating with a group. Typically individuals shift their individual recommendation in the direction of the group's decision.

groups have superior information to individuals and therefore make better decisions. We provide conditions under which this difference in information causes the group's actions to be more extreme than the optimal actions of individuals.[3]

To study polarization formally, it is useful to suppress the distinction between individual and group and simply compare the distribution of decisions as a function of the quality of the information available. Abstract results that state that decisions are more extreme when the information is more precise then imply that group decisions are more extreme than individual decisions (since groups have more precise information). Consequently for much of the paper we study decision rules as a function of the quality of information.

We introduce the basic ideas in Section 2, which describes several examples. In the examples, it is possible to compute explicit formulas for the decision rule as a function of the precision of the information. The examples introduce the distinction between ex ante and ex post variation. The ex ante distribution of actions is the distribution induced without conditioning on the realized state of nature. The distribution of actions becomes more variable[4] when the quality of information improves: The decision maker is more able to adjust decisions to the state of the world as her information improves. We also examine the distribution of actions conditional on the true state. One might expect better information to lead to less variable decisions. For example, a decision maker with perfect information will always make the right decision. On the other hand, a completely uninformed decision maker's action rule will also be constant. For the examples in Section 2, variability of the ex post distribution is first increasing and then decreasing in the available

[3]Our model deviates from the typical experimental design in an important way. In experiments, agents receive standardized information. In our model, agents receive different information. If agents have identical preferences and receive identical information, then there would be no variation in individual recommendations and the group's decision would agree with the (common) individual recommendation. It is possible to attribute the variations observed in experiments to mistakes or to heterogeneous preferences. To support our modeling assumptions, we attribute heterogeneous behavior to different information processing capabilities of different individuals. Our model is appropriate if agents have different information processing capability (so that different agents pay attention to different aspects of identical signals) or different prior information. We do not model these differences explicitly, but believe that they are a sufficiently plausible source of heterogeneity to justify our approach.

[4]We provide several ways to formulate variability in the paper. If the reader requires a precise definition now, identify variability with the variance of the action rule.

information.

Having illustrated our results, we introduce the formal model in Section 3. Section 4 presents some general results. If actions respond linearly to changes in beliefs, then the group actions are ex ante more variable than individual actions in the sense of second-order stochastic dominance. The linearity property is a joint restriction on preferences and information. It holds in the examples of Section 2.2. With weaker restrictions on preferences, we derive weaker results on variability.

Section 4.2 demonstrates two ways in which the distribution of actions conditional on the true state becomes more extreme as information improves. First, when there are two states of the world, we provide conditions under which improvements in information increase the probability of low actions when the state of the world is low and of high actions when the state of the world is high. Second, we study problems in which preferences satisfy a uniform single-crossing condition. This condition guarantees that a perfectly informed agent would (almost always) want to take an extreme action. We show that as the decision maker's information approaches perfect information, the probability that the decision maker takes an extreme decision approaches one.

Section 5 demonstrates how the model is compatible with experimental results. The results on ex post variability in Section 4.2 show how the classical results on group polarization and the Schkade, Sunstein, and Kahneman [22] results on jury deliberations involving guilt or innocence are consistent with our model. In particular, when it is optimal for a fully informed decision maker to take an extreme action, it is natural to expect better information to lead to more extreme behavior.

Section 5 also contains a discussion of the punitive awards task studied by Schkade, Sunstein, and Kahneman. This task differs somewhat from the standard example in the group-polarization literature because it placed no upper bound on the action set. Group decisions exhibit two unusual properties. In several groups, the group punitive awards is much higher than the maximum award recommended by an individual.[5] Second, the median of the individual punitive award decisions is more predictable than the group decision. Informally, the decisions obtained through group deliberation are more variable than those obtained through a majority vote. Our analysis identifies circumstances in which these properties would arise. In particular,

[5] This result was also observed in other experiments.

the group's decision will be more variable ex post than the individual decisions when individuals make their decision based on poor information. This conclusion follows from the non-monotonicity of ex post variability that we illustrate in Section 2.

Section 6 discusses related literature and, in particular, compares our approach to prominent theories from social psychology. Section 7 is a conclusion.

2 Examples

This section illustrates the main theoretical results using a series of examples. All examples feature a decision maker who attempts to estimate an unknown number (the state of nature, θ). She holds a prior over the state. Before reporting her estimate (her action, a), the decision maker receives a signal informative for the state based on which she updates the prior. In all examples losses are quadratic, i.e. the utility function is $u(a, \theta) = -(a - \theta)^2$. As a result she reports her posterior expectation over the state. The examples make different assumptions about the agent's information structure.

Before the state of nature is drawn, one should expect the distribution of actions to be more variable, properly defined, as the information of the decision maker improves. A poorly informed decision maker is only mildly influenced by her signal. As a result, whatever the state of nature, her action stays close to the optimal action based on the prior belief. As her information's quality improves, the decision maker's actions become more responsive to her signals, which are themselves more responsive to variations of the state of nature. So one expects the better informed group to take actions that are more variable ex ante.

The distribution of actions ought to be closer (in some sense) to the ex post optimal decision the better is the information of the decision maker. This intuition suggests that ex post, the distribution of actions will place more weight on extreme actions when extreme actions are optimal. We investigate this intuition in the general analysis of Section 4.2. The relation between information precision and the ex post variability of actions is typically non-monotonic. To see this note that both a perfectly informed decision maker and a decision maker receiving uninformative signals exhibit no variability in their actions. The perfectly informed individual always knows the state of nature and therefore (assuming that there is a unique optimal response given the state) always plays the same action. The decision maker with an uninformative information structure always takes the ex ante optimal action. These observations suggest that whether an improvement in information induces an increase or decrease in the actions' variability depends on the relative weight of two opposing effects: the increased responsiveness of actions to signals on the one hand and the decreased noisiness of signals on the other.

In this section, we study how the properties of distributions of optimal actions vary with the quality of the decision maker's information. The examples all assume that there is a single decision maker and a one-parameter family of information structures. The information structures vary in their precision. When we apply our theory, we treat groups as being better informed than individuals. So we associate a higher-precision information structure with a group and a lower-precision information structure with an individual.

The examples in this section all illustrate ways one might expect the distribution of actions to become more variable as the information structure improves. Formal statements of these results require precise definitions of variability of distributions and improvement of information. We provide definitions in Section 3. Here our presentation is less formal.

2.1 A Binary Example

Example 1 There are two states of nature $\Theta = \{0, 1\}$, two signals $S = \{0, 1\}$ and agents have to choose an action in the unit interval, $A = [0, 1]$. The prior probability distribution puts equal weights on the two states, $\pi(0) = \pi(1) = \frac{1}{2}$. The conditional distributions over signals are $\alpha(1 \mid 1) = \alpha(0 \mid 0) = c$, $c \in [1/2, 1]$. It follows that the decision rule $a^*(\cdot)$ satisfies $a^*(0) = 1 - c$ and $a^*(1) = c$. Ex ante, each action is taken with probability 1/2. The ex ante variance of $a^*(\cdot)$ is $(2c^2 + 2(1 - c)^2 - 1)/4$, which is strictly increasing in c: An improvement of the informativeness of the signal produces an increase in the variance of the induced optimal decision. When c increases, the optimal responses to each signal become more extreme because the decision maker is more confident. Conditional on a particular realization of θ, the actions will continue to be more extreme the better the information, but, given θ, the signal $s = \theta$ will increase in likelihood and this increase is relatively great the better is the decision maker's information. This effect tends to reduce the variance of the ex post decision rule as α increases. A computation demonstrates that the variance of the optimal action given

$\theta = 1$ is $(1-c)(1-2c)^2$, which is increasing over $[1/2, \hat{p}]$ and decreasing over $[\hat{p}, 1]$ where $\hat{p} \in (1/2, 1)$.

2.2 Examples using Exponential Families

In this subsection, we assume that an exponential family of distributions describes the information structure. For this family, the decision rule is a linear function of the signals. This property provides a natural way to compare decisions as the quality of information improves.[6]

Example 2 Suppose that $\theta \in \mathbb{R}$, $\pi(\cdot)$ is normal with mean μ and precision[7] $\tau > 0$, and that given θ, s is a normal distribution with mean θ and precision $r > 0$. The posterior distribution of θ given I independent signals $\mathbf{s} = (s_1, \ldots, s_I)$[8] is a normal distribution with mean $\mu^*(\mathbf{s})$ and precision $\tau + Ir$, where

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

In this example, the posterior distribution depends on the average signal. The optimal recommendation is simply the conditional mean of θ, $a^*(\mathbf{s}) = \mu^*(\mathbf{s})$. For this example, it is natural to treat r as a parameter that measures the precision of the information. With this interpretation, the optimal action is given by $a^*(s) = (\tau\mu + rs) / (\tau + r)$. (It follows from (1), that going from an individual to a group with I members who each receive independent, identically distributed signals, increases the precision by a factor of I.)

The ex ante variance of the action distribution is $r/[\tau(\tau + r)]$, which increases in r while the ex post variance (conditioned on the realization of θ) is $r/(\tau + r)^2$, which first increases and then decreases in r. As in the first example, the variability of the ex ante distribution of actions (as measured by variance) is greater the higher is r, while the variability of the ex post distribution is concave. Low r leads to low variability because the action is not sensitive to any information; high r leads to low variability because the action is perfectly suited to the state.

[6] Kaas, and Dannenburg, and Goovaerts [14] describe the linearity property used in our analysis. Darmois [8], Koopman [15], and Pitman [19] discovered the classic relationship between exponential families and sufficient statistics.

