Group Polarization in a Model of Information Aggregation

Nicolas Roux*and Joel Sobel†

JOB MARKET PAPER

November 18, 2012

Abstract

Group Polarization refers to the tendency of groups to make more extreme decisions than individuals. In decision problems under incomplete information, individuals make decisions based on their private information while groups gather the information of their members. As a result, groups are typically better informed than individuals. This paper sees group polarization as an outcome of the difference in access to information. We consider two decision makers, a group and an individual, facing a monotone decision problem. They share similar preferences and prior belief but the group is better informed than the individual in the sense the monotone information order (Athey and Levin, 2001). This paper presents sufficient conditions under which the group's decisions polarize as compared to the individual's. Group polarization is considered from two points of view: ex ante, the group's decisions are more dispersed in the sense of second order stochastic dominance if optimal actions are a linear function of beliefs. Conditional on a particular state, the group's actions are more extreme (in a stochastic sense) if optimal actions are linear in beliefs and the group has more information conditional on that state.

*Paris School of Economics–University of Paris 1
†University of California, San Diego

We present natural environments where the linearity condition fails and in which group polarization does not fully occur. We relate our results to existing experimental evidence. *Journal of Economic Literature* Classification Numbers: A12, D01; Keywords: statistical decision problem; group polarization; behavioral economics; psychology.

1 Introduction

Group polarization refers to the tendency of groups to make more extreme decisions than individuals. This phenomenon was first discovered by Stoner (1961) and has been extensively studied since then in the field of experimental social-psychology. This paper studies group polarization in decision problems under incomplete information. As groups gather the private information of their members, they are typically better informed than individuals. We attempt to see group polarization as a consequence of this difference in access to information.

An example of group polarization experiment is an experimental court case (Kaplan, 1977; Myers and Kaplan, 1976; Schkade, Sunstein and Kahneman, 2000). Jurors have to decide upon the severity of punishments to be inflicted to a defendant. After being exposed to the evidence, each juror makes up his mind and record an individual recommendation. Later on, jurors deliberate until they agree as to which verdict the jury will sentence. The comparison of individual and group (distribution of) verdicts reveals that groups reinforce the initial tendencies of individuals: in a case where isolated jurors recommend severe (lenient) punishments on average, the average juries' sentence is even more severe (lenient). Juries' verdicts are then more dispersed across cases. Group decisions are then said to polarize.

Section 2 points out that this pattern is consistent with simple models of information aggregation. In the models, jurors try to estimate the appropriate level of punishment (the state). They share the prior belief that the average defendant deserves intermediate punishments. Each juror then extracts a private signal from the trial. A juror's verdict results from the (convex) combination of the prior belief and his signal. Polarization arises because the relative weight attributed to the signal is increasing in its reliability. A relatively culpable defendant makes it likely for jurors to draw incriminating signals which make their verdicts moderately increase on average relative to intermediate punishments. Juries' signals are more reliable since they come from multiple sources. They are consequently weighted more heavily which makes juries increase their verdicts more strongly. In cases where the defendant is more culpable than initially

expected, juries sentence higher punishments than isolated jurors on average. The same mechanism holds for relatively innocent defendants so juries sentence lower punishments than isolated jurors in those cases. Group decisions then are more extreme than individual decisions conditional on essentially every state (ex post) and more dispersed across states (ex ante). The models in which those two results arise are specific: the utility function is quadratic and the signals are distributed according to standard distributions. In the rest of the paper, we consider a more general framework and determines conditions under which the two results follow from groups being better informed than individuals.

Section 3 presents the framework. We consider statistical decision problems. A decision maker faces a choice under incomplete information about future payoffs. Payoff relevant information is captured by a state variable whose realization is unknown. A decision maker receives private signals that convey information about the state. The informational content of signals is described by the joint distribution on signals and states, which will be referred to as the decision maker's information structure. Upon receiving his signal, the decision maker updates his prior belief and makes a utility maximizing decision. We consider two decision makers, a group and an individual, with similar utility function and prior belief. The group receives signals that are more informative for the state than the individual's. Intuitively, the group's signals are more correlated to the state than the individual's are. We aim to compare the group's and individual's actions.

We do so in monotone environments. Actions, states and signals are real numbers. The utility function and information structure are constrained so as to ensure that the decision maker responds to a higher signal by increasing his action. In the court case presented above, the state describes the level of culpability of the defendant. A juror (or a jury) receives signals that are more or less incriminating and responds to higher signals by increasing his punishment verdict. Working in monotone decision problems is restrictive but this framework seems like a reasonable approximation of the experimental tasks from the group polarization literature. Moreover, Sobel (2012) shows that in general decision problems, the mere assumption of one decision maker being better

2

informed than another does not constrain the way their decisions compare.

In monotone problems, a natural way to describe the assumption that the group is better informed is to use the *monotone information order* (Athey and Levin, 2001). The monotone information order can be thought of as an adaption of Blackwell information criterion (Blackwell, 1953) to monotone environments. An information structure Blackwell dominates another if it provides greater ex ante expected utility whatever the utility function considered. The monotone information order operates similarly on restricted classes of information structures and utility functions. In sections 4 and 5, we determine conditions under which the two polarization results introduced above results from the group's information being more precise than the individual's with respect to the monotone information order.

Section 4 studies conditions under which the group's decisions are more dispersed ex ante. The main result shows that higher dispersion of the group's actions is guarantied if actions respond linearly to changes in belief. The important point is that the precision of an information structure can be characterized by the ex ante dispersion of the posterior beliefs it induces. The better informed a decision maker, the more variable his posterior beliefs. Intuitively, a poorly informed juror does not extract much information from a trial so that he expects his posterior belief to remain close to the prior. His posteriors then are relatively constant across cases. Equipped with superior information, a jury's belief is likely to be more affected by a trial so its posteriors are more variable across states. Whether a well informed group's actions are more variable than a poorly informed individual's then depends upon the relation between beliefs and actions. If this relation is linear, then the group's actions are more variable in the sense of the *mean preserving increase in risk* order (Rothschild and Stiglitz, 1970). [1] This linearity property is satisfied for the class of quadratic utility functions used in the models of section 2. So the higher dispersion of group actions in those examples does not depend on the particular information structures we choose. Quadratic utility functions describe situations where a decision maker wants to hit a target. Losses incurred when failing to hit the

---

[1]The mean preserving increase order is equivalent to second order stochastic dominance when the two distributions being compared have equal means.

target only depends on the distance between the action and the target. There are natural situations however where this assumption is not sensible. For instance, in a court case context, it would typically be the case that overestimating punishments is more costly than underestimating them. The relation between belief and action would then be convex, which implies that an improvement in information shifts actions upward on average. Intuitively, it requires confidence (in one's information) about a defendant's culpability to sentence high punishments, which makes a well informed jury more likely to sentence high punishments. In those cases, we show that the group's decisions will be "asymmetrically more variable". The weaker the restrictions on the utility function, the weaker the sense in which the group's decisions are more variable. In fact, we show that in the general case, the group's decision could be less variable than the individual's.

Section 5 attempts to derive conditions under which the group's actions are more extreme conditional on every state. We treat the particular case where the state space contains 2 elements, so that there is a high (the defendant is culpable) and a low (the defendant is innocent) state. The group's actions will be higher (lower) conditional on the high (low) state if actions respond linearly to changes in beliefs and if the information structures are symmetric. The symmetry condition implies that an improvement in information ex ante translates to improvements conditional on both states. Therefore, it guaranties that the group is better informed whether the state is high or low. Provided that those two conditions are met, the sense in which the group's actions are more extreme is expressed by an integral condition on the distribution of actions. This condition implies in particular that the group's actions are more extreme on average. If actions are either concave or convex in beliefs, then the group's actions become more extreme only conditional on one state. Intuitively, if it is safer to sentence low punishments, then a poorly informed juror may sentence low punishments irrespective of the defendant. Conditional on the defendant being innocent, a well informed jury would not sentence lower punishments.

Section 6 discusses alternative accounts of the group polarization phenomenon.

