The Value of Information for Groups

Philippos Louis*

November, 2012

Abstract

The paper examines the possibility of comparing different information structures in terms of informativeness, in the context of collective decision making. We set up a fairly general model of collective decision making through voting. Attention is restricted to groups in which members share a common objective. An information structure's informativeness is measured by the expected aggregate value it offers the group. Our first result shows that a comparison is possible in some cases, for any such like-minded group and any possible voting rule. Still, we show that the instances where such comparisons are possible are very limited. The set of information structures that can be compared is extended if one poses restrictions upon the profile of group members preferences or the voting rule. We apply some of our results to a model in which information is endogenous.

*Chair for Organizational Design, Department of Economics, University of Zürich, Blümlisalpstrasse 10, CH-8006, Zürich, Switzerland. Email: filippos.louis@econ.uzh.ch.
I gratefully acknowledge financial support from the Spanish Ministry of Science and Innovation through grant Consolidated Group-C ECO2008-04756 and FEDER, as well as from the Swiss National Science Foundation (SNSF 135135) and the European Research Council (ERC Advanced Investigator Grant, ESEI-249433). I wish to thank Joan de Martí, Dimitris Xefteris, Riste Gjorgjiev, Salvador Barberà, Guillaume Haeringer, Matthias Messner, Larbi Alaoui and participants in the PhD semniar in Microeconomics at the University of Zürich for helpful comments and suggestions.

1 Introduction

Juries, FDA committees, hiring committees are all examples of groups of individuals that must make a decision: to convict or to acquit, to approve a drug or not, to hire a particular candidate. Furthermore, all members of these groups are *like-minded* because they share a common interest. If they had perfect information about the problem at hand they would all agree on what decision the group should take: acquit the innocent, reject a dangerous drug, hire an appropiate candidate for the job. Disagreement may arise because of imperfections in the information that the group members hold.

Groups can usually choose between a number of different sources of information that can reduce uncertainty about the outcome of their decision. Juries can hear different witnesses or admit specific evidence; an FDA committee can choose from an array of clinical trials for a drug; a hiring committee can ask for references from different sources, or have the candidate take specific tests. Given limited resources, these like-minded groups often face the choice of a specific source of information over another. Answering the question "what information source is better for the group?" is not straightforward. In this paper we show why not and under what conditions one can give a clear answer to that question.

It is well known that such an answer is not straightforward even for individual decision makers. They can not be hurt by any additional information, but the seminal work of Blackwell (1951) [5] shows that one cannot always rank two alternative sources of information. They can be ranked when one source A is equivalent to source B plus some noise. In that case source B is preferred because it is, unambiguously, a more precise statistic of the state of the world than A. But, in general, which one is better depends on the decision problem at hand. Subsequent literature has tried to provide partial answers by looking at specific families of decision problems. We further discuss these studies in the literature review section.

When one looks at the problem from the point of view of a group, the problem of aggregating preferences conflicts with the comparison of informativeness provided by different sources. The focus of this paper is the understanding and resolution of these conflicting interests from the group's point of view. Therefore, we set up a collective choice problem, general enough to encompass different economic situations. The group faces a binary choice problem. The collective choice problem is resolved through voting. All information is public. In such a model and with preferences being common knowledge there is no scope for strategic manipulation of votes. What makes the problem of ranking preferences over information harder for a group? The following example illustrates the relevant difficulties.

Example 1. *Consider a hiring committee comprised by two members: Anne and Bob. They face the choice of whether or not to hire a candidate for a job. The candidate may, or may not, be a good fit for the job. Any of the two possibilities is equally likely. The table in figure 1 gives Anne and Bob's valuations of any possible outcome of their decision.*

(Anne, Bob)		Candidates condition	
		Fit for job	Not fit
Committee's decision	Hire	$(1,1)$	$(0,0)$
	Do not hire	$(0,0)$	$\left(\frac{1}{4},\frac{2}{3}\right)$

Figure 1: Anne's and Bob's valuation of possible outcomes

Both members of the committee would agree on the best decision if they knew whether the candidate is appropiate for the job or not. Note also that while both agree on their valuation of making a wrong decision in either case, they also agree that the opportunity cost of not hiring a good candidate is higher than the opportunity cost of hiring an unfit candidate. That is, they both have a bias towards hiring. The bias is higher in Anne's case. Let us assume that unanimity is required in order to hire the candidate.