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

[8] We can imagine that the I signals come from I different agents. With this interpretation, I is the size of the group.

Example 3 Take $A = [0,1]$, $\Theta = [0,1]$ and assume that the prior is a Beta distribution with parameters rt and $r(1-t)$, where t is the expected value of θ and r is the precision. Let $S = \{0,1\}$ and let the conditional distribution of signals satisfy: $\alpha(1 \mid \theta) = \theta$ and $\alpha(0 \mid \theta) = 1 - \theta$. In this example, if each member of a group of size I receives an independent and identically distributed signal and $\mathbf{s} = (s_1, \ldots, s_I)$ and $\bar{\mathbf{s}} = \sum_{i=1}^{I} s_i/I$ then the action rule is

$$a^*(\bar{\mathbf{s}}) = \frac{r}{r+I}t + \frac{I}{r+I}\bar{\mathbf{s}}.$$

The same qualitative properties of Example 2 hold. As the group size increases, the action shifts away from the prior mean and towards the average signal. The expected action does not depend on the group size. The ex ante variance of the action is $I[r^2 t(1-t)]/[(I+r)(1+r)]$, which is increasing in I. The ex post variance of the action is $[I^2\theta(1-\theta)]/(r+I)^2$, which is increasing when $I > r$ and decreasing when $I < r$.

Example 4 Take $A = [0,1]$, $\Theta = [0,1]$ and assume that the prior is a Gamma distribution parameters ρ and β. Let $S = \mathbb{N}$ and let the conditional distribution of signals be Poisson with parameter θ. In this example, if each member of a group of size I receives an independent and identically distributed signal and $\mathbf{s} = (s_1, \ldots, s_I)$ and $\bar{\mathbf{s}} = \sum_{i=1}^{I} s_i/I$ the action rule is

$$a^*(\bar{\mathbf{s}}) = \frac{\rho}{\beta+I} + \frac{I\bar{\mathbf{s}}}{\beta+I},$$

then the same qualitative properties of Example 3 hold. As the group size increases, the action shifts away from the prior mean and towards the average signal.

The posterior distribution is Gamma with parameters $\rho+I$ and $\beta+\bar{\mathbf{s}}$. The expected action is $(\rho + I\bar{\mathbf{s}})/(\beta + I)$ does not depend on the group size. The ex ante variance of the action is $[I^2\rho(\beta+1)]/[(\beta+I)\beta]^2$, which is increasing in I. The ex post variance of the action is $I\theta/(I+\beta)^2$, which has an interior maximum.

Section 3 presents a general framework that admits these examples as special cases.

3 The Framework

We model information aggregation as a monotone statistical decision problem. Decision makers recommend an action $a \in [\underline{a}, \overline{a}]$. Their choice depends upon an underlying state of the world θ that is drawn from an ordered set Θ according to a prior distribution $\Pi(\cdot)$ ($\pi(\cdot)$ denotes the corresponding density or probability mass function). Decision makers receive a signal s informative for the state of nature that is drawn from the set $S = [\underline{s}, \overline{s}]$. A joint distribution $\mathcal{P}$ defined on $\Theta \times S$ describes the information structure. $\mathcal{A}(\cdot \mid \theta)$ denotes the conditional distribution of signals given that the state is θ ($\alpha(\cdot \mid \theta)$ is the corresponding density or probability mass function).

The ex ante distribution over signals, $\mathcal{D}(\cdot)$, is given by

$$\mathcal{D}(s) = \int_{\underline{s}}^{s} \int_{\theta \in \Theta} \alpha(s \mid \theta)\pi(\theta) d\theta.$$

A decision maker who receives signal s updates her prior belief according to Bayes's Rule and obtains a posterior distribution denoted $P(\cdot \mid s)$ (or $P(s)$):

$$P(\theta \mid s) = \frac{\alpha(s \mid \theta)\pi(\theta)}{\int_{\omega \in \Theta} \alpha(s \mid \omega) d\Pi(\omega)}. \tag{2}$$

The state space Θ, the signal space S, and the joint probability distribution $\mathcal{P}$ determine the information available to the decision maker. Since we hold the state space and the prior fixed, we refer to $\mathcal{I} = (S, \{\mathcal{D}, \{P(\cdot \mid s)\}_{s \in S})$ as the information structure of the decision maker. The information structure is **perfect** if for all (s, θ), $P(\theta \mid s) > 0$ implies $P(\theta' \mid s) = 0$ for all $\theta' \neq \theta$.

After the decision maker observes $s \in S$ and updates her prior, she chooses the action that maximizes her expected utility given the resulting posterior belief:

$$a^*(s) \in \text{argmax}_{a \in A} \int_{\theta \in \Theta} u(a, \theta) dP(\theta | s),$$

where $a^*(\cdot)$ is referred to as the **action rule** of the decision maker.[9]

We study **monotone** environments in which the action rule is an increasing function of signals. Athey and Levin [3] derive monotonicity of the action

[9] We assume that the action rule is single-valued.

rule from joint assumptions on the utility function and the information structure. A class of utility functions is defined by properties of the incremental return function, $r(\theta) = u(a', \theta) - u(a, \theta)$, for $a' > a$. The more restrictive the assumptions on $r(\cdot)$, the larger the class of information structures that will induce a monotone structure. We pay special attention to information structures ordered by the **monotone likelihood ratio** (MLR) property:[10]

$$\text{for all } \theta' > \theta \text{ and } s' > s, \ \frac{\alpha(s' \mid \theta')}{\alpha(s \mid \theta')} \geq \frac{\alpha(s' \mid \theta)}{\alpha(s \mid \theta)}. \tag{3}$$

If Inequality (3) holds, then we write $P(\cdot \mid s') \succeq_{MLR} P(\cdot \mid s)$.[11] We write $P(\cdot \mid s') \succ_{MLR} P(\cdot \mid s)$ if the inequality between likelihood ratios in (3) is strict. If the utility function satisfies the **single-crossing property** ($r(\theta) \geq 0$ implies $r(\theta') \geq 0$), then the action rule is increasing if posteriors are ordered by the monotone likelihood ratio order.[12] If the utility function can be further assumed to be **supermodular** ($r(\cdot)$ increasing), then posteriors need only be increasing in the sense of first-order stochastic dominance. When incremental returns are concave in θ, actions are increasing with respect to second-order stochastic dominance shifts in beliefs. The examples in the previous section all use monotone environments.

In this framework, we attempt to establish a relation between the precision of an information structure and the distribution of induced actions. We consider two monotone information structures, $\mathcal{I}_G$ and $\mathcal{I}_I$, and we assume that the former is more precise than the latter. Athey and Levin [3] introduce a natural way of modeling information precision in monotone environments. The **monotone information order** (MIO) relates the precision of an information structure to the ex ante expected utility it provides. Given a class of utility functions, $\mathcal{I}_G$ is more **precise** than $\mathcal{I}_I$ if it provides a greater ex ante expected utility for any utility function in this class.[13]

This framework generalizes the examples of the previous section. In Examples 2-4, adding independent, identically distributed signals increased the

[10]Our results do not depend on the specific class of monotone problems considered.

[11]Equation (2) describes the relationship between $P(\cdot)$ and $\alpha(\cdot)$.

[12]Quah and Strulovici [20] show that actions are increasing as a function of changes of belief in monotone likelihood ratio for a larger class of utility functions.

[13]The most famous criterion for information precision is due to Blackwell [4]. Blackwell's criterion is applicable to any kind of statistical decision problem, that is $\mathcal{I}_G$ Blackwell dominates $\mathcal{I}_I$ if and only if a decision maker with arbitrary preferences prefers holding $\mathcal{I}_G$ to $\mathcal{I}_I$. Given the analysis is restricted to ordered information structures, Blackwell's criterion is more restrictive than the MIO.

precision of the information. Hence we can view associate a less precise information structure with an individual and a more precise information structure with a group that consists of individuals who perfectly aggregate iid signals. In this section, the notion of information precision is weaker. The more precise (group) information structure need not be derived from an individual information structure. For the examples of Section 2.2 the information structure obtained by aggregating individual signals is monotone. This is because the mean of individual signals is a sufficient statistic for the group's information. In general, an information structure derived by aggregating iid signals need not be monotone.

It is restrictive to assume that the underlying decision problem has a monotone structure. These restrictions permit us to use powerful techniques. The assumptions appear in many common economic applications and seem natural in our applications. For example, in the case of experiments on jury decision making, θ may be an index of culpability of the defendant and we require that all jurors prefer to impose higher punishments on when evidence indicates that the defendant is more culpable. Furthermore, there is a theoretical reason for imposing a monotone structure. Sobel [24] demonstrates that without assumptions on the information structure there will be no general relationship between group and individual decisions. Hence some restrictions are needed in order to describe conditions under which groups take systematically more extreme decisions than individuals.

4 Information Precision and Polarization

We consider two decision makers with common preferences but distinct information structures. One decision maker (the group) holds a more precise information than the other (the individual). We investigate to what extent the group makes decisions that are more extreme than the individual's decisions. The model yields results in terms of comparison of the decision makers' distributions of actions. We distinguish comparisons from ex ante and ex post points of views. The examples introduced in the Section 2 illustrated these properties. We now make the intuition more precise and the analysis more general.

4.1 Ex Ante Analysis

In this section, we present general results that demonstrate that increasing the quality of the information structure increases the ex ante variability of actions.