This paper follows Sobel (2012) who studies the relationship between the decision of a group and the recommendations of its members provided that the

group aggregates the private information of its members. Sobel (2012) shows that without restrictions on the decision problem, there is no relationship between the recommendations of group members and the group's decision. Focusing on monotone problems, he shows that it requires strong restrictions in order for the group's decision to be bounded by the recommendation of its members, and concludes that the fact that a group tends to make relatively extreme decisions need not be evidence of systematic biases in group decision making. Our paper only focuses on monotone problems and studies conditions under which group polarization is likely to occur. Moreover, we do not constrain the group's information to be a combination of its members' information. Instead, the group and the individual have access to two distinguished information structures, and the sense in which the group is better informed is weaker than in Sobel (2012).

2 Specific Models of Group Polarization

This section motivates the subsequent analysis by presenting two simple models of information aggregation in which group polarization arises naturally and by showing how the models help organizing existing experimental evidence.

Most of the experimental research on group polarization was done in a task referred to as *choice dilemma questionnaire*. The task presents a hypothetical situation where a subject faces two alternative courses of action. One action, the cautious action, leads to a certain outcome. The alternative risky action leads to two possible outcomes, one better and one worse than the certain outcome. In the following example from Stoner (1968), the safe action is to go to the hospital, and the risky one is to board on plane.

> Mr. B is about to board a plane at the airport at the beginning of his overseas vacation. Although he has been looking forward to this trip for some time, he is troubled because he awoke in the morning with a severe abdominal pain. Because he was never flown before, he thinks that the pain may simply be an upset stomach brought on by anticipation of the flight. Although he is not far from a hospital

where he knows he will obtain quick attention, he realizes that a
visit to the hospital will cause him to miss his flight which in turn
will seriously disrupt his vacation plans. The pain has gotten more
severe in the last few minutes.

Subjects are presented with several cases similar to this one. For each case, a
subject rates on a scale the extent to which he would recommend an alternative
over the other. [2] Subjects are then divided into small groups and asked to reach
a consensual group rating. Finally, subjects record their individual ratings once
again after the group decision had been made. The comparison of group and
(initial) individual distributions of actions within cases reveals that individual
and group responses are related in two ways: first, in those cases where the
average individual response is relatively risky (cautious), the average group
response will be even more risky (cautious) (Stoner, 1968; Brown, 1986). Sec-
ond, the more extreme the average individual response, the more strongly the
average group response shifts toward the extreme. (Brown, 1986; Teger and
Pruitt, 1967).

The following model of information aggregation accounts for both results.
Faced with a case, subjects have to pick a response $a \in [0,1]$, where 1 corre-
sponds to the extreme preference for the risky alternative. A case is described
by an unknown state $\theta \in [0,1]$. Subjects share similar preferences represented
by the utility function $u(a,\theta) = -(a - \theta)^2$ so that the state can be thought of
as the appropriate response. Subjects hold a common prior belief on θ which is
Beta distributed on $[0,1]$ with parameters tr and $(1-t)r$. t is the initial mean
of θ and r the precision of the prior. [3] Each subject reads the description of
the problem and forms an opinion as to which response he should make. We
model this process by assuming that he receives k binary signals independently
drawn from the set $S = \{0,1\}$ where the probability of receiving signal $s = 1$ is
θ. Signals 0 and 1 can be thought of as arguments pro-caution and pro-risk (re-

---

[2]In the early literature subjects were asked to provide the minimum probability of success
in the risky alternative required to choose the risky alternative. Using this probability scale was
proved to be equivalent to using a scale indicating the extent of agreement with the cautious
alternative (Fraser, Gouge and Billig, 1971).

[3]The precision is the inverse of the variance

spectively) a subject comes up with to motivate his choice. [4] Subjects exchange their opinions during the group discussion. [5] As a result an isolated subject makes a decision based on k signals while a group composed of I members has Ik signals.

Equipped with a set of n signals, a decision maker updates the prior belief using Bayes rule and makes a decision

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

where $\bar{s}$ is the mean of the n signals. The decision rule is a weighted average of t, the initial mean of θ, and $\bar{s}$ the mean of the received signals. Group polarization arises naturally in this model because the weight placed on signals is increasing in the number of signals received. In a particular case θ, the group and each subject receive the same proportion of pro-risk signals on average (that is θ) but the group's expected response is more extreme relative to the initial mean t because it weights his information more strongly. In those cases where the situation requires a riskier response than initially expected ($\theta > t$), the group tends to make more severe punishments than an isolated subject, while the converse holds when $t > \theta$. Moreover, the more extreme the state θ (relative the initial mean t), the stronger the shift of the group response toward extremity. The difference between the average group response and the average individual response is $\frac{r(I-1)}{(r+I)(r+1)}(\theta - t)$ which is increasing (in absolute value) with respect to $\theta - t$.

Given that the group is better informed than the individual, it is able to better discriminate the appropriate response from one case to another. This translates to his decisions being more variable across cases. The ex ante variance of a decision maker's responses is $n[r^2t(1-t)]/[(n+r)(1+r)]$ which is indeed increasing in the number of signals received.

Group polarization has been observed in various contexts. In Myers, Woj-

---

[4]Silverthorne (1971) and Vinokur and Burstein (1974) (among others) observe that subjects motivate their response using arguments favoring either alternatives. They show that the ratio of pro-caution vs pro-risk arguments collected in a population of subjects predicts the average individual response in the case. The group discussion then consists in exchanging those arguments.

[5]Since they have identical preferences, there are no problems of information revelation.

cicki and Aardema (1977), subjects are simply asked to rate the extent to which they agree with some proposition such as "should ministers take a stand on political issues". Kaplan (1977) and Schkade, Sunstein and Kahneman (2000) study group polarization in experimental court cases. Subjects are presented with a description of a court case and are asked to rate the severity of punishment on a (bounded) scale. The model presented above can be applied to these situations as an individual's response involves some private information (his view on the problem) which is exchanged during group discussion.

It is also possible to consider slightly different tasks. In Schkade, Sunstein and Kahneman (2000), subjects evaluate the punitive awards (in dollars) to be inflicted to a defendant. The action set therefore is the set of positive numbers. In the following model of punitive awards evaluation, group polarization arises as in the previous model. Subjects try to estimate the appropriate amount of punitive awards which is captured by a state $\theta \in [0, \infty)$. They have quadratic preferences, $u(a, \theta) = -(a - \theta)^2$. The prior is a Gamma distribution with parameters α and β. Each subject draws a signal from the set $S = \mathbb{N}$. The conditional distribution of signals is a Poisson distribution with parameter θ. An individual makes a decision based on one signal while a group of size I receives I signals. Let us note $\bar{s} = \sum_{i=1}^{I} s_i / I$ the mean of the group's signals. As in the previous model, the action rule is weighted average of the initial mean of θ, α/β and the mean of the signals received (where the weights on signals increases with the number of signals):

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

Therefore, the same group polarization results hold in this model. The difference between group and individual expected actions conditional on θ writes $\frac{(I-1)}{(1+I/\beta)(1+1/\beta)}(\theta - \alpha/\beta)$ which is positive when θ is greater than initially expected. Moreover, it is linearly increasing in the difference between θ and t. More extremely culpable or extremely innocent defendants produce greater shifts of group decisions. However, the prior has a mean greater than its median. It follows that the average state among cases that induce upward polarization is more extreme (relative to the mean) than the average state among

8

cases than produce downwards polarization. As a result, upward shifts appear to be stronger than downward shifts. Schkade, Sunstein and Kahneman (2000) observe this result in their experiment and refer to it as severity shift. [6]

The two models above have the common feature that group decisions are more extreme on average conditional on essentially every state and are more variable ex ante. They feature specific information structures and quadratic utility function. In the rest of the paper, we generalize those two results to a wider class of decision problems, namely monotone problems. We identify conditions under which the decisions of a well informed group are likely to polarize as compared to those of a poorly informed individual.

