UC San Diego
UC San Diego Previously Published Works

Title
Group Polarization in a Model of Information Aggregation

Permalink
https://escholarship.org/uc/item/78t320tr

Journal
AMERICAN ECONOMIC JOURNAL-MICROECONOMICS, 7(4)

ISSN
1945-7669

Authors
Roux, Nicolas
Sobel, Joel

Publication Date
2015

DOI
10.1257/mic.20140252

Peer reviewed

AQ1 # Group Polarization in a Model of Information Aggregation[†]

By Nicolas Roux and Joel Sobel*

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. (JEL D71, D83)

Groups make decisions that are more extreme than some central tendency of the individual positions of the members of the group. Stoner (1968) first observed this phenomenon.[1] Other researchers have replicated and refined Stoner's insight, which the literature calls group polarization. Following Bordley (1983), 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 (2000). 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. Subjects then were randomly assigned to groups of six; these groups deliberated and decided on punitive verdicts. In those cases where individual punishment ratings are severe, the group ratings tend to be more severe

than the median individual rating. In those cases where individual punishment ratings are lenient, the group ratings tend to be more lenient than the median individual rating. As a result, group punitive verdicts are more polarized across cases than individual verdicts.[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 states 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, 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.

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 recommendation to (unmodeled) differences in the cognitive abilities. 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.[3]

AQ3   Individuals use different arguments to justify their decisions. They share these arguments during deliberation. Section V describes persuasive arguments theory, which psychologists have used to explain how balancing novel arguments during deliberation can lead to polarization. Our model can be seen as a statistical foundation for this theory.

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.

---

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

[3] Hong and Page (2009) present a model in which agents use different models to interpret information, leading to differences in beliefs. This approach provides conceptual support for the way we interpret our model.

Section IV also contains a discussion of another task studied by Schkade, Sunstein, and Kahneman. In addition to asking subjects to choose a punitive rating, they made them assign a damage award in dollars (a nonnegative number) for each case. This task differs somewhat from the standard example in the group-polarization literature because it places no upper bound on the action set. Group decisions exhibit two unusual properties. In several groups, the group punitive award 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, 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 I.

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

I. 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 an informative signal 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 of 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 IIIB. The relation between information precision and the ex post variability of actions is typically nonmonotonic. To see this note that both a perfectly informed decision maker and a decision maker receiving uninformative

---

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

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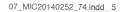
The examples in this section all illustrate ways one might expect the distribution of actions to become more variable as the information structure improves. Formal statements of these results require precise definitions of variability of distributions and improvement of information. We provide definitions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

A. *A Binary Example*

Example 1: There are two states of nature $\Theta = \{0, 1\}$, two signals $\mathcal{S} = \{0, 1\}$, and agents have to choose an action in the unit interval, $\mathcal{A} = [0, 1]$. The prior probability distribution puts equal weights on the two states, $\pi(0) = \pi(1) = \frac{1}{2}$. The conditional distributions over signals are $f(1 \,|\, 1) = f(0 \,|\, 0) = c, c \in [1/2, 1]$. 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 greater the better is the decision maker's information. This effect tends to reduce the variance of the ex post decision rule as c increases. A computation demonstrates that the variance of the optimal action given $\theta = 1$ is $c(1 - c)(1 - 2c)^2$, which is increasing over $(1/2, \hat{p})$ and decreasing over $(\hat{p}, 1)$ where $\hat{p} \in (1/2, 1)$.

B. *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 t and precision[7] $r > 0$, and that given θ, s is normally distributed 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 $r + Ir'$, where

$$(1) \qquad \mu^*(\mathbf{s}) = \frac{rt + r' \sum_{i=1}^{I} s_i}{r + r'I}.$$

In this example, the posterior distribution depends on the average signal. The optimal recommendation is simply the conditional mean of θ, $a^*(\mathbf{s}) = \mu^*(\mathbf{s})$.

In this example, a group of size I that receives I independent and identically distributed signals $(s_1, \ldots, s_I)$ each with precision r' acts the same as an individual who receives the signal $\bar{s} = \sum_{i=1}^{I} s_i/I$ with precision $r'I$. In particular, it follows from (1) that the action of the group can be written

$$\frac{rt + r'I\bar{s}}{r + r'I}.$$

Hence, it is natural to treat I as a parameter that measures the precision of the information as a function of group size.

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

Example 3: Take $\mathscr{A} = [0, 1]$, $\Theta = [0, 1]$, and assume that the prior is a Gamma distribution with parameters ρ and β. Let $\mathscr{S} = \mathbb{N}$ 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 $\bar{s} = \sum_{i=1}^{I} s_i/I$, the action rule is

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

---

[6] Kaas, Dannenburg, and Goovaerts (1997) describe the linearity property used in our analysis. Darmois (1935), Koopman (1936), and Pitman (1936) 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.