Consider two monotone information structures $\mathcal{I}_i = (S_i, \mathcal{D}_i, \{P_i(\cdot \mid s)\}_{s\in S_i})$ where $S_i = [\underline{s}_i, \overline{s}_i]$ is a nontrivial interval and $\mathcal{D}_i$ is strictly increasing for $i = I, G$.[14] We assume that $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$.[15]

An information structure can be described as a set of signals, a family of posterior beliefs induced by those signals, and an ex ante distribution on signals. An information structure therefore induces a random variable taking values in the set of posteriors $\Delta(\Theta)$. Athey and Levin [3] characterized the MIO-SC order as a measure of variability of this set of random variables:

Fact 1 *[Athey and Levin [3]] $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if for all $z \in [0,1]$, $P_I(\cdot \mid \mathcal{D}_I \leq z) \succ_{MLR} P_G(\cdot \mid \mathcal{D}_G \leq z)$.*

Given that both the group and the individual receive a low signal (a signal in the bottom fraction z of signals), the group is more confident that the state of nature is low since its information if more precise. Consequently, the group posterior is lower than the individual posterior with respect to the MLR order. It follows that the group chooses a lower action than the individual. An equivalent characterization can be obtained by considering good information. This leads to the alternative formulation that $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if $P_I(\cdot \mid \mathcal{D}_I > z) \prec_{MLR} P_G(\cdot \mid \mathcal{D}_G > z)$ for all $z \in [0,1]$. In this case, the group is more confident that the state of nature is high. Therefore, the observation that group information structures will dominate individual information according to the MIO-SC order provides a sense in which group beliefs are more dispersed than individual beliefs.

A monotone information structure $\mathcal{I} = (S, \mathcal{D}(\cdot), \{P(\cdot \mid s)\}_{s\in S})$ and utility function u give rise to an action rule $a^*(\cdot)$ that is increasing in the signal s and a distribution of actions, Λ, in which $\Lambda(a) = \text{Prob}(\{s : a^*(s) \leq a\})$.

[14] This assumption is without loss of generality as for any information structure one can find an informationally equivalent information structure with continuous signal set and strictly increasing ex ante distribution (see Lehmann [16]).

[15] Since the Blackwell criterion is more restrictive than the MIO-SC criterion, all the results derived in this section therefore hold for Blackwell [4]'s notion of informativeness. The counterexamples also hold for the Blackwell criterion when they use a binary state space because in that case the Blackwell and MIO-SC orderings are equivalent.

Defining $s^*(a)$ by $s^*(a) \equiv \sup\{s : a^*(s) \leq a\}$ if $\{s : a^*(s) \leq a\}$ contains at least one element and $s^*(a) \equiv \underline{s}$ if $\{s : a^*(s) \leq a\}$ is empty, we obtain $\Lambda(a) = \mathcal{D}(s^*(a))$. The value $a^*(\mathcal{D}^{-1}(\cdot))$ is the quantile function associated to Λ. Holding the utility function fixed, we seek conditions under which the induced distribution of actions from the group information Λ_G is more variable than the distribution of actions of the individual, Λ_I.

A natural way to formalize the notion of "more variable" is second-order stochastic dominance. Λ_I **second-order stochastically dominates** Λ_G if

$$\int_{-\infty}^{\tilde{a}} \Lambda_G(a)da \geq \int_{-\infty}^{\tilde{a}} \Lambda_I(a)da \text{ for all } \tilde{a}. \tag{4}$$

If Inequality (4) holds, then we write $\Lambda_I \succeq_{icv} \Lambda_G$ because the condition is equivalent to the property that $\int \phi(x)d\Lambda_I \geq \int \phi(x)d\Lambda_G$ for all increasing concave functions ϕ (see, Shaked and Shanthikumar [23, Chapter 4]). If Λ_I and Λ_G have the same mean, then the inequality in (4) will hold as an equation when $\tilde{a} = \overline{a}$. Further, it is well known that if Λ_I and Λ_G have the same mean, then Λ_I second-order stochastically dominates Λ_G if and only if Λ_G can be obtained from Λ_I through a sequence of mean-preserving spreads. We say that the group's action distribution is more variable than the individual's if (4) holds.

Proposition 1 presents conditions under which the distribution of the group's actions is more dispersed according to SSD than the distribution of the individual's actions. Given preferences $u(a, \theta)$, the **action function** δ associates to each posterior P in $\Delta(\Theta)$ an optimal decision $a = \delta(P)$, where

$$\delta(P) = \text{argmax} \int_{\theta \in \Theta} u(a, \theta)dP(\theta).$$

The action function is defined over the beliefs, rather than a linear space. It does not depend on the information structure, so the group and the individual share the same δ. Since the set of posteriors induced by a monotone information structure is ordered by MLR and preferences are single-crossing, the function $\delta(\cdot)$ is non-decreasing over $\{P(s)\}_{s \in S}$. We say that $\delta(\cdot)$ is strictly increasing if $P \succ_{MLR} P'$ implies $\delta(P) > \delta(P')$. We abuse terminology by calling $\delta(\cdot)$ linear for a given information technology if $\delta(\gamma P(s) + (1-\gamma)P(s')) = \gamma\delta(P(s)) + (1-\gamma)\delta(P(s'))$ for $\gamma \in (0, 1)$ and any signals s and s'.

Proposition 1 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ and the action functions are linear, then the group's action distribution is more variable than the individual's action distribution.*

The Appendix contains a proof of Proposition 1 and all subsequent results that require proof.

Proposition 1 follows from a change-of-variables argument. When the action function is linear, variability of the posteriors induced by the information structure translates directly into variability of the distribution of actions. Since group posteriors are more variable ex ante than the individual posteriors, so will be the associated distributions of actions.

Since $\delta(P_i(s)) = a_i^*(s)$ linearity of the action function is not sufficient for the linearity of the action rule. The additional requirement for obtaining a linear action rule is that posterior beliefs be a linear function of the signals, i.e. $P(\gamma s + (1-\gamma)s')) = \gamma P(s) + (1-\gamma)P(s')$. Quadratic preferences, $u(a,\theta) = -(a-\theta)^2$, generate linear action function since the optimal action is the expected value of θ according to the posterior belief of the decision maker. The property that posterior beliefs be a linear function of signals is restrictive, but holds in Examples 2-4.

If the action function is linear, then the distribution of actions generated by the group information will have the same mean as the distribution generated by individual information. If the action function is not linear, then there is no reason to expect the mean of the two distributions of actions to be equal. A possible conjecture is that the mean-adjusted distributions could be ranked by second-order stochastic dominance (that is, the distributions could be ranked by the *dilation* order in which $X' \succeq_{dil} X$ if $X' - EX'$ second-order stochastically dominates $X - EX$ (Shaked and Shanthikumar [23, page 200])). However, the following example shows that this is not true.

Example 5 Assume $\Theta = \{\theta_0, \theta_1\}$. When Θ contains two elements, posteriors can be represented by $q \in [0,1]$, the probability placed on θ_1 and an information structure $\mathcal{I}$ can be identified with a distribution of belief on $[0,1]$ that we denote by Γ. Specifically, $\Gamma(p) \equiv \text{Prob}(\{s : P(s) \leq p\})$. Let $\pi \in (0,1)$ be the prior belief on θ. The expected belief $\int_0^1 p d\Gamma(p)$ is necessarily equal to the prior belief π. It follows that $\mathcal{I}_G$ dominates $\mathcal{I}_I$ with respect to the MIO-SC order if and only if $\Gamma_I \succeq_{icv} \Gamma_G$.

Let

$$u(a,\theta) = \begin{cases} -(a)^2 & \text{if } \theta = \theta_0, \\ -\lambda(a-1)^2 & \text{if } \theta = \theta_1 \end{cases} \tag{5}$$

for $\lambda > 0$. In this case, $\delta(q) = \lambda q / (\lambda q + (1-q))$. It follows that $\delta(\cdot)$ is convex, linear, or concave depending on whether λ is less than, equal to, or greater than 1. For these preferences, λ measures the cost of a mistake in State $\theta = 1$ relative to the cost in State $\theta = 0$. When λ is close to zero, the decision will be biased towards $a = 0$ (in particular, the mean action will be less than the expected state) while as λ approaches ∞ the action will tend to be close to one.

Take a distribution of individual posteriors Γ_I and apply a single mean preserving spread to obtain Γ_G, so that the group information structure dominates the individual information structure with respect to the MIO-SC order. Suppose further that the MPS is such that the two distributions remain equal on some range $[0, \tilde{q}]$. The distributions of actions Λ_I and Λ_G are obtained though the change of variable $a = \delta(q)$, that is, $\Lambda_i(a) = \Gamma_i(\delta^{-1}(a))$ for $i = I, G$. Since $\Gamma_I(q) = \Gamma_G(q)$ for $q \in [0, \tilde{q}]$, $\Lambda_I(a) = \Lambda_G(a)$ for all $a \in [0, \delta^{-1}(\tilde{q})]$. Suppose now that $\lambda < 1$ so that δ is concave. It follows that the expected action of the group is strictly lower than the expected action of the individual: $\int_0^1 \delta(q) d\Gamma_G(q) < \int_0^1 \delta(q) d\Gamma_I(q)$. Therefore, the mean-adjusted distributions $\Lambda_i^c(a) = \Lambda_i(a + \int a d\Lambda_i)$ for $i = I, G$ satisfy $\Lambda_I^c(a) > \Lambda_G^c(a)$ for all $a \leq \delta^{-1}(\tilde{q}) - \int a d\Lambda_I$. So the distribution of the group actions is not more *dilated* than the distribution of individual actions.