3   General Framework

An agent recommends an action a in an interval $A = [\underline{a}, \overline{a}]$. His 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). The agent receives a signal s informative for the state of nature that is drawn from an interval $S = [\underline{s}, \overline{s}]$. A joint distribution $\mathcal{F}$ defined on $\Theta \times S$ describes the information structure (f denotes the density). We conduct our analysis using the fixed prior, $\Pi(\cdot)$ and the conditional distribution of signals $\mathcal{A}(\cdot|\theta)$ given that the state is θ ($\alpha(\cdot|\theta)$ is the corresponding density or probability mass function). Let us note the ex ante distribution on signals $\mathcal{D}$ where $\mathcal{D}(\cdot)$ is obtained as follows:

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

We assume without loss of generality the ex ante distribution of signals is strictly increasing. [7] Given that the prior is fixed, we will refer to $\mathcal{I} = \{S, \{\mathcal{A}(\cdot|\theta)\}_{\theta \in \Theta}\}$ as the information structure.

A decision maker who receives signal s updates his prior belief according

---

[6] We provide a complementary explanation of the severity shift in section 5

[7] Assuming that $\mathcal{D}$ is strictly increasing and S is an interval 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 (Lehmann, 1988).

to Bayes rule and obtains a posterior distribution denoted $P(\cdot|s)$ (or $P(s)$):

$$P(\theta|s) = \frac{\alpha(s|\theta)\pi(\theta)}{\int_{\omega \in \Theta} \alpha(s|\omega)\pi(\omega)d\omega}.$$

He then chooses the action that maximizes his expected utility given the resulting posterior belief, i.e.

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

where $a^*(\cdot)$ is referred to as the *action rule* of the decision maker. [8]

We restrict the analysis to monotone environments, i.e. problems where the action rule is increasing with respect to signals. Athey and Levin (2001) show that the monotonicity of the action rule obtains 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 $a' > a$. For instance, incremental returns may be assumed to be increasing with respect to θ, in which case the utility function is supermodular. The class of information structures is characterized by the sense in which the set of induced posteriors, $\{P(\cdot|s)\}_{s \in S}$, is ordered. Increasing the signal may for instance induce a first order stochastic dominance shift of the posterior.

The more restricted the class of utility functions, the larger the class of information structures inducing an increasing action rule. Those information structures will be called *monotone* (for a given class of utility functions). If the utility function satisfies the single-crossing property, i.e. $r(\theta) \geq 0$ implies $r(\theta') \geq 0$, then the action rule is increasing if posteriors are ordered by the *monotone likelihood ratio* order. [9] If the utility function can be further assumed to be supermodular, 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.

In this paper, we do not restrict attention to a particular class of monotone decision problems. We will therefore not specify which classes of utility func-

---

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

[9]Quah and Strulovici (2009) show that actions are increasing as a function of changes of belief in monotone likelihood ratio for a larger class of utility functions.

tions and information structures we are referring to.

In this framework, we attempt to establish a relation between the precision of an information structure and the extremity of the 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. A natural way of modeling information precision in monotone environments is to use the *monotone information order* (MIO) introduced by Athey and Levin (2001) which 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. [10]

Holding the utility function fixed, we compare the group's actions induced by $\mathcal{I}_G$ to the individual's actions induced by $\mathcal{I}_I$. We are interested in deriving the sense and conditions in which the group's decisions polarize as compared to the individual's. We do so from two points of view. Ex ante, the group's decisions are more dispersed. A well informed group is able to better discriminate a state from another than a poorly informed individual. Consequently, the individual's actions are relatively constant across states as compared to the group's actions. Conditional on a particular state, the group's actions are higher conditional on a high state and lower conditional on a low state.

4   Information Precision and Ex Ante Variability of Actions

The examples presented in section 2 suggest that a well informed group makes more variable decisions ex ante than a poorly informed individual. This section derives sufficient conditions on the utility function for this result to hold in monotone problems. Information precision is measured by ex ante variability of beliefs: the more precise the information structure, the more variable (ex ante) the beliefs it induces. As a result, whether group actions are more variable

---

[10]The most famous criterion for information precision is due to Blackwell (1953). 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.

ex ante than individual actions depends upon the relation between beliefs and actions, which is in turn related to properties of the utility function.

Proposition 1 and 2 relate the sense in which the group's actions are more variable to the familiar *second order stochastic dominance* order. Those two propositions make restrictive assumptions on the utility function as actions must be concave, convex (proposition 2) or linear (proposition 1) with respect to beliefs. If actions are only assumed to be strictly increasing in beliefs, proposition 3 shows that group actions are more variable with respect to a tail ordering. Example 1 shows that without any restriction, the group may exhibit less variable actions than the individual in the sense of second order stochastic dominance.

Lemma 1, which is due to Athey and Levin (2001) formalizes the intuition that the precision of an information structure can be measured by the ex ante variability of its posteriors. A monotone information structure $\mathcal{I}$ induces a set of posteriors $\{P(\cdot|s)\}_{s\in S} \subset \Delta(\Theta)$ which is ordered in some relevant stochastic sense noted $\geq_{st}$ and an ex ante distribution of signals, $\mathcal{D}(\cdot)$. As a result, it induces a random variable taking value in the belief space $\tilde{P}$ whose p-percentiles are noted $P(\cdot|\mathcal{D}^{-1}(p))$.

Lemma 1 (Athey and Levin (2001)). *$\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if for all $q \in [0,1]$, $P_I(\cdot \mid \mathcal{D}_I \leq q) \geq_{st} P_G(\cdot \mid \mathcal{D}_G \leq q)$.*

Given that the group and the individual receive informations $\mathcal{D}_G(s) \leq q$ and $\mathcal{D}_I(s) \leq q$ respectively, the group should be more confident that the state of nature is low since its information is more precise. This is represented by the fact that the group's posterior is lower than the individual's with respect to the relevant stochastic order. The group would then choose a lower action as compared the individual. Since the prior is fixed, it follows that $P_I(\cdot \mid \mathcal{D}_I \leq 1) = P_G(\cdot \mid \mathcal{D}_G \leq 1) = \Pi(\cdot)$. The condition of lemma 1 can therefore be rewritten the other way around: for all $q \in [0,1]$, $P_I(\cdot \mid \mathcal{D}_I \geq q) \leq_{st} P_G(\cdot \mid \mathcal{D}_G \geq q)$. Provided that that the signal is "high" for both the group and the individual, the group should be more confident that the state is high.

To see that this condition captures the idea that group posteriors are more variable ex ante, it is convenient to restrict attention to problems with two states of nature. When Θ contains two elements, the belief space is the unit interval (a belief is represented by the probability of the high state). The random

12

variable $\tilde{P}$ can then be conveniently represented by its distribution over $[0,1]$, denoted Γ. Then $P(\cdot|\mathcal{D}_G \leq q)$ corresponds to $\int_0^q \Gamma^{-1}(z)dz$, where Γ^{-1} is the quantile function of $\tilde{P}$. Athey and Levin (2001)'s condition then reduces to $\int_0^q \Gamma_G^{-1}(z)dz \leq \int_0^q \Gamma_I^{-1}(z)dz$ which is the quantile formulation of *mean preserving increase in risk* relation (Rothschild and Stiglitz, 1970). So, Athey and Levin (2001)'s condition on $\tilde{P}_I$ and $\tilde{P}_G$ is a multivariate generalization of the *mean preserving increase* ranking. [11]

Since an information structure and a utility function give rise to a (ex ante) distribution of signals $\mathcal{D}$ and an action rule $a^* : S \to A$, they generate ex ante a random variable, $\tilde{A}$, taking value in the action set A. Let us note the associated distribution $\Lambda(\cdot)$ in which $\Lambda(a) = \mathrm{Prob}(\{s : a^*(s) \leq a\})$. Defining $s_a^*(a)$ by $s_a^*(a) \equiv \sup\{s : a^*(s) \leq a\}$ if $\{s : a^*(s) \leq a\}$ contains at least one element and $s_a^*(a) \equiv \underline{s}$ if $\{s : a^*(s) \leq a\}$ is empty, we obtain $\Lambda(a) = \mathcal{D}(s_a^*(a))$. Note that $a^*(\mathcal{D}^{-1}(\cdot))$ is the quantile function associated to $\Lambda(\cdot)$.