Before taking a decision they have the choice to either have somebody interview the candidate or have him take a test. Each of these procedures can give some additional information. To keep the example simple, suppose that the outcome of both procedures can be deduced to a binary noisy signal, in the form of a recommendation: 'hire' or 'don't hire'. This recommendation is public: the result of the interview or the test is common knowledge for both individuals. The tables in the following figures show the likelihood of each recommendation in each possible case, for each one of the two procedures.

Likelihood of recomendation		Candidates condition	
		Fit for job	Not fit
Recomendation	Hire	$\frac{1}{2}$	$\frac{1}{3}$
	Do not hire	$\frac{1}{2}$	$\frac{2}{3}$

Figure 2: Interview

Likelihood of recomendation		Candidates condition	
		Fit for job	Not fit
Recomendation	Hire	$\frac{3}{4}$	$\frac{1}{2}$
	Do not hire	$\frac{1}{4}$	$\frac{1}{2}$

Figure 3: Test

The question we examine is which of the two procedures gives a higher expected value to the committee. After receiving a recommendation, Anne and Bob update their beliefs about the candidate's fitness for the job. They use their posterior to calculate their expected value from each of the two possible actions and vote for the action that gives them a higher expected value. It turns out that following an interview, both Anne and Bob would vote to hire the candidate. That is, a negative recommendation from an interview is not strong enough to overcome their bias towards hiring the candidate. On the other hand, if instead of an interview they use a test, Bob would vote according to the recommendation while Anne would still vote for hiring the candidate, regardless of the recommendation. Given that unanimity is required for a hire, the committee always hires the candidate after an interview, but only hires him if this is recommended after a test. We can thus calculate the expected social welfare for the committee after each of the two procedures:

$$W(interview) = \frac{1}{2}(1+1) + \frac{1}{2}\cdot 0 = 1$$
$$W(test) = \frac{1}{2}\left[\frac{3}{4}(1+1) + \frac{1}{4}\cdot 0\right] + \frac{1}{2}\left[\frac{1}{2}\cdot 0 + \frac{1}{2}\left(\frac{1}{4}+\frac{2}{3}\right)\right] = \frac{43}{48}$$

Social welfare is higher when using the interview.

It is interesting to notice that the test is actually a more precise procedure. The posterior belief after a recommendation to hire is the same in both procedures. But a recommendation not to hire from the test gives a higher posterior belief that the candidate is not fit, compared to the same recommendation from an interview. This means that less precise public information has a higher social value.

In this example it is easy to see that this happens because the less precise recommendation does not affect the committee's decision either way: Anne's and Bob's bias is too strong and the recommendation is too weak. The strong recommendation from the interview may affect the vote of Bob but not the vote of Anne, since she is more biased. Still, given the decision rule here, this means that with the test, the final outcome may change, Anne does not want.

We do not examine how Anne and Bob decide whether to use an interview or a test. We just make the comparison between the expected value generated by each of these procedures.

In this example we perform the exercise of comparing two different sources of public information for a particular group. It is always possible to perform such a comparison for a given pair of information sources, a particular group and a particular decision rule. In the remainder of the paper we show when such a comparison is possible for different combinations of individuals valuations and decision rules. As in the example we focus on situations with a binary choice and two possible states of nature. Unlike the example, we do not limit the analysis to binary signals but allow for the public signal to have any general form. Groups are allowed to have any possible number of members (even infinite). We only restrict individual valuations to be such that under perfect information all individuals would agree on the optimal choice.