The example demonstrates that without linearity, an increase in the precision of a distribution need not increase the variability of the action rule.

The next proposition describes a weak sense in which the group actions can be said to be more variable than the individual actions: the tails of the group's distribution of actions are fatter than those of the distribution of individual actions. The result replaces linearity of the action function with the assumption that the action function is strictly increasing: for $i = I, G$, $P' \succ_{MLR} P$ implies $\delta(P') > \delta(P)$ for $P, P' \in \{P_i(s)\}_{s \in S_i}$.

Proposition 2 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ and the action functions are strictly increasing, then the support of the individual's action distribution is strictly contained in the convex hull of the support of the group's action distribution.*

The proposition states that the distribution of actions generated by better information places positive probability on more extreme actions. Intuitively, the condition that $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ implies that the lowest signals induce a lower posterior (with respect to the MIO-SC order) for the group than for the individual. These lower posteriors lead to lower actions. Unlike the properties of Examples 1-4, Proposition 2 does not imply that the group's and individual's action distributions can be ranked by variance.

The following example demonstrates that the conclusions of Proposition 2 require the assumption that the action function is strictly increasing. The example has three actions. The perfectly informed group takes the intermediate action when it is appropriate, but the individual avoids the intermediate action because he never learns when the intermediate state is the most likely to occur. The group takes extreme actions less often, but with more confidence, than the individual.

Example 6 Let $\Theta = A = \{0, 1, 2\}$ with all three states equally likely ex ante. Let $u(a, \theta) = -(a - \theta)^2$. The group has perfect information (it receives signal $s = i$ when the state is $\theta = i$). The individual receives the lowest signal with probability one half: whenever the true state is $\theta = 0$ and half of the time the true state is $\theta = 1$. Otherwise the individual receives the highest signal. The group takes each action with probability one third. The individual takes the extreme actions with probability one half each. The conclusions of Proposition 2 fail to hold because the action function is not strictly increasing.[16]

There is a straightforward, but weak, conclusion that does not depend on strict monotonicity. Since $P_I(\mathcal{D}_I^{-1}(0))$ dominates $P_G(\mathcal{D}_G^{-1}(0))$ with respect to MLR, the minimum action of the group must be no higher than the minimum action of the individual. A similar result holds at the other extreme.

Proposition 3 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, then the support of the individual's action distribution is contained in the convex hull of the support of the group's action distribution.*

4.2 Ex Post Analysis

We now discuss how the distribution of actions changes with information conditional on the true state. It is still our interest to investigate the extent to which better information leads to more extreme actions.

We consider two ways in which to frame the analysis. The most straightforward situation is when the state space contains two elements. When there are only two states, there are only two possible ex post optimal actions.

[16] Similar examples can also be constructed with a binary state.

If better information leads to actions that are closer to an ex post optimal action, then one would expect more extreme actions as the information structure becomes more precise. Our results formalize this intuition. When there are more than two elements in the state space, the optimal ex post action need not be extreme. There is, however, an important special class of preferences for which the only ex post optimal actions are extreme. We demonstrate a sense in which better information leads to extreme actions in this special case.

There is a third way in which improvements of information lead to systematic changes in the action rule. The examples in Section 2.2 demonstrate situations in which the ex post distribution of actions first becomes more variable and then less variable as the precision of the decision maker improves. We have been unable to find substantive generalizations of these properties.

4.2.1 Binary State Space

We work in the case where the state space contains two elements. Lemma 1 provides conditions under which information precision translates into posteriors being higher conditional on the high state and lower conditional on the low state. If the information structure satisfies a symmetry condition (described below), then increasing precision shifts posteriors in the sense of the increasing concave order (4). Without the symmetry condition, posteriors are only more extreme in the sense that the tail of the ex post distributions of posteriors are fatter. As in Section 4.1, the properties of posteriors translate to properties about actions under appropriate conditions.

Conditional on each state θ, an information structure generates a family of posterior beliefs, $\{P(\cdot \mid s)\}_{s \in S}$. To do the ex post analysis, we assume that the distribution of signals is taken conditional on θ (so it is given by $\mathcal{A}(\cdot \mid \theta)$).

As we pointed out in Example 5, in the binary case we can describe an information structure $\mathcal{I}$ with a distribution $\Gamma(\cdot)$ on $[0,1]$. The corresponding conditional distributions are $\Gamma(p \mid \theta_i) \equiv \text{Prob}(\{s : P(s \mid \theta_i) \leq p\})$ for $j = 0, 1$.

When the state space is binary, we need a symmetry condition to guarantee that ex ante improvement in information leads to an ex post improvement for both states. An information structure on a binary state space $\Theta = \{\underline{\theta}, \overline{\theta}\}$ if $-\underline{s} = \overline{s} > 0$ (so that the domain of signals is symmetric) and $\mathcal{P}(\underline{\theta}, s) = \mathcal{P}(\overline{\theta}, -s)$. Lemma 1 states that higher dispersion of group posteriors translates to group posteriors being more extreme.

Lemma 1 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, both information structures are symmetric, and posteriors that are strictly increasing with respect to signals, then*

$$\int_0^q \Gamma_G(z \mid \theta_0)dz \geq \int_0^q \Gamma_I(z \mid \theta_0)dz \text{ for all } q \in [0,1]$$

and

$$\int_q^1 \Gamma_G(z \mid \theta_1)dz \leq \int_q^1 \Gamma_I(z \mid \theta_1)dz \text{ for all } q \in [0,1].$$

Lemma 1 states that group posteriors are more concentrated at the top of the unit interval conditional on the high state and more concentrated at the bottom of the unit interval conditional on the low state. As in the ex ante case, linearity of the action function makes it possible for these stochastic orders to rank ex post distributions of actions. Proposition 4 states this result.

Recall that in the binary case we identify the posterior distribution with the probability that the state is equal to θ_1. Consequently, the action function is a function of a single variable and we can define concavity, linearity, and convexity in the standard way. Example 5 describes a one-parameter family of preferences that exhibit convexity, linearity, and concavity.

Let $\Lambda(a \mid \theta) = \mathcal{A}(s_a^*(a) \mid \theta)$ be the conditional distribution of actions.

Proposition 4 *Assume that $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, both information structures are symmetric, and posteriors are strictly increasing with respect to signals. If the action function is concave, then*

$$\int_{\underline{a}}^a \Lambda_G(z \mid \theta_0)dz \geq \int_{\underline{a}}^a \Lambda_I(z \mid \theta_0)dz \text{ for all } a \in A,$$

while if the action function is convex, then

$$\int_a^{\overline{a}} \Lambda_G(z \mid \theta_1)dz \leq \int_a^{\overline{a}} \Lambda_I(z \mid \theta_1)dz \text{ for all } a \in A.$$

When the action function is linear, Proposition 4 implies that group actions are more extreme on average conditional on both states. If the action function is convex, group actions are guaranteed to be higher conditional on the high state, but need not be lower conditional on the low state. The intuition for this finding comes from thinking about preferences given by (5). In this case, when $\lambda < 1$, the losses for taking a high action when the state is

	0	1
l	.7	.5
m	.3	.5
h	0	0

Table 1: $\mathcal{I}_I$

	0	1
l	.6	.4
m	.4	.4
h	0	.2

Table 2: $\mathcal{I}_G$

low are large relative to the losses associated with taking a low action when the state is high. Consequently, incomplete information will bias the action function towards low actions. This tendency will create a downward bias in the individual's distribution of actions relative to the group's. On the other hand, Lemma 1 states that posterior beliefs become more extreme when information improves. These two forces work in the same direction when the state is θ_1, leading to the conclusion that as precision increases group actions are stochastically higher. When the state is θ_0, it is not clear which of the two forces dominate. Similarly, when the action function is concave, better information lead to (stochastically) lower actions in the low state.

It is not possible to order the ex post distribution of actions by first-order stochastic dominance. A decision maker with precise information is likely to take extreme actions based on the information. If the information is imperfect, the decision maker could receive an extreme signal that is unrepresentative of the true state. If the true state is high, this could lead to an action that is lower than the action a poorly informed decision maker would take. That is, a decision maker with precise information may make bigger mistakes ex post than a less informed individual.

Without the symmetry assumption, there is no guarantee that an ex ante improvement in information leads to an ex post improvement in information for all states. Hence the conclusions of Lemma 1 need not hold. Example 7 illustrates this.

Example 7 Assume that there are two states and three signals. The states are equally likely ex ante. The information structures $\mathcal{I}_I$ and $\mathcal{I}_G$ are given in Tables 1 and 2.

The group receives the completely informative signal $s_2 = m$ with positive probability. On the other hand, when $\theta = 0$, the individual's signals appear to be more informative. We show that $\mathcal{I}_G$ is more informative than $\mathcal{I}_I$, but that conditional on $\theta = \theta_0$, the group's information is not more precise than the individual's.