We relate the precision of the information structure to the variability of $\tilde{A}$. To do so, it is convenient to define the following function: given some preferences $u(a,\theta)$, the *action function* δ associates to each posterior P in $\Delta(\Theta)$ an optimal decision $a = \delta(P)$, where

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

The action function is defined over the beliefs, rather than a linear space. It only depends on the utility function so the group and the individual share the same δ. The action function is increasing with respect to posteriors by construction. We abuse terminology by calling $\delta(\cdot)$ linear, concave or convex for a given information structure if $\delta(\gamma P + (1-\gamma)P)$ is equal, higher or lower (respectively) than $\gamma\delta(P) + (1-\gamma)\delta(P)$ for $\gamma \in (0,1)$ and for any posteriors P, P' induced by this information structure.

Proposition 1 presents a sufficient condition on the action function for the distributions of group and individual actions to be ranked in terms of *mean preserving increase in risk*.

---

[11]A random variable is a mean preserving increase of another if and only if the two random variables have the same mean and are ranked by second order stochastic dominance.

Proposition 1. *Suppose $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$. If the action function is linear for $\mathcal{I}_G$ and $\mathcal{I}_I$, then $\int_{a \leq \tilde{a}} \Lambda_G(a)da \geq \int_{a \leq \tilde{a}} \Lambda_I(a)da$ for all $\tilde{a} \in A$ with equality for $\tilde{a} = \bar{a}$.*

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.

Claim 1 presents sufficient conditions on the utility function for the action function to be linear for any information structure.

Claim 1. *The action function is linear for any information structure if the utility function is quadratic, i.e. $u(a, \theta) = -(a - h(\theta))^2$.*

If the marginal return to action satisfy the too conditions set in claim 1, then they take the form: $u_a'(a, \theta) = a + h(\theta)$. The optimal action associated to signal s equalizes the expected marginal returns to action to 0, i.e. $\int_{\theta \in \Theta} a + h(\theta)dP(\theta|s) = 0$. The optimal action is linear in the expected value of $h(\theta)$, so is linear in P.

Quadratic utility functions model situations where decision makers can be seen as trying to hit a target $h(\theta)$. The costs associated to missing the target only depends on how far from the target the action is.

This assumption may be sensible in the context of the choice dilemma questionnaire presented in section 2. Subjects are presented with two alternative courses of actions and indicate the extent to which one is preferable to the other. On the other hand, when the tasks consists in evaluating the punitive awards to be inflicted to a defendant as in Schkade, Sunstein and Kahneman (2000), it may be natural to assume that over-estimating the appropriate punishments is more costly than under-estimating them. [12] The verdicts of the a juror then tend to be biased downward. Example 1 presents a decision model describing this bias and shows that it induces the action function to be convex with respect to beliefs.

Example 1. Consider that a juror wants to match his verdict a with the appropriate amount of punitive awards θ. a and θ are both positive numbers.

---

[12]This assumption would have more bite in an actual court case where verdicts have consequences than in an experimental court case.

His payoffs are modeled by the function $U(a - \theta)$ which is concave,twice differentiable and reaches its maximum at 0. The asymmetry in losses is modeled by assuming that $U'(\cdot)$ is concave. To see that the jurors' actions are convex with respect to beliefs, consider two degenerate beliefs on θ and θ' that we will note (abusing notations) θ and θ'. The optimal actions associated to those beliefs are $\delta(\theta) = \theta$ and $\delta(\theta') = \theta'$ respectively. Now, consider that the juror believes that each state occurs with probability $1/2$ and evaluates the punitive awards at $1/2\theta + 1/2\theta'$. His expected marginal gains then is $\frac{1}{2}U'(\frac{1}{2}(\theta - \theta')) + \frac{1}{2}U'(\frac{1}{2}(\theta' - \theta))$ which is strictly negative since U' is concave. The optimal evaluation therefore is lower than $\frac{1}{2}\theta + \frac{1}{2}\theta'$ so the action function is convex.

There also are situations where missing the target is more serious when the target is high than when it is low. Consider a police commissioner who must decide how many policemen are needed to ensure security around a strike. He only has noisy information about the number of participants. Suppose that sending too many or too few policemen is equally costly: in the first case, the media will critic a disproportionate use of force while in the other they will blame his recklessness. However, the bigger the strike, the greater the media coverage so the greater the costs of the mistake. The police commissioner's action would then be biased upwards. Example 2 presents a model that captures this situation and shows that it leads to a concave action function.

Example 2. The action a and the state θ are positive numbers. The utility function is $u(a, \theta) = -\theta(a - \theta)^2$. The marginal returns to action function (which satisfies the single-crossing property) writes $-2\theta(a - 1)$. $\delta(P)$ is therefore equal to the ratio of the second order moment of P to the expected value of P:

$$\delta(P) = \frac{E_P(\theta^2)}{E_P(\theta)}.$$

The action function is concave in P if for all P, P' induced by the information structure, $\delta(P') > \delta(P) \Rightarrow E(P') \geq E(P)$.

When the utility function exhibits a bias (towards high or low actions) that is captured by a concave or convex action function, the less precise an agent's

15

information, the more biased his decisions. In example 1, it requires confidence about the defendant's culpability to sentence high punitive awards. Equipped with superior information, a jury is more likely to be confident enough to sentence high awards than an isolated juror. As a result, the group and individual distributions of actions have different means and cannot be compared as in proposition 1. Proposition 2 formalizes these intuitions in the case where the state space contains two elements.

Proposition 2. *Consider that Θ has two elements. Suppose $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, and the suppose that the action function is strictly increasing for $\mathcal{I}_G$ and $\mathcal{I}_I$:*

- *If the action function is concave for $\mathcal{I}_G$ and $\mathcal{I}_I$, then $\int_{a \leq \tilde{a}} \Lambda_G(a) da \geq \int_{a \leq \tilde{a}} \Lambda_I(a) da$.*

- *If the action function is convex for $\mathcal{I}_G$ and $\mathcal{I}_I$, then $\int_{a \geq \tilde{a}} \Lambda_G(a) da \leq \int_{a \geq \tilde{a}} \Lambda_I(a) da$.*

Proposition 2 provides a sense in which the group's actions are "asymmetrically more variable" than the individual's. In proposition 1 the two conditions set in proposition 2 held simultaneously. This was a way of saying that the group's actions are more extreme on both sides of the action set. When the action function is convex (concave), then the group's actions are only more extreme at the top (bottom) of the action set in the sense of the integral condition set in proposition 2. At the other hand of the action set, the sense in which the group's decisions are more extreme is expressed through a (weaker) tail ordering, as explained in proposition 3. [13]

Proposition 3 shows that the tails of the group's distribution of actions are fatter than those of the individual's distribution of actions. This result only requires the action function to be strictly increasing, i.e. $P' >_{st} P$ implies $\delta(P') > \delta(P)$ for $P, P' \in \{P(s)\}_{s \in S}$. It is therefore applicable to a much wider class of monotone environments.

---

[13]One may think that adjusting the distributions of actions for their mean makes it possible to compare the resulting distribution according to the mean preserving increase in risk relation. To see that this is not true, consider the following characterization of the mean-preserving increase relation: when there are two states of the world, $\tilde{P}_G$ is a mean preserving increase of $\tilde{P}_I$ if and only if $E(\psi(\tilde{P}_G)) \geq E(\psi(\tilde{P}_I))$ for any convex function ψ. This latter condition does not apply on the mean adjusted distributions $\delta(\tilde{P}_G) - E(\delta(\tilde{P}_G))$ and $\delta(\tilde{P}_I) - E(\delta(\tilde{P}_I))$ in general.

Proposition 3. *Suppose the action function is strictly increasing for $\mathcal{I}_G$ and $\mathcal{I}_I$. Suppose also that both information structures induce a finite number of posteriors. If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, then:*

- *if $\Lambda_G(a) < \Lambda_I(a)$ for some $a \in A$, then there exists $a' < a$ such that $\Lambda_G(a') > \Lambda_I(a')$.*

- *if $\Lambda_G(a) > \Lambda_I(a)$ for some $a \in A$, then there exists $a' > a$ such that $\Lambda_G(a') < \Lambda_I(a')$.*

The conclusions of Proposition 3 require the assumption that the action function is strictly increasing. Example 3 demonstrates that this assumption is necessary. In example 3, payments are quadratic but the action set is discrete. As a result, different posteriors induce the same action so that the greater variability of group beliefs is hidden at the level of actions. In fact, the point of example 3 is even stronger: the result of proposition 1 is reversed, i.e. an improvement in information precision leads to a decrease in action variability in the sense of second order stochastic dominance.