As in Example 2, as the group size increases, the action shifts away from the prior mean and toward the average signal.

The posterior distribution is Gamma with parameters $\rho + I$ and $\beta + \bar{s}$. The expected action 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 II presents a general framework that admits these examples as special cases.

II. The Framework

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

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

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

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 $P(s)$):

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

The state space Θ, the signal space $\mathcal{S}$, 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 $\mathcal{I} = (\mathcal{S}, \mathcal{D}, \{P(\cdot \mid s)\}_{s \in \mathcal{S}})$ 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 $s \in \mathcal{S}$ and updates her prior, she chooses the action that maximizes her expected utility given the resulting posterior belief

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

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

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

Athey and Levin (2001) characterized the MIO-SC order as a notion of variability of this measure:

FACT 1 (*Athey and Levin* 2001): $\mathcal{I}_G$ *is more precise than* $\mathcal{I}_I$ *if and only if for all* $z \in [0, 1]$, $P_I(\cdot \,|\, \mathcal{D}_I \leq z) \succ_{MLR} P_G(\cdot \,|\, \mathcal{D}_G \leq z)$.

$P_i(\cdot \,|\, \mathcal{D}_i \leq z)$, for $i = I, G$, stands for the posterior belief i holds when learning that his signal belongs to $\{s : \mathcal{D}_i(s) \leq z\}$. Given that both the group and the individual receive a low signal, the group is more confident that the state of nature is low since its information is 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 $P_I(\cdot \,|\, \mathcal{D}_I > z) \prec_{MLR} P_G(\cdot \,|\, \mathcal{D}_G > z)$ 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 $\mathcal{I} = (\mathcal{S}, \mathcal{D}(\cdot), \{P(\cdot \,|\, s)\}_{s \in \mathcal{S}})$ 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 $\Lambda(a) = \text{Prob}(\{s : a^*(s) \leq a\})$. 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" uses second-order stochastic dominance. Λ_I **second-order stochastically dominates** Λ_G if

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

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 2007, Chapter 4). If Λ_I and Λ_G have the same mean, then the inequality in (4) will hold as an equation when $\tilde{a} = \bar{a}$. 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 and the two distributions have equal means.

Proposition 1 presents conditions under which the distribution of the group's actions is more variable 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

$$\delta(P) = \operatorname*{argmax}_{a \in \mathcal{Q}} \int_{\theta \in \Theta} u(a, \theta) \, dP(\theta).$$

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 nondecreasing over $\{P(s)\}_{s \in \mathcal{S}}$. 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'.

[Margin note: Replace "technology" with "structure" Structure]

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

[Margin note: is; Replace "functions are" by "function is"]

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 a linear action function since the optimal action is the expected value of θ according to the posterior belief of the decision maker. It is straightforward to show that the action rule is linear if, in addition, there exists a cumulative distribution function $Q(\cdot)$ such that $P(\theta|s) \equiv Q(\theta - s)$. The action rule is linear in Examples 2 and 3.

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 2007, 200)). However, the following example shows that this is not true.

Example 4: 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 beliefs 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

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

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 toward $a = 0$ (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 function is 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 MLR order) for the group than for the individual. These lower posteriors lead to lower actions. Unlike the properties of Examples 1–3, 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 so the action function is not strictly increasing. 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 con-
fidence, than the individual.

Example 5: Let $\Theta = \mathscr{A} = \{0, 1, 2\}$ 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 prob-
ability 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 $P_I\big(\mathcal{D}_I^{-1}(0)\big)$ dominates $P_G\big(\mathcal{D}_G^{-1}(0)\big)$ 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 individu-
al's action distribution is contained in the convex hull of the support of the group's
action distribution.*

B. *Ex Post Analysis*

We now discuss how the distribution of actions changes with information condi-
tional 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 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. 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 IB 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.

---

[16] Similar examples can also be constructed in a two-state model.

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 $\int_0^{0.5} \Gamma_G(z\,|\,\theta_0)\, dz < \int_0^{0.5} \Gamma_I(z\,|\,\theta_0)\, dz$.