1.1 Results

We show first that it is possible to establish a partial order on information sources such that a particular source gives a higher value than another, for any possible profile of individual valuations and any possible voting rule. Alas, the cases where such an order is possible are limited. We show for instance that the only binary signal that dominates other binary signals in such a way is the perfectly revealing signal.

The partial order can be extended if one restricts the domain of possible individual valuations. In particular, if we restrict attention to groups that a priori receive the same expected value from both alternatives, it is possible to establish an extended partial order on information sources, for any possible voting rule. We give examples of parametric families of such sources that can be ordered according to some parameter. These include power distribution functions and exponential functions.

For any group, there may exist an optimal voting rule that maximizes social welfare. Such an optimal rule pinpoints a particular individual within the group as a decisive voter. If there exists an ordering on information sources according to the expected value they give to the decisive voter, then this order holds for the group if the optimal voting rule is applied.

Finally we use some of our results in an application of collective choice design. We look at a case where the group may obtain information before making a decision. This depends stochastically on how many of the group members want the group to receive this public information. We show how this demand for information depends on the voting rule used and we characterize the rule that maximizes the expected social value for the group. We then use the fact that some information structures can be ordered for any voting rule and perform comparative statics. It turns out that the more informative the available information structure, the more conservative is the optimal voting rule.

Throughout our analysis we assume that individual values from different outcomes represent the individuals' true valuations. This does not affect our analysis since we only perform comparisons of information structures without considering any mechanism for choosing a particular one among the available alternatives. For the cases covered in our first result, individuals would have no incentive to misrepresent their valuations: when information sources can be ordered in such a way, all group members agree with the group's ordering. When this is not possible, individuals could have incentives to misrepresent their valuations if this could affect the choice of information source. Such manipulation would not make any sense in order to affect the group's final choice: given binary choices and the type of preferences we consider, voting is a strategy-proof method. By misrepresenting their valuations the agents cannot obtain a better outcome.

1.2 Literature review

The seminal contribution to the literature on information structure comparisons is Blackwell (1951) [5]. There, an information structure is more informative than another if it is preferred by any decision maker for any possible decision problem.This strong condition induces a partial order on information structures. A more comprehensive order is offered by Lehmann (1988) [17], who focuses on specific decision problems that are monotone, and on information structures that satisfy a monotone likelihood condition. Persico (2000) [21], Athey and Levin (2001) [1] and Jewitt (2007) [16] move in the same direction and extend Lehmann's ordering for more general classes of monotone problems. In a recent contribution, Cabrales, Gossner and Serrano (2012) [6] provide a complete order on information structures, based on a measure of entropy in the decision maker's beliefs, in a class of investment problems. Ganuza and Penalva (2010) [9] take a different approach than these papers. They provide an ordering that is not based on any class of decision problems. Instead, they order different information structures in terms of the variability of conditional expectations they generate. They use this order to study the incentives of an auctioneer to disclose information. Similarly to all these papers, the present one shares the aim of comparing different information structures. It differs in doing so from the point of view of a group of agents instead of a single decision maker.

A series of papers by Gersbach (1991 [11], 1992 [12], 1993 [13], 2000 [14]) study the value of public information for groups that face a collective choice problem. Through different examples it is shown that it is possible for public information to be harmful for even a majority of voters. There are important differences between our work and Gersbach's contribution. Gersbach's approach is valid for a more general family of collective choice problems, with multiple alternatives and no restriction on preferences, but it only considers comparisons between two extreme cases: perfect uncerainty or perfect information. Our institutional restrictions in the choice problem allow us to compare situations with different degrees of uncertainty

Messner and Polborn (2012) [19] and Strulovici (2010) [24] are two papers that look at the attitude towards experimentation of groups of individuals, that make decisions collectively. In both of them the setup is dynamic. In the first one, the group faces the option to take a decision immediately, or to wait to obtain more information. In the second one, the group decides through voting whether it wants to continue experimenting with a particular policy. Continuous experimentation allows voters to learn about the policy's effects on their welfare. Both papers focus on the choice of the collective decision making rule and on how it determines the degree to which group members learn. Similarly, one way to look at what we do in this paper is to think that the group is engaged in a one stage experimentation game in which experimentation can take different forms and we look at which type of experimentation offers the highest value to the group.