Example 3. There are two states of nature 0 and 1. The action set contains three actions: $\{0, 1/2, 1\}$ and preferences are quadratic $u(a, \theta) = -(a - \theta)^2$. The intermediate action is never optimal under perfect information. It is therefore natural to expect a well informed group not to play this action as often as a poorly informed individual. This example shows that this insight is not accurate. The group receives three signals, high medium and low:

	0	1
l	8/10	1/10
m	1/10	1/10
h	1/10	8/10

The individual receives only two signals high and low:

	0	1
l	3/4	1/4
h	1/4	3/4

The group's information is more precise than the individual's. However, the individual only plays extreme actions (with equal probability) while the group plays the intermediate action with positive probability. Indeed, the group posteriors about state 1 are 1/9, 1/2 and 8/9, where the intermediate posteriors occurs with probability 1/10. The individual posteriors are 1/4 and 3/4. Given that preferences are quadratic, a decision maker should play 0 if and only if his belief about state 1 is lower than 1/4 and should play 1 if and only if his belief is higher than 3/4. The individual has a more dispersed distribution of action in the sense of second order stochastic dominance.

There is a straightforward, but weak, conclusion that holds for all monotone decision problems.

Proposition 4. *If the group information structure is more precise than the individual's, then the support of the ex ante distribution of individual decisions is included in the range of the support of the ex ante distribution of group decisions.*

5 Information Precision and Ex Post Extremity of Decisions

The models presented in section 2 have the property that the group's actions are more variable ex ante than the individual's. The ex ante analysis of section 4 shows that quadratic utility functions are important to obtain this result. This section considers the other property of the models of section 2, namely that the group's actions are more extreme on average as compared to the individual's conditional on essentially all states of the world.

We work in the case where the state space contains two elements. As in section 4, precision of an information structure is associated to properties of posteriors and therefore to properties of actions. In the ex ante case, we could use the known characterization of information precision as ex ante dispersion of posteriors. In this section, we have to determine how information precision translates to the conditional distributions of posteriors. Lemma 2 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 which implies that an ex ante improvement in information precision translates to improvements in precision conditional on both states, then posteriors are higher conditional on the high state and lower conditional on the low state. Posteriors being higher or lower is expressed using the integral condition of proposition 2. Without the symmetry condition, posteriors are only more extreme in the sense that the tail of the ex post distributions of posteriors are fatter. The extent to which those properties on posteriors translate to actions is discussed similarly as in section 4.

Conditional on each state θ, an information structure generates a random variable $\tilde{P}_\theta$ on beliefs whose support is the same as $\tilde{P}$'s, i.e. $\{P(\cdot|s)\}_{s \in S}$. But the distribution of signals is now taken conditional on θ, i.e. $\mathcal{A}(\cdot|\theta)$. The p-percentile of $\tilde{P}_\theta$ is $P(\cdot|\mathcal{A}^{-1}(q|\theta))$. In the following definition, $\mathrm{P}(\cdot, \cdot)$ is the joint density on $\Theta \times S$.

Definition 1. *An information structure is symmetric if $\theta \in \Theta$ iff $-\theta \in \Theta$, $s \in S$ iff $-s \in S$ and $\mathrm{P}(s, \theta) = \mathrm{P}(-s, -\theta)$.*

The symmetry condition implies that an ex ante improvement in information precision translates to ex post improvements in information precision. As a result, the higher ex ante dispersion of group posteriors translates to group posteriors being more extreme conditional on both states. Lemma 2 provides a formal meaning of this statement.

Lemma 2. *Suppose $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ and both information structures are symmetric. Moreover, suppose that both information structures have posteriors strictly increasing with respect to signals. Then $\int_0^q \Gamma_G(z|\theta_0)dz \geq \int_0^q \Gamma_I(z|\theta_0)dz$ for all $q \in [0,1]$, and $\int_q^1 \Gamma_G(z|\theta_1)dz \leq \int_q^1 \Gamma_I(z|\theta_1)dz$ for all $q \in [0,1]$.*

The stochastic orders involved in lemma 2 provide a sense in which the group's posteriors should be 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 5 states this result.

Let us note the conditional distribution of actions $\Lambda(\cdot|\theta)$ where $\Lambda(a|\theta) = \mathcal{A}(s_a^*(a)|\theta)$. The associated random variable is noted $\tilde{A}_\theta$. $a^*(\mathcal{A}^{-1}(\cdot)|\theta)$ is the quantile function associated to $\Lambda(\cdot|\theta)$.

Proposition 5. *Suppose $\mathcal{I}_G$ and $\mathcal{I}_I$ satisfy the assumptions of lemma 2.*

- *If the action function is concave for $\mathcal{I}_G$ and $\mathcal{I}_I$, then $\int_{\underline{a}}^a \Lambda_G(z|\theta_0)dz \geq \int_{\underline{a}}^a \Lambda_I(z|\theta_0)dz$ for all $a \in A$*

- *If the action function is convex for $\mathcal{I}_G$ and $\mathcal{I}_I$, then $\int_a^{\overline{a}} \Lambda_G(z|\theta_1)dz \leq \int_a^{\overline{a}} \Lambda_I(z|\theta_1)dz$ for all $a \in A$.*

In the particular case where the action function is linear, both conditions stated in proposition 5 apply, that is group actions are more extreme than individual actions conditional on both states. [14] In particular, group actions are more extreme on average conditional on both states. If the action function is convex, i.e. agents are biased towards low actions, then the group's actions are expected to be higher conditional on the high state, but not necessarily lower conditional on the low state (the converse holds for a concave action function). This result can be related to the severity shift of Schkade, Sunstein and Kahneman (2000) presented in section 2. When subjects are asked to evaluate the amount of punitive awards, the upward shift in group decisions (relative to individual decisions) are much stronger than the downward shift. As argued, in example 1, it may sensible to assume that subjects prefer underestimating rather than overestimating the amount of punitive awards and as result have a convex action function. The decisions of group of subjects then become more visibly extreme when the defendant is culpable than when he is innocent.

If symmetry cannot be assumed, then the conclusions of lemma 2 must be weakened. An ex ante improvement in information does not guaranty that information has improved conditional on both states as example 4 shows.

---

[14]Note that the greater extremity of group actions is not expressed by first order stochastic dominance: the reason is that conditional on the high state, the group may sometimes play actions lower than any action potentially played by the individual. Intuitively, since the group is more confident in his information, it sometimes makes bigger mistakes than the individual could do.

Example 4. Assume that there are two states 0 and 1 and three signals l, m and h. The states are equally likely ex ante. The information structure $\mathcal{I}_I$ is described by the following table of conditional probabilities:

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

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

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 not inferior when $\theta = 0$. Ex ante, the group's more precise information given 1 compensates for lack of precision given 0.

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 $\int_0^{.5} \Gamma_G(z|\theta_0)dz < \int_0^{.5} \Gamma_I(z|\theta_0)dz$. The conclusions of lemma 2 do not hold conditional on θ_0.

If the information structures cannot be assumed symmetric, then the following weaker implications on conditional distributions of posteriors still holds:

Lemma 3. *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, then:*

- *if $\Gamma_G(q|\theta_0) < \Gamma_I(q|\theta_0)$ for some $q \in [0,1]$, then there exists $q' < q$ such that $\Gamma_G(q'|\theta_0) > \Gamma_I(q'|\theta_0)$.*

- *if $\Gamma_G(q|\theta_1) > \Gamma_I(q|\theta_1)$ for some $q \in [0,1]$, then there exists $q' > q$ such that $\Gamma_G(q'|\theta_1) < \Gamma_I(q'|\theta_1)$.*

Lemma 3 states that, conditional on the high state, the upward tail of the group's distribution of beliefs is fatter than the individual's. 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 though, so it does not guaranty polarization.