VERTICAL SPACE [TO SIGNAL END OF EXAMPLE]

In the example, the group's information is plainly superior to the individual's when the state is θ_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 $\theta = \theta_0$. The added precision of the group's information in state θ_1 compensates for lack of precision given θ_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 $\Gamma_G(q\,|\,\theta_0) \geq \Gamma_I(q\,|\,\theta_0)$ for $q \leq \underline{q}$ with strict inequality for a set of positive measure in $(0, \underline{q})$, and $\Gamma_G(q'\,|\,\theta_1) \leq \Gamma_I(q'\,|\,\theta_1)$ for $q' \geq \overline{q}$ 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 exist $a', a'' \in [\underline{a}, \overline{a}]$ such that*

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

and

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

Replace a' by a'' (twice)

If actions are not necessarily strictly increasing with respect to beliefs, then Proposition 5 does not hold. Consider a 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 is lower than that of the better informed decision maker conditional on the low state.

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 section, we derive a result under an alternative notion of information precision.

Single-Crossing Preferences.—In this section, we assume that preferences satisfy a **uniform single-crossing condition**. Specifically, we assume that $\mathscr{A} = [0, 1]$

and that there exist 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 $1 - \pi(\theta^*)$ the decision maker's posterior places probability 1 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. We omit the formal statement of this result and its proof (which is a direct consequence of the law of large numbers).

IV. Applications

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

A. *Application to Jury Decision Making*

Sunstein, Schkade, and Kahnemann (2000—henceforth, SSK) report on a study of jury decision making. In their experiment, subjects receive descriptions of 15 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 9 point rating scale going from 0 (none) to 8 (extremely severe) and they choose the damage award that the defendant should pay (a nonnegative 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

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 $(I - 1)(\theta - \alpha/\beta)/((1 + I/\beta)(1 + 1/\beta))$. 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.[18]

We have pointed out that the relation between information precision and the ex post variability of actions is typically nonmonotonic. In our model, the relation between the number of signals received and the ex post variability of actions is increasing over a 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 information 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 than 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

---

[18] A complementary explanation relates the severity shift to asymmetric losses. As explained in Section IIIB, 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.

A first approach, to which our model belongs, argues that it is the exchange of private information that makes groups exacerbate the predominant individual opinions of their members. Bordley (1983) studies a model of information aggregation in which there are two states. He uses this model to describe choice shifts identified in Stoner's experiments. In his model, each individual has a belief about the relative likelihood of the two states. This belief determines the individual's recommendation. Bordley (1982) provides an axiomatic framework that implies that the aggregate belief is a function of individual beliefs. Bordley (1983) demonstrates that this representation can give rise to the polarization in Stoner's experiments. These results capture the basic intuition for our results. Our results go beyond Bordley's because they give a formal definition of polarization, they apply to more than the two-state environment, and contain an explicit description of the complete statistical decision problem facing individuals and the group. There is a separate literature that examines the possibility that observing a common signal will cause the priors of different individuals to diverge.[20]

Sobel (2014) discusses the relationship between individual and group decisions without restrictions on information structures. He shows that 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.

Social psychologists also proposed an information exchange-based theory of group polarization. The persuasive arguments theory[21] 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 according to the number of arguments in favor of the risky alternative relative to the arguments in favor of the cautious alternative (for choice dilemma questionnaires). The proportion of risky to cautious arguments in a case is hypothesized 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, take $\mathscr{A} = [0, 1]$, $\Theta = [0, 1]$ 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 $\mathscr{S} = \{0, 1\}$, and let the conditional distribution of signals

---

[20] See Acemoglu, Chernozhukov, and Yildiz (2006), Andreoni and Mylovanov (2012), and Dixit and Weibull (2007) for additional models of this kind. Clemen and Winkler (1990) and Winkler and Clemen (1992) 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.

[21] See Brown (1986) for a clear exposition.

satisfy $f(1\,|\,\theta) = \theta$ and $f(0\,|\,\theta) = 1 - \theta$. 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

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

Interpret signals 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 $s = 1$ are 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.

The persuasive arguments 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, we assume 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 describe the difference between a group and an individual decision as a shift in a risky or more cautious direction. 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.

The persuasive arguments theory also differs from our model with respect to how information is measured. In our model, private information is summarized in actions (provided that different signals trigger different actions, and that preferences are known). PAT on the other hand assumes that private information amounts to the pro and con arguments an individual comes up with to justify his action. Measuring information as arguments is restrictive. For instance it predicts that two individuals with identical arguments should not change their opinions after deliberating. It thereby neglects the possibility that additional information may be brought by the mere fact that another individual independently came up with the same arguments. Baron et al. (1996) show that mere corroboration pushes people's opinions toward the extremes.

Information exchange based approaches consider that group decisions only differ from individual decisions with respect to the amount of information on which they are based. Group decisions are therefore superior. By contrast, the subsequent theories view group polarization as a symptom of poor decision making by groups.