For the individual, the probability that $\theta = 1$ is equal to $5/12 = .5/(.7+.5)$ if the signal is zero and $5/8 = .5/(.3+.5)$ if the signal is one. It follows that

$$\Gamma_I(\cdot) = \begin{cases} 0 & \text{if} \quad p \in [0, \frac{5}{12}) \\ \frac{3}{5} & \text{if} \quad p \in [\frac{5}{12}, \frac{5}{8}) \\ 1 & \text{if} \quad p \in [\frac{5}{8}, 1] \end{cases} \tag{6}$$

and

$$\Gamma_I(\cdot \mid \theta_0) = \begin{cases} 0 & \text{if} \quad p \in [0, \frac{5}{12}) \\ \frac{7}{10} & \text{if} \quad p \in [\frac{5}{12}, \frac{5}{8}) \\ 1 & \text{if} \quad p \in [\frac{5}{8}, 1] \end{cases} \tag{7}$$

For the group, the probability that $\theta = 1$ is equal to $2/5$ if the signal is zero, $1/2$ if the signal is one, and one when the signal is two.

$$\Gamma_G(\cdot) = \begin{cases} 0 & \text{if} \quad p \in [0, \frac{2}{5}) \\ \frac{1}{2} & \text{if} \quad p \in [\frac{2}{5}, \frac{1}{2}) \\ \frac{9}{10} & \text{if} \quad p \in [\frac{1}{2}, 1] \end{cases} \tag{8}$$

and

$$\Gamma_G(\cdot \mid \theta_0) = \begin{cases} 0 & \text{if} \quad p \in [0, \frac{2}{5}) \\ \frac{3}{5} & \text{if} \quad p \in [\frac{2}{5}, \frac{1}{2}) \\ 1 & \text{if} \quad p \in [\frac{1}{2}, 1] \end{cases} \tag{9}$$

It is straightforward to check that $\int_0^q \Gamma_G(z)dz \geq \int_0^q \Gamma_I(z)dz$ for all $q \in [0, 1]$, but $\int_0^{.5} \Gamma_G(z \mid \theta_0)dz < \int_0^{.5} \Gamma_I(z \mid \theta_0)dz$.

In the example, the group's information is plainly superior to the individual's when the state is 1 because in that case the group receives a fully informative signal with positive probability. On the other hand, the individual's information is superior when $\theta = 0$. The added precision of the group information in state 1 compensates for lack of precision given 0.

When the information structures are not symmetric, we have the following weaker result:

Lemma 2 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, then there exists $\underline{q}, \overline{q} \in [0,1]$ such that $\Gamma_G(q \mid \theta_0) \geq \Gamma_I(q \mid \theta_0)$ for $q \geq \underline{q}$ with strict inequality for a set of position measure in $(0, \underline{q})$ and $\Gamma_G(q' \mid \theta_1) \leq \Gamma_I(q' \mid \theta_1)$ for $q > \overline{q}$ with strict inequality for a set of positive measure in $(\overline{q}, 1)$.*

Lemma 2 states that, conditional on the high state, the upper tail of the distribution of group beliefs is fatter than the upper tail of the individual's distribution. The corresponding statement holds conditional on the low state. Therefore, as long as actions are strictly increasing in beliefs, the same property will hold for conditional distributions of actions. This proposition does not imply that the group's actions are more extreme on average.

Proposition 5 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ and individual and group actions are strictly increasing with respect to their respective beliefs, then there exists $a', a'' \in [\underline{a}, \overline{a}]$ such that*

$$\Lambda_G(a \mid \theta_0) \geq \Lambda_I(a \mid \theta_0) \text{ for } a \leq a' \text{ and } \Lambda_G(a' \mid \theta_0) > \Lambda_I(a' \mid \theta_0)$$

and

$$\Lambda_G(a \mid \theta_1) \leq \Lambda_I(a \mid \theta_1) \text{ for } a \geq a'' \text{ and } \Lambda_G(a' \mid \theta_1) < \Lambda_I(a' \mid \theta_1).$$

If actions are not necessarily strictly increasing with respect to beliefs, then Proposition 5 does not hold. Consider an problem with two states and two actions. Compare the distribution of actions generated by a poorly informed agent who always takes the low action and a better informed agent whose decision depends nontrivially on the signal. Plainly the distribution of actions of the less informed decision maker can be lower than that of the better informed decision maker.

Information improvements with respect to the MIO weakly constrain the conditional belief distributions even when the state space contains only two elements. In the next sections, we derive results under alternative notions of information precision.

4.2.2 Single-Crossing Preferences

In this section we assume that preferences satisfy a **uniform single crossing condition**. Specifically, we assume that $A = [0, 1]$ and that there exists θ^*

such that for all $1 \geq a' > a \geq 0$, the incremental utility $r(\theta) = u(a', \theta) - u(a, \theta)$ is negative for $\theta < \theta^*$ and positive for $\theta > \theta^*$. When preferences satisfy uniform single crossing, a fully informed agent will take an extreme action (either 0 or 1) for all $\theta \neq \theta^*$. We exploit this property to demonstrate a sense in which better informed agents take more extreme actions ex post.

A representative example of preferences that satisfy the uniform single-crossing condition is a **portfolio model**. The problem is to determine the share of wealth to allocate over a safe asset, which yields θ^*, and a risky one, which yields θ. Individuals must pick the fraction a of the portfolio to invest in the risky asset. So in this model $u(a, \theta) \equiv U(a\theta + (1-a)\theta^*)$ where $U(\cdot)$ is a concave function defined over monetary outcomes. Risk averse agents will typically select $a < 1$ even when their information suggests that the mean of θ exceeds θ^*. On the other hand, if positive information is sufficiently precise, one expects to see higher investments in the risky asset.

Recall that an information structure is perfect if the distribution $\mathcal{P}$ on states and signals has the property that the conditional probability of θ given s is either zero or one. If the information structure is perfect, then with probability $1 - \theta^*$ the decision maker's posterior places probability one on a state $\theta \neq \theta^*$. Consequently, the decision maker will select an extreme action (either 0 or 1) with probability (at least) $1 - \pi(\theta^*)$. If the information structure is approximately perfect in that with high probability the posterior distribution is concentrated on the true state, then the decision maker will select an action close to 0 or 1 unless the true state is close to θ^*. Consequently when the prior distribution is atomless, one would expect large groups aggregating independent and identically distributed signals to take extreme actions with high probability.

5 Applications

In this section we discuss the relationship between our model and motivating examples.

5.1 Application to Jury Decision Making

Sunstein, Schkade and Kahnemann [22] (henceforth SSK) report on a study of jury decision making. In their experiment, subjects receive descriptions of fifteen court cases (in the form of written material and video tapes). For each case, subjects make two decisions individually: they decide on the severity of the punishment that should be given to the defendant on a nine-point rating scale going from 0 (None) to 8 (Extremely severe) and they chose the damage award that the defendant should pay (a non-negative number). Groups of six subjects then form and make the same two decisions for each case (on a consensual basis). SSK report the classic group-shift phenomenon: in those cases where the individual decisions tend to be relatively severe (lenient), the group decisions tend to be even more severe (lenient). The puzzling result in SSK's experiment concerns the comparison of the variability of group and individual actions in the two tasks. While the severe shift in punishment ratings was accompanied by a decrease in variability, the severe shift in damage awards was accompanied by an dramatic increase in variability. Lenient shift in punishment ratings also went along with a decrease in variability, but lenient shift in damage awards did not produce an increase in variability as decisions tended to cluster at 0.

SSK's findings for the punishment task are consistent with the classical experiments on polarization. They are also consistent with the theoretical implications of our two-state or uniformly single-crossing analyses. Complete information would lead to an extreme recommendation (guilt or innocence). To the extent that groups have superior information than individuals, it is natural to expect group decisions to be more extreme. We consider the results on deliberation over punitive damages in more detail.

Roux [21, Chapter 3] analyzes SSK's data and identifies five stylized facts about deliberation over punitive damages.

1. The distributions of punitive awards are skewed to the right.

2. More severe cases generate more dispersed distributions of actions.

3. Severe cases generate severe shifts whereas lenient cases induce no shift. Moreover, more severe cases generate more intensive shifts.

4. Group actions are never less dispersed than individual actions. More severe cases generate a greater difference in the dispersions of group and individual actions.

5. On rare ocassions, the maximum of the individual awards is less than the group's award.

To organize these results, consider the model of Example 4. The state and action space are $[0, \infty)$ and subjects have quadratic preferences, $u(a, \theta) = -(a - \theta)^2$. The prior is a Gamma distribution with parameters α and β. Let $S = \mathbb{N}$ and the conditional distribution of signals be Poisson with parameter θ. Each member of a group of size I receives an independent and identically distributed signal. Let $\mathbf{s} = (s_1, \ldots, s_I)$ and $\bar{\mathbf{s}} = \sum_{i=1}^{I} s_i/I$. The action function then is

$$a^*(\bar{\mathbf{s}}) = \frac{\alpha}{\beta + I} + \frac{I\bar{\mathbf{s}}}{\beta + I}. \tag{10}$$

Conditional on θ, the expected group action then is $(\alpha + I\theta)/(\beta + I)$ and its variance is $I\theta/(I + \beta)^2$. The corresponding individual variables are obtained by setting $I = 1$.

The first fact follows since signals are Poisson distributed and the decision rule is an affine transformation of signals the distribution of actions is skewed to the right.

Since expected actions are increasing in θ, we associate increasing θ with increasing severity. $I\theta/(I + \beta)^2$ is linearly increasing in θ so more severe cases indeed induce greater variance of actions, which is consistent with the second finding. As θ increases, the variation of signals increases. In this way the information structure captures the intuition that it is more difficult to evaluate the punitive damages with accuracy as the defendant's culpability increases.