Proposition 6. *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:*

- *if $\Lambda_G(a|\theta_0) < \Lambda_I(a|\theta_0)$ for some $a \in A$, then there exists $a' < a$ such that $\Lambda_G(a'|\theta_0) > \Lambda_I(a'|\theta_0)$.*

- *if $\Lambda_G(a|\theta_1) > \Lambda_I(a|\theta_1)$ for some $a \in A$, then there exists $a' > a$ such that $\Lambda_G(a'|\theta_1) < \Lambda_I(a'|\theta_1)$.*

If actions are not necessarily strictly increasing with respect to beliefs, then proposition 6 does not hold. This is illustrated in example 5.

Example 5. Consider an problem with two states and two actions. One decision maker receives noisy information while the other receives no information at all. Suppose that the utility function is such that it is optimal for the uniformed agent to play the low action. Then conditional on the low state, the distribution of actions of the uninformed agent is lower in any sensible sense.

6   Alternative Theories

Group polarization has been extensively studied in the past 50 years, so that several different accounts of the phenomenon have been proposed. We distinguish theories based on information sharing and theories based on preference aggregation. The decision of a group can be either seen as the outcome of the aggregation of its members' preferences or the combination of its members' information. During the group's discussion, subjects exchange their points of view. Presenting one's point of view to the group can sensibly be interpreted as a preference statement or information sharing.

If the preference approach is retained, then the decision of the group results from a compromise between fixed individual positions. To the best of our knowledge, there are two different preference based account of group polarization. The first relates polarization to the social decision rule that groups

use. While group polarization is usually measured using the mean of individual and group distributions of actions, there is no reason to expect a group to choose the mean of its members' decisions. In particular if groups adopt pluralistic voting, the group's decision will be the median of its members' responses (assuming that preferences are single-peaked). A case then generates a shift of group decisions toward the risky alternative if the median of the distribution of individual responses is more risky than its mean. [15]

An alternative account due to Eliaz, Ray and Razin (2006) relates group polarization (in the problems involving decisions between a risky and a cautious alternative) to the possibility that an individual reveals a different preference in a group than in isolation. 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. 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.

Existing evidence suggests that group polarization can only be partially 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 as compared to the initial individual decisions. So the group discussion seems to bring some kind of relevant information to group members which produces a polarization of individual opinions. [16]

---

[15]More generally, the analysis of the social decision rule used by a group is referred to as social decision scheme theory (Davis et al., 1974)

[16]Actually, if one wants to check whether part of the group polarization phenomenon can be

The alternative set of theories is therefore based on the assumption that group polarization results from the information subjects obtain during discussion. The two major psychological theories of group polarization (as well as ours) belong to this class.

According to social comparison theory, individuals evaluate their recommendations relative to a norm of behavior that is reflected in the recommendations of others. For a given choice 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.[17] Individuals make their original, pre-deliberation recommendation according to their prior perception of the ideal choice. During deliberations, 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. Social comparison theory has not been formalized, and its description may lead to different modeling approach. A possibility amounts to see social comparison theory as a particular case of our informational approach. Assume that the state of nature captures a social norm. Every individual receives private information about this norm (his opinion). A group has a more precise idea of what the norm is and group polarization arises as described in this paper.

Persuasive Argument Theory (PAT) (Silverthorne, 1971; Vinokur and Burstein, 1974) requires a discussion between group members and is based on the observation of arguments exchanged during discussion. In a choice dilemma questionnaire, a subject determines his attitudes by thinking about reasons (arguments) favoring either alternatives. A response then reflects the balance between conflicting reasons. During group discussion, subjects exchange arguments and an individual's attitude is affected by arguments that are novel to this individual. Systematic group polarization occurs in some decision prob-

---

explained by a preference based account, one needs to compare the individual decisions after the information exchange to the group decision. Stoner (1968) notes that groups' decisions polarize slightly more than post deliberation individuals' decisions.

[17]Brown (1986) 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."

lems because arguments supporting one alternative are more numerous than those favoring the other. Each subject randomly draws a subset of k arguments. Since arguments favoring the risky alternative are more numerous, subjects favor the risky alternative on average. During the deliberation, an individual receives on average more novel arguments favoring the risky alternative than arguments favoring the safe one. A risky shift naturally obtains on average. Persuasive argument theory is very close in spirit to our account of group polarization. We thought of arguments as statistical signals, and relate group polarization to the pooling of those signals within the group.

7   Conclusion

This paper derives conditions under which a well informed group's decisions polarize as compared to a poorly informed individual's in a monotone decision problem. We show that polarization is likely to occur in problems where decision makers try to hit a target and incur losses that only depends on the distance between the action and the target. When payoffs bias decisions upward or downward, we provide a (weaker) sense in which the group's actions polarize.

Our results point out that polarization need not be inconsistent with the assumption that groups make better decisions on average. Nevertheless, the existing experiments do not yield a clear normative benchmark for evaluating decisions so that it is difficult to draw any conclusions. Glaeser and Sunstein (2009) discuss many situations where agents seem to be prone to systematic biases in the way they revise their opinions during a discussion. It may be the case for instance that groups fail in accounting for the possible correlations in their members' information so that the resulting polarization it too strong. We therefore believe that the framework presented here may be useful as a benchmark for deriving possible evidence of systematic group failures.

References

Athey, Susan, and Jonathan Levin. 2001. "The Value of Information in Mono-

tone Decision Problems." Stanford University, Department of Economics Working Papers 01003. 1, 3, 12, 13, 28, 31

Blackwell, David. 1953. "Equivalent Comparisons of Experiments." *The Annals of Mathematical Statistics*, 24(2): pp. 265–272. 3

Brown, Roger. 1986. *Social Psychology, the Second Edition.* The Free Press. 6, 24

Davis, James H., Norbert Kerr, Mario Sussmann, and A. Kent Rissman. 1974. "Social decision schemes under risk." *Journal of Personality and Social Psychology*, 30(2): 248 – 271. 23

Eliaz, Kfir, Debraj Ray, and Ronny Razin. 2006. "Choice Shifts in Groups: A Decision-Theoretic Basis." *The American Economic Review*, 96(4): pp. 1321–1332. 23

Fraser, Colin, Celia Gouge, and Michael Billig. 1971. "Risky shifts, cautious shifts, and group polarization." *European Journal of Social Psychology*, 1(1): 7–30. 6

Glaeser, Edward L., and Cass R. Sunstein. 2009. "Extremism and Social Learning." *Journal of Legal Analysis*, 1(1): 263–324. 25

Jewitt, Ian. 2007. "Information Order in Decision and Agency Problems." Nuffield College Working Papers. 30

Kaplan, Martin F. 1977. "Discussion Polarization Effects in a Modified Jury Decision Paradigm: Informational Influences." *Sociometry*, 40(3): pp. 262–271. 1, 8

Lehmann, E. L. 1988. "Comparing Location Experiments." *The Annals of Statistics*, 16(2): pp. 521–533. 9, 30

Levy, Haim, and Yoram Kroll. 1978. "Ordering Uncertain Options with Borrowing and Lending." *The Journal of Finance*, 33(2): 553–574. 28

Myers, David G., and Martin F. Kaplan. 1976. "Group-Induced Polarization in Simulated Juries." *Personality and Social Psychology Bulletin*, 2(1): 63–66. 1

Myers, David G., Sandra Brown Wojcicki, and Bobette S. Aardema. 1977. "Attitude Comparison: Is There Ever a Bandwagon Effect?1." *Journal of Applied Social Psychology*, 7(4): 341–347. 7

Quah, John K.-H., and Bruno Strulovici. 2009. "Comparative Statics, Informativeness, and the Interval Dominance Order." *Econometrica*, 77(6): 1949–1992. 10

Rothschild, Michael, and Joseph E Stiglitz. 1970. "Increasing risk: I. A definition." *Journal of Economic Theory*, 2(3): 225 – 243. 3, 13

Schkade, David, Cass R. Sunstein, and Daniel Kahneman. 2000. "Deliberating about Dollars: The Severity Shift." *Columbia Law Review*, 100(4): pp. 1139–1175. 1, 8, 9, 14, 20

Shaked, Moshe, and George Shanthikumar. 2007. *Stochastic Orders and their Applications. Springer Series in Statistics*, Springer. 28, 32

Silverthorne, Colin P. 1971. "Information input and the group shift phenomenon in risk taking." *Journal of Personality and Social Psychology*, 20(3): 456 – 461. 7, 24