The diffusion-of-responsibility theory (DOR) was introduced as an explanation for the early finding in choice dilemma questionnaires that groups tend to make riskier choices than individuals (see Wallach, Kogan, and Bem 1964 for an early reference). As group members share the responsibility in case their choice led to a poor outcome, they are more willing to take risks as a group than when choosing in isolation. A referee provided one way to formalize this idea. Assume that preferences

depend on the size of the group. Consider a two-state model in which actions are elements of $[-1, 1]$. It is plausible to identify extreme actions (-1 and 1) as risky and the action $a = 0$ as cautious. One can specify preferences (as a function of group size) so that the distance between the optimal action and $a = 0$ is increasing in the size of the group (and the sign of the optimal action depends on the state).

Eliaz, Ray, and Razin (2006) present a different model of choice shifts that relates to the diffusion-of-responsibility theory. 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 (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 and are based on a behavior that has a flavor of diffusion of responsibility. 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. 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.

Existing evidence suggests that group polarization can only be partially explained by a diffusion-of-responsibility argument. This kind of argument is based on preference aggregation, and therefore implies that individual opinions do not change in the course of group decision making. Yet, it is systematically the case that individual decisions collected after the discussion also polarize compared to the initial individual decisions. Group discussion then appears to aggregate information to some extent.[22]

For the subsequent theories, deliberation brings in new information, that makes individual decisions polarize. In contrast to our model, this information may not be about the problem, may not be fully communicated, or may not be rationally aggregated. According to the social comparison theory, individuals evaluate their actions relative to a norm of behavior that is reflected in the actions of others.[23] 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.[24] 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

---

[22] Stoner (1968) notes that groups' decisions polarize slightly more than post deliberation individual decisions.

[23] There are different versions of that theory, which are presented in Isenberg (1986).

[24] Brown (1986, 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."

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. 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.

Our 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 often make bad decisions for systematic reasons. The decision-making environment at NASA has been blamed for tragedies in the US space program. Janis's (1982) discussion of groupthink among President Kennedy's national security advisors foreshadows more recent failures of intelligence agencies in the United States. Sunstein (2000), Sunstein and Hastie (2008), and Glaeser and Sunstein (2009) argue that group polarization is the result of behavioral decision making. These papers assume that individuals exchange information about the problem, but the process of aggregation is subject to two sorts of distortions. First, the sample of information that is communicated tends to be biased in favor of the opinion that is initially dominant. This is due to social motives, which makes group members reluctant to provide information that contradicts an established consensus. Second, in Glaeser and Sunstein (2009) individuals fail to account for the fact that the information provided by other group members is biased in favor of the dominating opinion and neglect correlation between signals. The combination of these distortions makes individuals' opinions become extreme when they should not. They also lead to the possibility that the group's decision is inferior to individual choices.[25]

All the aforementioned mechanisms may be working simultaneously at generating group polarization. The extent to which each type of mechanism matters depends on the type of task and the population considered. As far as we know, disentangling information exchange theories from alternative theories based on the proto-typical group polarization dataset (including individual and group decisions) has yet to be done. There is an enormous experimental literature in social psychology that attempts to distinguish between the predictions of the psychological mechanisms behind polarization.[26] Proponents of the persuasive arguments theory attempted to show that exposure to arguments alone could ~~predict~~ *induce* change in attitude. In contrast, proponents of the social-comparison theory attempted to show that exposure to the decisions of others, without any explicit informational content, was sufficient to produce a change of opinion. Both undertakings were successful, which lead social psychologists to conclude that both mechanisms are at work.

[margin note: Replace "predict" with "induce"]

[margin note: Delete hyphen]

---

[25] In our model, the group's decision always leads to an ex ante superior decision than individual decisions because all actors are rational and the group has superior information. On the other hand, our formal results that provide conditions under which group decisions are more variable than individual decisions depend only on properties of the beliefs used to make decisions. Hence, if a group mistakenly acted as if it had superior information than individual group members, the group's decision would be more variable that individual decisions, but need not be superior to an individual group member's decision.

[26] See Brown (1986) or Isenberg (1986) for reviews.