The study of collective choice by groups of individuals that share a common goal, in environments with incomplete information, goes back to Condorcet in the 18th century and was studied later as well by Marshak and Radner (1972) [18] in their theory of teams. More recently, Austen-Smith and Banks (1996) [2] showed that strategic considerations may not allow the correct aggregation of information in such groups. Feddersen and Pesendorfer (1997) [8] show that if the size of the group goes to infinity, information is correctly aggregated. Persico (2004) [22], Gerardi and Yariv (2008) [10] and Gershkov and Szentes (2009) [15] study how the incentives of group members to acquire information depend on the design of the decision mechanism. Bergemann and Välimaki (2005) [4] survey the literature on information acquisition in the context of committees and other mechanism design problems. In all of this line of the literature, any information that the agents have or may acquire is a priori private. This gives rise to particular strategic considerations on their part when making decisions on whether or not to acquire information, on how they communicate with others or on how to vote. All these are absent in our setting: attitudes towards information depend strictly on individuals' valuations and the design of the decision process, not on the possible existence of any private information.

The possible value of public information to a set of individuals is studied in Morris and Shin (2002) [20]. In their setting, individuals are involved in a game where actions have strategic complementarities. They find that more public information is socially beneficial when agents have no private information. When this is not true, more public information may hurt society. In their paper decision making is decentralized. Public information can help agents make better decisions but may also serve as a coordination device. In our setting, decision making is centralized. Public information's value lies solely in its instrumental function in improving decision making by reducing uncertainty about the state of the world.

2 The model.

In this section we setup a model of collective choice. Any results we obtain concerning the possibility of ordering information sources refer to this model.

2.1 Like-minded groups.

Consider a set of agents I (possibly infinite) that have to choose, jointly, between two possible collective actions, $x \in X = \{0, 1\}$. These may represent, for example, two alternative policies, or two different candidates for a post. We may refer to these simply as the "low" and "high action". There is an unknown state of nature that can have two possible values, $\theta \in \{0, 1\}$. We may also refer to these as the "low" and "high" state respectively. Agents share a common prior regarding the state of nature. Let π represent the ex-ante probability agents assign to the state of nature being high: $\pi = Pr(\theta = 1)$.

Individual valuations depend both on the collective action x and the state of nature θ and are given by a function $u_i(x, \theta, t_i)$. Notice that we only assume that all individuals' utility functions depend on the same variables, namely x, θ and an idiosyncratic parameter $t \in T \subseteq \mathbb{R}$. This parameter represents the possible bias of an individual towards either one of the actions the group may take. In particular, let $\lambda_i(t) = \frac{u_i(1,1,t)-u_i(0,1,t)}{u_i(0,0,t)-u_i(1,0,t)}$ be an agent's *bias function*. That is the ratio of the opportunity cost from choosing the low action in the high state, over the opportunity cost from choosing the high action in the low state. We make the following assumption:

Assumption 1. The bias function $\lambda_i(t)$ increasing in t and
for $t > t'$,

$$\lambda_i(t) > \lambda_j(t')\,,\ \forall i, j \in I$$

What this says is that individuals of a higher type are more biased in favor of the high action. Furthermore, it states that it is possible to order individuals in terms of their bias simply by knowing their type t and without any further information about their valuations. We take advantage of this fact to lighten notation and we from now on refer to the bias of individual i simply as $\lambda(t_i)$, omitting the subscript for λ.