The difference between group and individual expected actions conditional on θ is equal to $(I - 1)(\theta - \alpha/\beta)/\left((1 + I/\beta)(1 + 1/\beta)\right)$. So severe cases $(\theta > \alpha/\beta)$ produce severe shifts, while lenient cases $(\theta < \alpha/\beta)$ produce lenient shifts. Moreover, the intensity of the shift (defined as previously) is increasing in the absolute value of the difference between θ and α/β. Hence more severe cases generate more intensive shifts. We cannot account for the finding that lenient cases produce no shift at all, but the fact that θ is bounded below predicts that severe shifts should generate greater shifts than lenient shifts.[17]

We have pointed out that the relation between information precision and the ex post variability of actions is typically non-monotonic. In our model,

[17] A complementary explanation relates the severity shift to asymmetric losses. As explained in Section 4.2.1, if it is more costly to overestimate than to underestimate damages, then poorly informed individuals will be less likely than better informed groups to make high damage awards.

the relation between the number of signals received and the ex post variability of action is increasing over small number of signals and then decreasing. Consequently, the group's actions will be more variable ex post than the individual's actions if the individual's and group's informations structures are sufficiently poor. Specifically, the difference in the variance of group and individual actions conditional on θ is equal to

$$\frac{(1-I)(I-\beta^2)}{(1+\beta)^2(I+\beta)^2}$$

and is positive if and only if $\beta^2 \geq I$, where β^2 is inversely related to the variance of the prior belief. Intuitively, if the information available to individuals is so weak that it has little influence on decisions, individual actions may well be less variable ex post that those of better informed groups.

When $\beta^2 \geq I$, the finding that more severe cases generate a greater difference in the dispersions of group and individual actions is consistent with the model since the difference in the variance of group and individual actions conditional on θ is increasing in θ.

Finally, while the maximum of the individual awards is almost always greater than the group award, this is not true in all cases. It is clear from (10) that a group's decision may be more extreme than the recommended awards from all of the individuals in a group.

SSK argue that the judicial system should treat similar cases similarly and hence are concerned by the ex post relative unpredictability of group decisions. They interpret their results as a reason to reduce the role of juries in the assignment of punitive damages. It is also important, however, that the system treat different cases differently. Our analysis shows that the lack of variation of individual recommendations within cases may be due to a lack of variation across cases. In terms of our model, if the damage-award decision signals were sufficiently poor, the improvement of information at the level of the group could lead to an increase in variability of actions. On the contrary, if individual punishment ratings decisions were based on sufficiently informative signals, then further improvement would lead to a decrease in the variability of actions. This interpretation has the flavor of SSK's argument that subjects find it difficult to measure punitive damages in monetary units.

5.2 Choice Dilemmas

While we discussed group polarization in the context of jury decision making, the phenomenon has been observed in a wider class of problems.

Stoner [25] presented subjects with a description of a situation in which a decision maker must choose between two actions. The safe action led to a certain outcome. The alternative (risky) action led to two possible outcomes, one better and one worse than the certain outcome. The experimental subject was instructed to offer advice to the decision maker by indicating the minimum probability that the risky action would lead to the best outcome needed in order to take the risky action.

Stoner's subjects were given twelve problems. They make individual recommendations on each problem. They were then divided into groups of six and asked to make recommendations. Finally, subjects recorded their individual recommendation after the group decision had been made. A control group of individuals were not asked to participate in the group recommendation process, but were asked to make two individual recommendations. Stoner found that the group recommendation was riskier than the average individual recommendation in most cases. Individual recommendations collected after group deliberations also exhibited a shift towards more risky recommendations (although not as pronounced as the group's shift). That is, subjects changed their individual recommendation after deliberation. Individuals in the control group did not exhibit a systematic change in recommendations. Stoner's results have been replicated in many settings. They are an established part of the literature on social psychology.[18]

The finding of a risky shift was quickly refined to include a richer set of findings. While Stoner's subjects typically make more risky recommendations as groups than as individuals, in some cases, their recommendations were more cautious. Moreover, the shifts were predictable in two related ways. The direction of the shift could be predicted from the average individual response and the average individual response could be predicted from the nature of the choice dilemma. These features are fully consistent with our basic model.

[18] Brown [5] devotes 50 pages of a textbook to the topic of group polarization.

6 Related Literature

Sobel [24] discusses the relationship between individual and group decisions without restrictions on information structures. He shows the group's optimal decision is not constrained by individual recommendations. That is, in general polarization is consistent with rational decision making, but not a consequence of rational decision making. The current paper demonstrates that polarization is a natural outcome in monotone decision problems.

Eliaz, Ray, and Razin [10] present the first, and to our knowledge, only other, theoretical model of choice shifts.[19] 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 do not arise. Eliaz, Ray, and Razin [10] 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. 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 our approach, this model does not rely on information aggregation. Eliaz, Ray, and Razin [10] 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 [10] approach is the connection it makes between systematic shifts in group decisions and systematic violations of the expected utility hypothesis.

Existing evidence suggests that group polarization can only be partially

[19]There is a separate literature than examines the possibility that observing a common signal will cause the priors of different individuals to diverge. See Acemoglu, Chernozhukov, and Yildiz [1], Andreoni and Mylovanov [2], and Dixit and Weibull [9] for models of this kind. Clemen and Winkler [7] and Winkler and Clemen [27] study a related model of combining forecasts. They point out that one should not expect an aggregate of forecasts to be a weighted average of individual forecasts.

interpreted as a result of preference aggregation. Preference based approaches imply that individual decisions do not change after discussion. Yet, it is systematically the case that individual decisions collected after the discussion also polarize compared to the initial individual decisions. Group discussion appears to aggregate information to some extent because it leads to changes in individual (post deliberation) decisions.[20]

Group polarization has received a great deal of attention in the psychology literature, which has proposed several theories to explain experimental findings.[21]

According to the social-comparison theory, individuals evaluate their actions relative to a norm of behavior that is reflected in the actions of others. For a given problem, there is an ideal choice that may depend on the choices of others. For example, in some problems individuals may wish to make a recommendation that is somewhat riskier than the average recommendation.[22] Individuals make their original, pre-deliberation recommendation according to their prior perception of the ideal choice. During deliberation, the group's distribution of choices becomes known. Some individuals will discover that their original position was not at its ideal location relative to the group and shift accordingly. This theory depends on several assumptions. Individuals must have preferences that depend on the preferences (or actions) of others. There must be uncertainty about the beliefs of others so that observing their recommendations conveys relevant information about the prevailing norm of behavior. The location of the ideal position must depend on the choice problem in order to account for shifts in different directions.

The persuasive-arguments theory[23] posits that for each choice problem there are many possible arguments in favor of any recommendation. Individuals use a subset of these arguments to support their initial recommendation. During deliberation, individuals share their arguments. Collecting the whole sample of arguments for each case makes it possible to distinguish cases ac-

[20]Stoner [25] notes that groups' decisions polarize slightly more than post deliberation individual decisions.

[21]Brown [5], Isenberg [5], Myers and Lamm [18], and Turner [26] contain detailed reviews.

[22]Brown [5, page 469] describes the process as follows: "People will be motivated to fall on one or the other side of the central tendency because they seek not to be average but better than average, or virtuous. To be virtuous, in any of an indefinite number of dimensions, is to be different from the mean – in the right direction and to the right degree."

[23]See Brown [5] for a clear exposition.

cording to the number of arguments in favor of the risky alternative relative to the arguments in favor of the cautious alternative. The proportion of risky to cautious arguments in a case is hypothesize to to predict the average individual response in the case and the direction and magnitude of the group polarization. Specifically, a case with a higher (lower) proportion of risky to cautious arguments induces risky (cautious) individual responses and produces a stronger shift of group responses in the direction of the risky (cautious) alternative.

The theory does not explain where arguments come from and why individuals do not recognize that the group may have a biased sample of arguments. We propose that our model gives a precise formulation of what an argument includes, a way to predict the tendencies of individuals, and an explanation of choice shifts. Specifically, consider Example 3. Interpret signal as arguments and distinguish between arguments that favor one option ($s = 1$) or the other ($s = 0$). The state of nature $\theta \in [0, 1]$ characterizes a case. The proportion signals $s = 1$ is equal to θ. The decision rule is a convex combination of the mean of the received signals $\bar{s}$ and the initially expected mean of θ, t. The expected individual response in a case with a proportion θ of risky arguments is $(rt + k\theta) / (r + k)$, which is increasing in the proportion of $s = 1$ arguments. Moreover, the magnitude of the shift, that is, the distance between the expected group and individual responses, $r(I-1)(\theta - t)/ ((r + I)(r + 1))$, is indeed increasing in the distance between θ and t.

Persuasive argument theory asserts that the reason why group decisions shift in a particular direction is that the proportion of arguments that support this direction during deliberations is relatively high. In contrast, is simply that groups have a more exhaustive pool of arguments. In our model, there is no reason to treat some actions as riskier than others so we have describe the effect as shifts towards one or the other extreme. A natural way to incorporate the notion that some decisions are riskier than others is to assume that losses associated with incorrect decisions are greater in (what we would call) the riskier direction.

There have been many experimental attempts to separate the predictions of the two theories. By its nature, the persuasive argument theory demands that the group has an opportunity to exchange opinions and not just recommendations.[24] While the evidence is not conclusive, it appears that mere

[24]Burnstein and Vinokur [6], however, argue that exposing an individual to the recom-

exposure to the recommendations of others is sufficient to generate choice shifts. This finding is certainly consistent with our model, in which (assuming that the information structure is known and recommendations are strictly increasing in the signal received) the informational content of a recommendation is contained in the recommendation itself.