PROOF OF LEMMA 3:

The first condition is based on Lehmann (1988, 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 $s \in S_I$,

$$(A5) \qquad \mathcal{F}_G^{-1}(\mathcal{F}_I(s; \theta_0)|\theta_0) \leq \mathcal{F}_G^{-1}(\mathcal{F}_I(s; \theta_1)|\theta_1).$$

Since $P_i(s)$ is continuous and strictly increasing for $i = I, G$, Condition (A5) can be written in terms of beliefs. That is, $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if $q \in [0, 1]$,

AQ7 $$(A6) \qquad P_G(\mathcal{F}_G^{-1}(\mathcal{F}_I(P_I^{-1}(q); \theta_0)|\theta_0)) \leq P_G(\mathcal{F}_G^{-1}(\mathcal{F}_I(P_I^{-1}(q); \theta_1)|\theta_1)).$$

Since $\Gamma_i(q) = \mathcal{F}_i(P_i^{-1}(q))$, (A1) follows from (A6). The proof of condition (A2) is similar.

To show condition (A3), we use the following characterization of the MIO-SC order: [27] $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if for all s and s',

$$(A7) \qquad \mathcal{F}_G(s'|\theta_0) \leq \mathcal{F}_I(s|\theta_0) \Rightarrow \mathcal{F}_G(s'|\theta_1) \leq \mathcal{F}_I(s|\theta_1).$$

Since beliefs are nondecreasing with respect to signals, (A7) implies (A3). The proof of (A4) is similar. ∎

PROOF OF LEMMA 1:

$\Gamma_i(k|\theta_0)$ 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 equal to the conditional probability (given θ_0) 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 $1 - \Gamma_i(1 - k|\theta_1)$. It follows that

$$(A8) \qquad \Gamma_i(k|\theta_0) = 1 - \Gamma_i(1 - q|\theta_1).$$

It follows from (A8) that

$$(A9) \qquad \Gamma_G(q|\theta_0) \leq \Gamma_I(q|\theta_0) \Leftrightarrow \Gamma_G(1 - q|\theta_1) \geq \Gamma_I(1 - q|\theta_1).$$

Conditions (A3) and (A9) imply

$$(A10) \qquad \Gamma_G(q|\theta_0) \leq \Gamma_I(q|\theta_0) \Leftrightarrow \Gamma_G(q|\theta_1) \geq \Gamma_I(q|\theta_1),$$

which in turn implies that

$$(A11) \qquad \Gamma_G(\tilde{q}|\theta_0) = \Gamma_I(\tilde{q}|\theta_0) \Leftrightarrow \Gamma_G(\tilde{q}|\theta_1) = \Gamma_I(\tilde{q}|\theta_1).$$

---

[27] This characterization was originally proposed by Jewitt (2007, 4). Since Jewitt (2007) does not contain the proof, we provide it here. To show that (A7) characterizes the MIO-SC, we prove that (A5) and (A7) are equivalent. Suppose (A5) does not hold. Then, applying the function $\mathcal{F}_G(\cdot|\theta_1)$ on both sides of (A5) yields $\mathcal{F}_G(s'|\theta_1) > \mathcal{F}_I(s|\theta_1)$, where $s' = \mathcal{F}_G^{-1}(\mathcal{F}_I(s|\theta_0)|\theta_0)$. As $\mathcal{F}_G(s'|\theta_0) = \mathcal{F}_I(s|\theta_0)$ by construction, (A7) does not hold either.
Suppose now that (A7) does not hold. It follows that $\mathcal{F}_G^{-1}(\mathcal{F}_I(s|\theta_1)|\theta_1) < s' < \mathcal{F}_G^{-1}(\mathcal{F}_I(s|\theta_0)|\theta_0)$, which contradicts (A5).

PLEASE ANSWER ALL AUTHOR QUERIES (numbered with "AQ" in the margin of the page). Please disregard all Editor Queries (numbered with "EQ" in the margins). They are reminders for the editorial staff.

AQ#   Question   Response

1.   Please provide shorter version of title   *Roux and Sobel :*
     for use as running head – 63 characters   *Information Aggregation*
     max, which includes author last names   *and Group Polarization*
     and spaces.

2.   Where you thank institutions for   *DONE*
     financial support, please include grant
     numbers if you have them.

3.   I deleted the hyphen and changed to
     persuasive arguments theory from   *OK*
     persuasive-argument theory. Notice
     I added an 's' to make 'argument'
     'arguments' based on the term in a
     psychology dictionary. Is this ok?

4.   I changed this from 'if more precise' to   *OK*
     'is more precise.'

5.   Please add Sunstein, Schkade, and   *CHANGED TO*
     Kahnemann(2000) to the references.   *Schkade, Sunstein, and*
                                                    Kahnemann in text

6.   Something appears to be missing from
     this sentence beginning "Assuming
     that an individual..', specifically in the   *FIXED*
     phrase 'when a pivotal member of the
     group..'' When a pivotal member of the
     group does what?

7.   Extra parens on both sides of (A6).   *FIXED*