It is helpful for our analysis to assume that agents are distributed over the type space T following a distribution $\xi(t)$ which may be either continuous or discrete. All agents have equal mass. The total mass of agents is normalized to 1: $\int_T \xi(t)dt = 1$. In the case of a discrete distribution the integral should be substituted by a summation[1].

We shall further assume that all agents together form what we name a *like-minded group (LMG)*[2].

Definition 1. *A set of individuals I is a like-minded group if:*

$$u_i(\theta, \theta, t_i) > u_i(1 - \theta, \theta, t_i),$$

$$\forall\ \theta \in \{0, 1\},\ t_i \in T, i \in I$$

Assumption 2. I is a like-minded group.

As can be seen from the definition, all members of a LMG agree on what the best action is in each state of nature. This does not mean that members of a LMG always agree on what action the group should take. Given the uncertainty about the state of nature there may be disagreement resulting from individual biases in favor of the higher or lower action.

Tables 4 and 5 show two examples of 2-member groups. The numbers in the cells represent the value for each group member given the action in a specific state. Group 1 is not a LMG. Agents disagree on the optimal choice in the high state. Group 2 is a LMG. Assuming the numbers represent the values of agents 1 and 2 respectively we can compute $\lambda_1(t_1) = 1 > \frac{1}{4} = \lambda_2(t_2)$, which means $t_1 > t_2$: agent 2 is biased towards the low action.

	$\theta = 0$	$\theta = 1$
$x = 0$	10, 10	10,8
$x = 1$	8, 8	8,10

Figure 4: Group 1 is not a LMG

	$\theta = 0$	$\theta = 1$
$x = 0$	1, 100	0, -10
$x = 1$	0, 20	1,10

Figure 5: Group 2 is a LMG

[1]Note that $xi(t)$ is not a probability distribution and the fact that $\int_T \xi(t)dt = 1$ is just a normalization. In our analysis we assume that agents' types are common knowledge

[2]The term 'committee' is often used in the literature to describe groups with such preferences. But besides preferences, the term also has connotations of relatively small groups and 'committee members' are often assumed to possess private information which they are expected to aggregate. In our model groups may be of any size and we assume there is no private information. We therefore prefer this alternative term.

2.2 Information.

The group receives a public signal $s \in S \subseteq \mathbb{R}$ (common to all individuals) about the state of nature before taking a decision. In particular, the signal s is distributed according to a cumulative distribution function $F_\theta(s)$ on the set S of possible signals. From now on we refer to either the information structure (the pair $\{F_0(s), F_1(s)\}$) or the distribution (the unconditional distribution $F(s;\pi)$) interchangeably or simply by F. We use $f_\theta(s)$ to denote the probability density function for continuous signals and use the same notation for discrete signals, implying $f_\theta(s) = Pr(s|\theta)$ for such signals. We assume that the distribution satisfies the *monotone likelihood ratio property (MLRP)*.

Assumption 3. $\frac{f_1(s)}{f_0(s)} \geq \frac{f_1(s')}{f_0(s')} \Leftrightarrow s > s'$.

In other words, higher public signals imply that the high state of nature $\theta = 1$ is more likely.

For the moment we make no further assumptions on S. While we use integrals over subsets of S in the analysis that follows, and unless mentioned otherwise, results also hold in the case of a discrete signal space and proofs can be obtained by substituting integrals with summations.

2.3 Individual and Collective Choice.

2.3.1 Individual choice

Before setting up our model of collective choice it is useful to understand how individuals behave in such a setup. Or, in different words, by looking at "groups" of a single individual.

Let $\phi_i : S \to X$ be a decision rule for an individual agent i. Given a decision rule ϕ_i and an information structure F, an agent's ex ante expected value is:

$$\mathcal{U}_i(\phi_i, F, t_i) = \pi \int_{s \in S} u_i(\phi_i(s), 1, t_i) f_1(s) ds + (1-\pi) \int_{s \in S} u_i(\phi_i(s), 0, t_i) f_0(s) ds$$

Now let $\hat{\phi}_i : S \times T \to X$ be the optimal decision rule i.e. the one that maximizes the individuals expected payoff given the public signal. It can be easily seen that the *MLRP* on the information structure implies that $\hat{\phi}_i(s, t_i)$ is a threshold function.