Sobel, Joel. 2012. "On the Relationship between Individual and Group Decisions." UC San Diego Working Papers. 2, 4, 5

Stoner, James. 1961. "A comparison of individual and group decisions involving risk." MIT, School of Industrial Management Unpublished Master's Thesis. 1

Stoner, James A.F. 1968. "Risky and cautious shifts in group decisions: The influence of widely held values." *Journal of Experimental Social Psychology*, 4(4): 442 – 459. 5, 6, 24

Teger, Allan I., and Dean G. Pruitt. 1967. "Components of group risk taking." *Journal of Experimental Social Psychology*, 3(2): 189 – 205. 6

Vinokur, Amiram, and Eugene Burstein. 1974. "Effects of partially shared persuasive arguments on group-induced shifts: A group-problem-solving approach." *Journal of Personality and Social Psychology*, 29(3): 305 – 315. 7, 24

Appendix

Proof of Proposition 1. Given Athey and Levin (2001)'s characterization of Lehmann criterion, $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if: For all $z \in [0,1]$, $P_I(\cdot \mid \mathcal{D}_I \leq z) \succ_{MLR} P_G(\cdot \mid \mathcal{D}_G \leq z)$. If δ is non-decreasing, then 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 $\int_0^z \delta(P_I(\mathcal{D}_I^{-1}(p)))dp \geq \int_0^z \delta(P_G(\mathcal{D}_G^{-1}(p))dp$ where $\delta(P_I(\mathcal{D}_I^{-1}(p)))$ and $\delta(P_G(\mathcal{D}_G^{-1}(p))$ are the p-percentile of the distributions of actions induced by $\mathcal{I}_I$ and $\mathcal{I}_G$ respectively. Levy and Kroll (1978) [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 where $\Lambda(a) = \sup\{p \in [0,1] \;:\; a^*(\mathcal{D}^{-1}(p)) \leq a\}$.

Proof of Proposition 2 Given that the state space has two elements, saying that $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ is equivalent to saying that the information structures are ranked according to Blackwell's criterion, i.e. that $\tilde{P}_G$ is more dispersed than $\tilde{P}_I$ in the convex order.

This last condition can be rewritten as (Shaked and Shanthikumar, 2007, chapter 3): for all convex function ψ, $E(\psi(\tilde{P}_G)) \geq E(\psi(\tilde{P}_I))$.

If δ is concave (and increasing), then it is true $\psi \circ \delta$ is decreasing convex for all ψ decreasing convex. Since it is true in particular, that for all decreasing convex function ψ, $E(\psi(\tilde{P}_G)) \geq E(\psi(\tilde{P}_I))$, it follows that for all decreasing convex function ψ, $E(\psi(\delta(\tilde{P}_G))) \geq E(\psi(\delta(\tilde{P}_I)))$. Therefore $\tilde{A}_G$ dominates $\tilde{A}_I$ in the decreasing convex order, i.e. $\int_{a \leq \tilde{a}} \Lambda_G(a)da \geq \int_{a \leq \tilde{a}} \Lambda_I(a)da$ for all $\tilde{a} \in A$.

If δ is convex (and increasing), then it is true $\psi \circ \delta$ is increasing convex for all ψ increasing convex. Since it is true in particular, that for all increasing convex function ψ, $E(\psi(\tilde{P}_G)) \geq E(\psi(\tilde{P}_I))$, it follows that for all increasing convex function ψ, $E(\psi(\delta(\tilde{P}_G))) \geq E(\psi(\delta(\tilde{P}_I)))$. Therefore $\tilde{A}_G$ dominates $\tilde{A}_I$ in the increasing convex order, i.e. $\int_{a \geq \tilde{a}} \Lambda_G(a)da \leq \int_{a \geq \tilde{a}} \Lambda_I(a)da$ for all $\tilde{a} \in A$.

Proof of Proposition 3. We prove the first statement. The second can be proven symmetrically. The proof only holds for discrete information structures, i.e. the set of posteriors induced by an information structure contains a finite number of elements.

Since the group information is more precise than the individual information, it must be the case that $P_G(\cdot|\mathcal{D}_G^{-1}(0))$ is weakly dominated (in the relevant stochastic order) by $P_I(\cdot|\mathcal{D}_I^{-1}(0))$. Since δ is strictly increasing, if the dominance is strict then the minimum group action is lower than the minimum individual action.

Suppose on the other hand that the beliefs are identical, $P_G(\cdot|\mathcal{D}_G^{-1}(0)) = P_I(\cdot|\mathcal{D}_I^{-1}(0))$. Suppose further that $P_G(\cdot|\mathcal{D}_G^{-1}(q))$ and $P_G(\cdot|\mathcal{D}_I^{-1}(q))$ are constant over $[0, q_G)$ and $[0, q_I)$ respectively. This implies that $P_G(\cdot|\mathcal{D}_G^{-1}(q_G))$ dominates $P_G(\cdot|\mathcal{D}_G^{-1}(0))$. Similarly for the individual, we have $P_I(\cdot|\mathcal{D}_I^{-1}(q_I))$ dominates $P_I(\cdot|\mathcal{D}_I^{-1}(0))$. As a result, q_G cannot be strictly smaller than q_I. If it was the case $\int_0^{q_I} P_G(\cdot|\mathcal{D}_G^{-1}(q))dq$ would strictly dominate $\int_0^{q_I} P_I(\cdot|\mathcal{D}_I^{-1}(q))dq$ which is in contradiction with the characterization of the monotone information order provided in lemma 1.

If $q_I < q_G$, then $P_G(\cdot|\mathcal{D}_G^{-1}(q_I))$ is strictly dominated by $P_I(\cdot|\mathcal{D}_I^{-1}(q_I))$. Therefore, $\delta(P_G(\cdot|\mathcal{D}^{-1}(q_I)))$ is strictly lower than $\delta(P_I(\cdot|\mathcal{D}^{-1}(q_I)))$ and for all $q \in [0, q_I)$, $\delta(P_G(\cdot|\mathcal{D}^{-1}(q_I)))$ is equal to $\delta(P_I(\cdot|\mathcal{D}^{-1}(q_I)))$.

Suppose finally that $q_I = q_G$. Then it must be the case that $P_G(\cdot|\mathcal{D}_G^{-1}(q_I))$ is weakly dominated by $P_I(\cdot|\mathcal{D}_I^{-1}(q_I))$ for otherwise, for ϵ arbitrarily small, $\int_0^{q_I+\epsilon} P_G(\cdot|\mathcal{D}_G^{-1}(q))dq$ would strictly dominate $\int_0^{q_I+\epsilon} P_I(\cdot|\mathcal{D}_I^{-1}(q))dq$. Either the dominance is strict in which case $\delta(P_G(\cdot|\mathcal{D}^{-1}(q_I)))$ is strictly lower than $\delta(P_I(\cdot|\mathcal{D}^{-1}(q_I)))$ and for all $q \in [0, q_I)$, $\delta(P_G(\cdot|\mathcal{D}^{-1}(q_I)))$ is equal to $\delta(P_I(\cdot|\mathcal{D}^{-1}(q_I)))$; or the two beliefs are equal in which case one can iterate with the same reasoning.

In order to prove lemma 2, it will be convenient to use lemma 4. Lemma 4 sets two equivalent implications of the MIO on conditional distributions of posteriors when the state space has two elements. Lemma 4 roughly states that a more precise information structure should have conditional distributions of posteriors more distant from each other.