The most naive version of the social comparison theory gives no reason to believe that recommendations converge after one opportunity to deliberate. After individuals revise their personal recommendations to conform better with the group ideal, the ideal shifts. This shift could generate another change in individual recommendations. This possibility is not consistent with the persuasive argument theory or our model (assuming that deliberation permits a complete discussion of available arguments and information). We do not know whether it has been investigated experimentally.

The theories are all extremely flexible. Social comparison theories have freedom to modify the location of the ideal position from problem to problem. The persuasive argument theory permits ad hoc adjustments in the number and quality of arguments available for each position. In our model, changes in the nature of uncertainty influence predictions. Our model embeds the intuitions of both social comparison theory and persuasive argument theory into a more general mathematical framework. The framework illustrates how polarization arises without biases in beliefs and provides a precision description of conditions that lead to polarization. The model lacks detailed representation of the underlying psychological mechanism that (might) cause choice shifts. It does identify specific situations in which a shift will occur even if individual recommendations all agree. This possibility stretches the logic of the alternative theories.

7 Conclusion

This paper studies how the distribution of actions made by an optimizing decision maker changes as a function of the precision of the decision maker's information. It formalizes the proposition that the ex ante distribution of decisions become more variable as the information of the decision maker improves. It further presents conditions under which the ex post distribution

mendations of others may be sufficient to stimulate the individual to construct persuasive arguments.

of actions becomes more extreme as the information of the decision maker improves.

We argue that the formal model can organize well established results on the tendency for groups to make decisions that are more extreme than decisions that would be made by individual members of the group. In order to do this we treat a typical group polarization problem as if it were an example of rational information aggregation. We interpret a group's decision as the outcome of an optimization problem that rationally uses all of the information available to individual group members. In particular, it is based on superior information than what is available to any individual member of the group. Our theoretical results provide reasons why one might expect group decisions to be systematically more extreme than individual decisions.

The information aggregation model provides a framework that can organize a broad range of experimental observations. It also provides an intuitive, rational foundation for behavior that others have interpreted as evidence of systematic errors in group decision making. Nevertheless, the results only show that polarization can be consistent with rational decision making.

The model assumes that groups have no problems aggregating information and reaching a joint decision. Anyone who has served on a committee will know that these assumptions are unrealistic. There is strong academic and popular evidence that convinces us that groups can ofter 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 tragedies in the U.S. space program. Janis's [12] discussion of groupthink among President Kennedy's national security advisors foreshadows more recent failures of intelligence agencies in the United States. Glaeser and Sunstein [11] discuss situations where agents seem to be prone to systematic biases in the way they revise their opinions during a discussion. Both Janis and Glaeser and Sunstein argue that groups fail to account for the possible correlations in their members' information. If groups tend to overestimate the precision of their information (by neglecting correlations), then group decisions need not be superior to individual decisions. Our theoretical analysis demonstrates how this bias would lead to group polarization.

References

[1] Daron Acemoglu, Victor Chernozhukov, and Muhamet Yildiz. Learning and disagreement in an uncertain world. Technical report, MIT, October 2006.

[2] James Andreoni and Tymofiy Mylovanov. Diverging opinions. *American Economic Journal: Microeconomics*, 4(1):209–232, February 2012.

[3] Susan Athey and Jonathan Levin. The value of information in monotone decision problems. Technical report, Stanford University, December 2001.

[4] David Blackwell. Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, 24(2):265–272, 1953.

[5] Roger Brown. *Social Psychology*. New York : Free Press, second edition, 1986.

[6] Eugene Burnstein and Amiram Vinokur. What a person thinks upon learning he has chosen differently from others: Nice evidence for the persuasive-arguments explanation of choice shifts. *Journal of Experimental Social Psychology*, 11:412–426, 1975.

[7] Robert T. Clemen and Robert L. Winkler. Unanimity and compromise among probability forecasters. *Management Science*, 36(7):767–779, July 1990.

[8] Georges Darmois. Sur les lois de probabilite a estimation exhaustive. *CR Acad. Sci. Paris*, 260:1265–1266, 1935.

[9] Avinash K. Dixit and Joërgen W. Weibull. Political polarization. *Proceedings of the National Academy of Sciences*, 104(18):7351–7356, 2007.

[10] Kfir Eliaz, Debraj Ray, and Ronny Razin. Choice shifts in groups: A decision-theoretic basis. *American Economic Review*, 96(4):1321–1332, September 2006.

[11] Edward L. Glaeser and Cass R. Sunstein. Extremism and social learning. *Journal of Legal Analysis*, 1(1):263–324, 2009.

[12] Irving L. Janis. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Houghton Mifflin Company, Boston, 1982.

[13] Ian Jewitt. Information order in decision and agency problems. Technical report, Nuffield College, 2007.

[14] Rob Kaas, Dennis Dannenburg, and Marc Goovaerts. Exact credibility for weighted observations. *ASTIN bulletin*, 27(2):287–295, 1997.

[15] Bernard Osgood Koopman. On distributions admitting a sufficient statistic. *Transactions of the American Mathematical Society*, 39:399–409, May 1936.

[16] Eric L. Lehmann. Comparing Location Experiments. *Annals of Statistics*, 16(2):521–533, 1988.

[17] Haim Levy and Yoram Kroll. Ordering uncertain options with borrowing and lending. *The Journal of Finance*, 2:553–574, May 33.

[18] David G. Myers and H Lamm. The group polarization phenomenon. *Psychological Bulletin*, 83(4):602–627, 1976.

[19] Edwin James George Pitman. Sufficient statistics and intrinsic accuracy. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 32, pages 567–579. Cambridge Univ Press, 1936.

[20] John K.-H. Quah and Bruno Strulovici. Comparative statics, informativeness, and the interval dominance order. *Econometrica*, 77(6):1949–1992, November 2009.

[21] Nicolas Roux. *Three Essays in Information Economics*. PhD thesis, Paris School of Economics and University of Paris I, November 2012.

[22] David Schkade, Cass R. Sunstein, and Daniel Kahneman. Deliberating about dollars: The severity shift. *Columbia Law Review*, 100:1139–1175, 2000.

[23] Moshe Shaked and J. George Shanthikumar. *Stochastic Orders*. Springer, 2007.

[24] Joel Sobel. Forecasting group decisions from individual behavior. *Theoretical Economics*, forthcoming.

[25] James A. F. Stoner. Risky and cautious shifts in group decisions: The influence of widely held values. *Journal of Experimental Social Psychology*, 4:442–459, 1968.

[26] John C. Turner. *Social Influence*. Open University Press, Milton Keynes, 1991.

[27] Robert L. Winkler and Robert T. Clemen. Sensitivity of weights in combining forecasts. *Operations Research*, 40(3):609–614, May-June 1992.

Appendix

Proof of Proposition 1. Since $\delta(\cdot)$ is non-decreasing, it follows from Fact 1 that for all $z \in [0,1]$,

$$\delta\left(P_I(s \mid \mathcal{D}_I(s) \leq z)\right) \geq \delta\left(P_G(s \mid \mathcal{D}_G(s) \leq z)\right).$$

Since δ is linear, it follows that for all $z \in [0,1]$,

$$\int_{\underline{s}}^{\mathcal{D}_I^{-1}(z)} \delta(P_I(s))d\mathcal{D}_I(s) \geq \int_{\underline{s}}^{\mathcal{D}_G^{-1}(z)} \delta(P_G(s))d\mathcal{D}_G(s).$$

A change of variable yields $\int_0^z \delta(P_I(\mathcal{D}_I^{-1}(p)))dp \geq \int_0^z \delta(P_G(\mathcal{D}_G^{-1}(p))dp$. Levy and Kroll [17, Theorem 5′]show that this condition is equivalent to $\int_{-\infty}^{\tilde{a}} \Lambda_G da \geq \int_{-\infty}^{\tilde{a}} \Lambda_I da$ for all a. ■

Proof of Proposition 2. Since $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, it follows from Fact 1 that $z \in [0,1]$, $P_I(\cdot \mid \mathcal{D}_I \leq z) \succ_{MLR} P_G(\cdot \mid \mathcal{D}_G \leq z)$. Since $\delta(\cdot)$ is strictly increasing, it follows that $\delta(P_I(\mathcal{D}_I^{-1}(0))) > \delta(P_G(\mathcal{D}_G^{-1}(0)))$. This guarantees that $\Lambda_G(a) < \Lambda_I(a)$ for $a \in (\delta(P_G(\mathcal{D}_G^{-1}(0))), \delta(P_I(\mathcal{D}_I^{-1}(0))))$. A symmetric argument establishes that $\Lambda_G(a) > \Lambda_I(a)$ for a sufficiently high. ■

The next result identifies two equivalent implications of the MIO on conditional distributions of posteriors.