Lemma 1. *There either exists a threshold $\tilde{s}_i(t_i)$ such that:*

$$\hat{\phi}_i(s, t_i) = \begin{cases} 1, & s > \tilde{s}_i(t_i) \\ 0, & \textit{otherwise} \end{cases}$$

or $\hat{\phi}_i(s, t_i)$ is constant.

Proof. All proofs of lemmas and propositions can be found in the appendix. □

In particular, the threshold $\tilde{s}(t)$ is defined as follows:

$$\tilde{s}(t) = \max\left\{ s \ : \ \frac{f_1(s)}{f_0(s)} \frac{\pi}{1-\pi} \lambda(t) \leq 1 \right\} \tag{1}$$

Furthermore, this threshold is decreasing in the agent's type.

Lemma 2.

$$\tilde{s}(t) \geq \tilde{s}(t') \quad for \quad t < t'.$$

At this point we make one last assumption on the set of possible signals that comes with no loss of generality, but makes sure that the threshold is always well defined:

Assumption 4. There exists $\underline{s} \in S$, with $\underline{s} < s$, $\forall\ s \in S \setminus \{\underline{s}\}$, and $F(\underline{s}) = 0$. We further assume $\underline{s} = -\infty$ if and only if $(-\infty, s') \in S$ for some $s' \in \mathbb{R}$

For the analysis that follows it may be useful to remark that given the way we define the threshold, $\tilde{s}(t_i)$ is the highest value of s such that i still prefers $x = 0$ over $x = 1$. This is particularly important in the case of discrete signals.

2.3.2 Collective choice

The group makes decisions using a voting rule in the following way: Action $x = 0$ is taken if a proportion of at least $q \in [0, 1]$ members of the group agree to take this action. Members of the group agree to take an action if it is the optimal decision they would take as individual decision makers[3]. Therefore it is given by the function $\hat{\phi}_i(s, t_i)$. Thus, the group's decision function is:

$$x(q, s) = \begin{cases} 1, & \int_T \hat{\phi}(s, t)\xi(t)dt < q \\ 0, & otherwise \end{cases}$$

Given the preferences, it is easily seen that after receiving the public signal the group divides into two ordered subgroups.

[3]As was mentioned in the introduction, given preferences individual preferences, voting under any q-rule is strategy proof. This means that voting for one's preferred choice is a dominant strategy.

Lemma 3. *Given a public signal $\hat{s}$, there exists $\tilde{t}(\hat{s})$ such that:*

$$\hat{\phi}(\hat{s}, t) = \begin{cases} 1, & t > \tilde{t}(\hat{s}) \\ 0, & otherwise \end{cases}$$

Furthermore, $\tilde{t}(s)$ is decreasing in s.

Some individuals of a low type with a threshold $\tilde{s}$ above the received signal consider $x = 0$ optimal. The rest of the group, that is higher type individuals, consider $x = 1$ as optimal. The group's final decision will depend on whether the mass of the first of these subgroups is larger than the quota q. If it is, then according to the decision rule, the group takes action $x = 0$. If not, it takes action $x = 1$. If the voting rule is simple majority we know that the group's decision always coincides with the vote of the median voter. The following lemma generalizes this idea for all possible q-rules.