Lemma 4. *If $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$, then for all $q \in [0,1]$,*

$$\Gamma_G^{-1}(\Gamma_I(q|\theta_0)|\theta_0) \leq \Gamma_G^{-1}(\Gamma_I(q|\theta_1)|\theta_1).$$

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

$$\Gamma_G(q'|\theta_0) \leq \Gamma_I(q|\theta_0) \Rightarrow \Gamma_G(q'|\theta_1) \leq \Gamma_I(q|\theta_1)$$

Proof of Lemma 4. The first condition is based on Lehmann (1988): $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if, for all $s \in S_I$,

$$\mathcal{A}_G^{-1}(\mathcal{A}_I(s;\theta_0)|\theta_0) \leq \mathcal{A}_G^{-1}(\mathcal{A}_I(s;\theta_1)|\theta_1)$$

Assuming that $P_i(s)$ is continuous and strictly increasing for $i = I, G$, the above characterization translates to the distribution of beliefs, i.e. for all $q \in [0,1]$,

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

Since $\Gamma_i(q) = \mathcal{A}_i(P_i^{-1}(q))$, the following property on ex post distribution of beliefs obtains:

$$\Gamma_G^{-1}(\Gamma_I(q;\theta_0)|\theta_0) \leq \Gamma_G^{-1}(\Gamma_I(q;\theta_1)|\theta_1) \quad \text{for all} \quad q \in [0,1] \quad (1)$$

The second condition is based on Jewitt (2007): $\mathcal{I}_G$ is more precise than $\mathcal{I}_I$ if and only if for all s and s', $\mathcal{A}_G(s'|\theta_0) \leq \mathcal{A}_I(s|\theta_0)$ implies $\mathcal{A}_G(s'|\theta_1) \leq \mathcal{A}_I(s|\theta_1)$.

Given that beliefs are non-decreasing with respect to signals, this property also holds for conditional distributions of beliefs, i.e. for all $q', q \in [0,1]$

$$\Gamma_G(q'|\theta_0) \leq \Gamma_I(q|\theta_0) \Rightarrow \Gamma_G(q'|\theta_1) \leq \Gamma_I(q|\theta_1)$$

.

Proof of Lemma 2 Note first that symmetry implies that $\alpha(s|\theta_1) = \alpha(-s|\theta_0)$, which in turn implies that $\mathcal{A}(s|\theta_1) = 1 - \mathcal{A}(-s|\theta_0)$. Provided that posteriors are strictly increasing with respect to beliefs, let us denote by $s(q)$ the inverse

of $P(q)$.

We start by showing that $\Gamma(q|\theta_1) = 1 - \Gamma(1 - q|\theta_0)$. Note that $\Gamma(q|\theta_1) = \mathcal{A}(s(q)|\theta_1)$. Since symmetry induces that $s(q) = -s(1 - q)$, it follows that $\mathcal{A}(s(q)|\theta_1) = \mathcal{A}(-s(1 - q)|\theta_1)$ which is also equal to $1 - \mathcal{A}(s(1 - q)|\theta_1)$ and so to $1 - \Gamma(1 - q|\theta_0)$.

Provided that the information structure is symmetric, the second proposition of lemma 4 implies that for $q \in [0, 1]$

$$\Gamma_G(q|\theta_0) \leq \Gamma_I(q|\theta_0) \Leftrightarrow \Gamma_G(1 - q|\theta_0) \geq \Gamma_I(1 - q|\theta_0)$$

that is to say

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

Condition 1 implies that $\Gamma_G(\cdot|\theta_0)$ and $\Gamma_I(\cdot|\theta_0)$ cross at $\tilde{q}$ if and only if $\Gamma_G(\cdot|\theta_1)$ and $\Gamma_I(\cdot|\theta_1)$ cross at $\tilde{q}$.

We now show that if $\Gamma_G(\cdot|\theta_0)$ and $\Gamma_I(\cdot|\theta_0)$ cross at $\tilde{q}$ then

$$\int_0^{\hat{q}} \Gamma_G(z|\theta_0) - \Gamma_I(z|\theta_0)]dz \geq \int_0^{\hat{q}} [\Gamma_G(z|\theta_1) - \Gamma_I(z|\theta_1)]dz. \tag{2}$$

To see this, note first that $\int_0^{\hat{q}} \Gamma_G(z|\theta_k) - \Gamma_I(z|\theta_k)dz = \int_0^{\Gamma_I(\hat{q}|\theta_k)} \Gamma_I^{-1}(u|\theta_k) - \Gamma_G^{-1}(u|\theta_k)du$. A change of variable yields $\int_0^{\hat{q}} q - \Gamma_G^{-1}(\Gamma_I(q|\theta_k)|\theta_k)d\Gamma_I(q|\theta_k)$. Therefore, $\int_0^{\hat{q}} [\Gamma_G(z|\theta_0) - \Gamma_I(z|\theta_0)]dz \geq \int_0^{\hat{q}} [\Gamma_G(z|\theta_1) - \Gamma_I(z|\theta_1)]dz$ can be equivalently rewritten as

$$\int_0^{\hat{q}} q - \Gamma_G^{-1}(\Gamma_I(q|\theta_0)|\theta_0)d\Gamma_I(q|\theta_0) \geq \int_0^{\hat{q}} q - \Gamma_G^{-1}(\Gamma_I(q|\theta_1)|\theta_1)d\Gamma_I(q|\theta_1)$$

which holds since $\Gamma_G^{-1}(\Gamma_I(q; \theta_0)|\theta_0) \leq \Gamma_G^{-1}(\Gamma_I(q; \theta_1)|\theta_1)$ for all $q \in [0, 1]$ is implied by $\mathcal{I}_G$ being more precise than $\mathcal{I}_I$.

Now, Athey and Levin (2001)'s characterization of the MIO order, i.e. $\int_0^q \Gamma_G(z)dz \geq \int_0^q \Gamma_I(z)dz$ for all $q \in [0, 1]$ can be equivalently rewritten as: for all $q \in [0, 1]$,

$$\int_0^q \Gamma_G(z|\theta_0) - \Gamma_I(z|\theta_0)dz \geq \int_0^q \Gamma_I(z|\theta_1) - \Gamma_G(z|\theta_1)dz. \tag{3}$$

Conditions 2 and 3 imply that $\int_0^q \Gamma_G(z|\theta_0) - \Gamma_I(z|\theta_0)dz$ cannot be negative if the two distributions cross at q. Finally, if the above integral is not negative

31

when the two distribution cross, then it is not negative at any point.

Proof of Proposition 5.

Following Shaked and Shanthikumar (2007, chapter 3), the conclusions of lemma 2 can be equivalently written as follows:

- for all increasing concave function ϕ, $E(\phi(\tilde{P}_{G,\theta_0})) \le E(\phi(\tilde{P}_{I,\theta_0}))$

- for all increasing convex function ϕ, $E(\phi(\tilde{P}_{G,\theta_1})) \le E(\phi(\tilde{P}_{I,\theta_1}))$

Since the composition of two increasing convex functions is convex and the combination of two increasing concave functions is concave, it follows that:

- if δ is increasing and concave, then for all increasing concave ϕ, $E(\phi(\delta(\tilde{P}_{G,\theta_0}))) \le E(\phi(\delta(\tilde{P}_{I,\theta_0})))$

- if δ is increasing and convex, then for all increasing convex ϕ, $E(\phi(\delta(\tilde{P}_{G,\theta_1}))) \ge E(\phi(\delta(\tilde{P}_{I,\theta_1})))$

Therefore,

- for all increasing concave function ϕ, $E(\phi(\tilde{A}_{G,\theta_0})) \le E(\phi(\tilde{A}_{I,\theta_0}))$

- for all increasing convex function ϕ, $E(\phi(\tilde{A}_{G,\theta_1})) \le E(\phi(\tilde{A}_{I,\theta_1}))$

Proof of Lemma 3 We prove the first part of lemma 3. The second part can proven symmetrically. Consider condition 7 of lemma 4. Suppose now that $\Gamma_G(\tilde{q}|\theta_0) < \Gamma_I(\tilde{q}|\theta_0)$ for some $\tilde{q} \in [0,1]$ and for all $q \le \tilde{q}$, $\Gamma_G(q|\theta_0) \le \Gamma_I(q|\theta_0)$. Then condition 2 implies that $\Gamma_G(q|\theta_1) \le \Gamma_I(q|\theta_1)$ for all $q \in [0,\tilde{q}]$.

Therefore it cannot be the case that $\pi(\theta_0) \int_0^{\tilde{q}} \Gamma_G(z|\theta_0) - \Gamma_I(z|\theta_0)dz$ be higher than $\pi(\theta_1) \int_0^{\tilde{q}} \Gamma_I(z|\theta_1) - \Gamma_G(z|\theta_1)dz$, as the former term is strictly negative and the latter is positive. So it cannot be the case that $\int_0^{\tilde{q}} \Gamma_G(z)dz \ge \int_0^{\tilde{q}} \Gamma_I(z)dz$.