Lemma 3 *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, then for all $q \in [0,1]$,*

$$\Gamma_G^{-1}(\Gamma_I(q \mid \theta_0) \mid \theta_0) \leq \Gamma_G^{-1}(\Gamma_I(q \mid \theta_1) \mid \theta_1) \tag{11}$$

and

$$\Gamma_G^{-1}(1 - \Gamma_I(q \mid \theta_0) \mid \theta_0) \leq \Gamma_G^{-1}(1 - \Gamma_I(q \mid \theta_1) \mid \theta_1) \tag{12}$$

Equivalently, for all $q, q' \in [0,1]$:

$$\Gamma_G(q' \mid \theta_0) \leq \Gamma_I(q \mid \theta_0) \Rightarrow \Gamma_G(q' \mid \theta_1) \leq \Gamma_I(q \mid \theta_1) \tag{13}$$

and

$$\Gamma_G(1 - q' \mid \theta_0) \geq \Gamma_I(1 - q \mid \theta_0) \Rightarrow \Gamma_G(1 - q' \mid \theta_1) \geq \Gamma_I(1 - q \mid \theta_1). \tag{14}$$

Proof of Lemma 3. The first condition is based on Lehmann [16, Theorem 5′]'s characterization of the MIO-SC order: $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if, for all $s \in S_I$,

$$\mathcal{A}_G^{-1}(\mathcal{A}_I(s;\theta_0) \mid \theta_0) \leq \mathcal{A}_G^{-1}(\mathcal{A}_I(s;\theta_1) \mid \theta_1). \tag{15}$$

Since $P_i(s)$ is continuous and strictly increasing for $i = I, G$, Condition (15) can be written in terms of beliefs. That is, $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if $q \in [0, 1]$,

$$P_G(\mathcal{A}_G^{-1}(\mathcal{A}_I(P_I^{-1}(q));\theta_0) \mid \theta_0)) \leq P_G(\mathcal{A}_G^{-1}(\mathcal{A}_I(P_I^{-1}(q));\theta_1) \mid \theta_1)). \tag{16}$$

Since $\Gamma_i(q) = \mathcal{A}_i(P_i^{-1}(q))$, (11) follows from (16). The proof of Condition (12) is similar.

To show Condition (13), we use the following characterization of the MIO-SC order: [25] $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if for all s and s',

$$\mathcal{A}_G(s' \mid \theta_0) \leq \mathcal{A}_I(s \mid \theta_0) \Rightarrow \mathcal{A}_G(s' \mid \theta_1) \leq \mathcal{A}_I(s \mid \theta_1). \tag{17}$$

Since beliefs are non-decreasing with respect to signals, (17) implies (13) The proof of (14) is similar. ■

Proof of Lemma 1 $\Gamma_i(k \mid \theta_0)$ is the conditional probability (given θ_0) of receiving a signal that leads to a posterior belief that θ_0 is no greater than k. By symmetry, this is equalto the conditional probability (given θ_1) of receiving a signal that leads to a posterior belief that θ_0 is no greater than k. When posteriors are strictly increasing in beliefs, this quantity is $1 - \Gamma_i(1 - k \mid \theta_1)$. It follows that

$$\Gamma_i(k \mid \theta_0) = 1 - \Gamma_i(1 - q \mid \theta_1). \tag{18}$$

It follows from (18) that

$$\Gamma_G(q \mid \theta_0) \leq \Gamma_I(q \mid \theta_0) \Leftrightarrow \Gamma_G(1 - q \mid \theta_1) \geq \Gamma_I(1 - q \mid \theta_1). \tag{19}$$

[25] This characterization was originally proposed by Jewitt [13, page 4]. Since Jewitt [13] does not contain the proof, we provide it here. To show that (17) characterizes the MIO-SC, we prove that (15) and (17) are equivalent. Suppose (15) does not hold. Then, applying the function $\mathcal{A}_G(\cdot|\theta_1)$ on both sides of (15) yields $\mathcal{A}_G(s'|\theta_1) > \mathcal{A}_I(s|\theta_1)$ where $s' = \mathcal{A}_G^{-1}(\mathcal{A}_I(s;\theta_0) \mid \theta_0)$. As $\mathcal{A}_G(s'|\theta_0) = \mathcal{A}_I(s|\theta_0)$ by construction, (17) does not hold either.

Suppose now that (17) does not hold. It follows that $\mathcal{A}_G^{-1}(\mathcal{A}_I(s|\theta_1)|\theta_1) < s' < \mathcal{A}_G^{-1}(\mathcal{A}_I(s|\theta_0)|\theta_0)$, which contradicts (15).

Conditions (13) and (19) imply

$$\Gamma_G(q \mid \theta_0) \le \Gamma_I(q \mid \theta_0) \Leftrightarrow \Gamma_G(q \mid \theta_1) \ge \Gamma_I(q \mid \theta_1), \tag{20}$$

which in turn implies that

$$\Gamma_G(\tilde{q} \mid \theta_0) = \Gamma_I(\tilde{q} \mid \theta_0) \Leftrightarrow \Gamma_G(\tilde{q} \mid \theta_1) = \Gamma_I(\tilde{q} \mid \theta_1), \tag{21}$$

We claim that if

$$\Gamma_G(\tilde{q} \mid \theta_0) = \Gamma_I(\tilde{q} \mid \theta_0), \tag{22}$$

then

$$\int_0^{\hat{q}} \left(\Gamma_G(z \mid \theta_0) - \Gamma_I(z \mid \theta_0)\right) dz \ge \int_0^{\hat{q}} \left(\Gamma_G(z \mid \theta_1) - \Gamma_I(z \mid \theta_1)\right) dz. \tag{23}$$

To establish (23), note first that if (22) holds, then by (21),

$$\int_0^{\hat{q}} \left(\Gamma_G(z \mid \theta_k) - \Gamma_I(z \mid \theta_k)\right) dz = \int_0^{\Gamma_I(\hat{q} \mid \theta_k)} \left(\Gamma_I^{-1}(u \mid \theta_k) - \Gamma_G^{-1}(u \mid \theta_k)\right) du. \tag{24}$$

A change of variable enables us to use (24) to rewrite (23) as

$$\int_0^{\hat{q}} \left(q - \Gamma_G^{-1}(\Gamma_I(q \mid \theta_0) \mid \theta_0)\right) d\Gamma_I(q \mid \theta_0) \ge \int_0^{\hat{q}} \left(q - \Gamma_G^{-1}(\Gamma_I(q \mid \theta_1) \mid \theta_1)\right) d\Gamma_I(q \mid \theta_1) \tag{25}$$

provided that (22) holds. The claim that (22) implies (23) now follows from (11) and (25).

By symmetry, Athey and Levin [3]'s characterization of the MIO-SC order, $\int_0^q \Gamma_G(z)dz \ge \int_0^q \Gamma_I(z)dz$ for all $q \in [0,1]$ is equivalent to: for all $q \in [0,1]$,

$$\int_0^{q} \left(\Gamma_G(z \mid \theta_0) - \Gamma_I(z \mid \theta_0)\right) dz \ge \int_0^{q} \left(\Gamma_I(z \mid \theta_1) - \Gamma_G(z \mid \theta_1)\right) dz. \tag{26}$$

Conditions (23) and (26) imply that

$$\int_0^{q} \left(\Gamma_G(z \mid \theta_0) - \Gamma_I(z \mid \theta_0)\right) dz \ge 0 \tag{27}$$

whenever $\Gamma_G(q \mid \theta_0) = \Gamma_I(q \mid \theta_0)$, which in turn implies that (27) is non-negative for all q, the desired result. ∎

Proof of Proposition 4. Standard results (for example, Shaked and Shanthikumar [23]) imply that the first condition of Lemma 1 is equivalent to $E(\phi(\tilde{P}_{G,\theta_0})) \leq E(\phi(\tilde{P}_{I,\theta_0}))$ for all increasing, concave ϕ. Since the composition of two increasing concave functions is concave, it follows that if $\delta(\cdot)$ is increasing and concave, then for all increasing concave ϕ, $E(\phi(\delta(\tilde{P}_{G,\theta_0}))) \leq E(\phi(\delta(\tilde{P}_{I,\theta_0})))$, which establishes the first part of the proposition. The second part follows from a symmetric argument. ■

Proof of Lemma 2 We prove the first part of Lemma 2. The second part follows from a symmetric argument. Let

$$\underline{q} \equiv \sup\{q : \Gamma_G(q' \mid \theta_0) \geq \Gamma_I(q' \mid \theta_0) \text{ for all } q' \leq q\}. \tag{28}$$

If $\underline{q} = 0$, then because $\Gamma_i(\cdot \mid \theta_0)$ is monotone, there exists $\hat{q}$ such that

$$\Gamma_G(q \mid \theta_0) < \Gamma_I(q \mid \theta_0) \tag{29}$$

for $q \in (0, \hat{q})$. Inequality (29) and Condition (13) of Lemma 3 imply that

$$\Gamma_G(q \mid \theta_1) < \Gamma_I(q \mid \theta_1) \tag{30}$$

for $q \in (0, \hat{q})$. Conditions (29) and (30) imply that $\Gamma_G(q) < \Gamma_I(q)$, which contradicts the assumption that $\mathcal{I}_G$ is more precise that $\mathcal{I}_I$. It follows that $\underline{q} > 0$ Furthermore, if $\Gamma_G(q \mid \theta_0) \equiv \Gamma_I(q \mid \theta_0)$ on $[0, \underline{q}]$, then the same argument that shows $\underline{q} > 0$ shows that there is an $\epsilon > 0$ such that

$$\Gamma_G(q' \mid \theta_0) \geq \Gamma_I(q' \mid \theta_0) \text{ for all } q' \leq \underline{q} + \epsilon,$$

which contradicts the definition of $\underline{q}$. ■

Proof of Proposition 5 We prove the first statement. The second statement follows from a symmetric argument. Let $a' = a^*(\mathcal{D}_G^{-1}(\underline{q}))$ (for the $\underline{q}$ described in Lemma 2). When actions are strictly increasing in beliefs, Lemma 2 implies that $\Lambda_G(a \mid \theta_0) \geq \Lambda_I(a \mid \theta_0)$ for all $a \leq a'$ and $\Lambda_G(a \mid \theta_0) > \Lambda_I(a \mid \theta_0)$. ■