Lemma 4. *The decisive type is the policy type $t_d(q)$ that satisfies:*

$$\int_{(-\infty, t_d)} \xi(t)dt < q \quad and \int_{(-\infty, t_d]} \xi(t)dt \geq q$$

The group's decision function can be written as follows:

$$x(q, s) = \begin{cases} 1, & t_d(q) > \tilde{t}(s) \\ 0, & otherwise \end{cases}$$

It is easy to see why this result holds: to know whether the group takes action 0 or 1 we just have to know whether an individual with the decisive type t_d prefers 0 or 1. Given lemma 3, we know that if an individual of type t_d prefers $x = 1$, so will all individuals to his right (of a higher type) and following the decision rule, $x(q, s) = 1$. If the individual prefers $x = 0$, then so will all individuals to his left as well, and their mass is more or equal to q and therefore $x(q, s) = 0$. Note that the *decisive type* is defined in a way that depends on the distribution of types for a particular group and is always well defined.

This concludes our modeling of collective choice for LMG's. In the following section we examine whether it is possible to order different distributions with respect to the value they provide to the group in this framework.

3 Comparing Information Structures

The value of information lies in the degree to which it allows an agent to take better decisions. Therefore, when comparing two information structures, say F and G, one sais that F is more informative than G when the expected payoff for the decision maker under F is higher than under G. Making such a comparison for a specific decision problem is not complicated. But when the same comparison is made for a different decision problem the previous ranking in terms of informativeness may not hold. A vast literature in economics and statistics studies the properties of information structures that allow such comparisons and gives orderings that are valid for more or less general families of decision problems. In this section we attempt a similar exercise. Instead of looking at individual decision problems, we focus on a collective choice problem. When is an information structure better for the group than another one?

We attempt to answer this question in terms of the collective choice model setup in the previous section. Still, for this question to have a content one must define the value of information for a group. As in Gersbach (1991 [11]; 1992 [12]; 1993 [13]; 2000 [14]), we choose to follow a utilitarian approach. We consider the sum of individual group members' valuations following the collective choice made under a particular information structure to represent the value of this information for the group.

Definition 2. *A group's expected value from an information structure F is:*

$$\mathcal{V}(\hat{\phi}_d, F, \pi, q) = \int_T \mathcal{U}_i(\hat{\phi}_d, F, \pi, t_i, q)\xi(t_i)dt$$

Having defined the value of information for groups we can now define informativeness for a group. Like in the literature concerning comparisons of information structures for individual decision makers, we consider a particular information structure to be more informative than another if it offers a higher expected value to the decision maker. In our case the decision maker is the group. Formally, we use the following definition:

Definition 3. *We say that F is* more informative for the group *than G and denote it as:*

$$F \geqslant_I G$$

if and only if:

$$\mathcal{V}(\hat{\phi}_d, F, \pi, q) \geq \mathcal{V}(\hat{\phi}_d, G, \pi, q)$$

3.1 Strong dominance

It follows from these definitions that how we rank two distributions depends not only on the their characteristics but also on the heterogeneity and distribution of group members' preferences, and the decision rule used to make the collective choice. Appropriate restrictions on these objects could deliver an answer to our question but with a loss in generality. We first attempt to provide a more general answer. This is in terms of conditions on distributions that, if satisfied, give a ranking of distributions that does not depend on the group's characteristics and the decision rule used. The only essential assumption is that of the group being like-minded.

The following definition formalizes the type of relation among distributions we look for:

Definition 4. *We say that F* strongly dominates *G if $F \neq G$ and $F \succcurlyeq_I G$ for any I and anny q. We denote such a relationship as*

$$F \gg G$$

We obtain a partial order of distributions in terms of strict dominance. Our first main result gives the conditions that allow us to rank two distributions in such a way.

Proposition 1. *Let I be a like-minded group with $t_i \in T$. Let F, G be two information structures and $\hat{s}_H(k) = argmax\{\frac{h_1(s_H)}{h_0(s_H)} \leq k\}$ for $(H, h) \in \{(F, f), (G, g)\}$.*
Then,

$$F \gg G$$

if and only if, for all $k > 0$,

$$F_0(\hat{s}_F(k)) \geq G_0(\hat{s}_G(k)) \tag{2}$$

and

$$F_1(\hat{s}_F(k)) \leq G_1(\hat{s}_G(k)) \tag{3}$$

To understand the conditions in Proposition 1 notice that: $F_0(\tilde{s}_F(t_d(q))$ is the probability that the group takes the right decision ($x = 0$) when the state is $\theta = 0$, while $1 - F_1(\tilde{s}_F(t_d(q))$ is the probability of taking the right decision ($x = 1$) when the state is $\theta = 1$. Changing the voting rule changes the type of the decisive voter and therefore also $\tilde{s}(t_d(q))$. If the conditions in the proposition hold, then for any possible voting rule and any possible group or, instead, any possible value of $\tilde{s}(t_d(q))$, the probability of making the right decision in any of the two states of nature under F is higher than under G. In other words, using F reduces both type I and type II errors.

As we show next, these are very strong requirements for comparing distributions. In particular, if $F \geqslant_I G$ for any group, then it must also be true for single individuals. That is, any individual decision maker with preferences as the ones in our model agrees on the ranking of F and G. Recall that in the example of the introduction, although individuals would separately agree on one ranking, the ranking for the group was the opposite.

Furthermore, given that by using F instead of G, both types of errors are reduced, all members in a group are better-off when the group makes the decision under F. It would be therefore useless for any member of the group to miss-report his true valuation if that could affect the choice of information structure. Still, although everybody would agree on the best distribution for the collective choice, there does not have to be agreement with the group's final decision. For a given public signal from F there may still be some individuals supporting the high action and others supporting the low action.

We now give an example of two distributions F and G for which the conditions in Proposition 1 are satisfied and hence $F \gg G$.

Example 2. *Let F and G be two distributions over four different values: $S_H = \{s_H^1, s_H^2, s_H^3, s_H^4\}$, for $H \in \{F, G\}$. The table in Figure 6 gives the complete description of the two distributions.*

	$\mathbf{F_0}$	f_0	$\mathbf{F_1}$	f_0	$\frac{f_1}{f_0}$		$\mathbf{G_0}$	g_0	$\mathbf{G_1}$	g_1	$\frac{g_1}{g_0}$
$\mathbf{s_F^1}$	**0.9**	0.9	**0**	0	0	$\mathbf{s_G^1}$	**0.3**	0.3	**0**	0	0
$\mathbf{s_F^2}$	**0.96**	0.06	**0.025**	0.025	$\frac{5}{12}$	$\mathbf{s_G^2}$	**0.8**	0.5	**0.15**	0.15	$\frac{3}{10}$
$\mathbf{s_F^3}$	**1**	0.04	**0.1**	0.075	$\frac{15}{8}$	$\mathbf{s_G^3}$	**1**	0.2	**0.55**	0.4	2
$\mathbf{s_F^4}$	**1**	0	**1**	0.9	∞	$\mathbf{s_G^4}$	**1**	0	**1**	0.45	∞

Figure 6: These distributions are such that $F \gg G$

To check whether the conditions of Proposition 1 are satisfied we calculate the pairs $\{\hat{s}_F(k), \hat{s}_G(k)\}$ for different values of k and see whether the inequalities

(2) and (3) hold.

$$\{\hat{s}_F(k), \hat{s}_G(k)\} = \begin{cases} \{s_F^1, s_G^1\}, & k \in (0, \frac{3}{10}) \\ \{s_F^1, s_G^2\}, & k \in [\frac{3}{10}, \frac{5}{12}) \\ \{s_F^2, s_G^2\}, & k \in [\frac{5}{12}, \frac{15}{8}) \\ \{s_F^3, s_G^2\}, & k \in [\frac{15}{8}, 2) \\ \{s_F^3, s_G^3\}, & k \in [2, \infty) \end{cases}$$

The graph of the two distributions in figure 7 allows to easily check that the inequalities (2) and (3) hold for each one of these pairs.

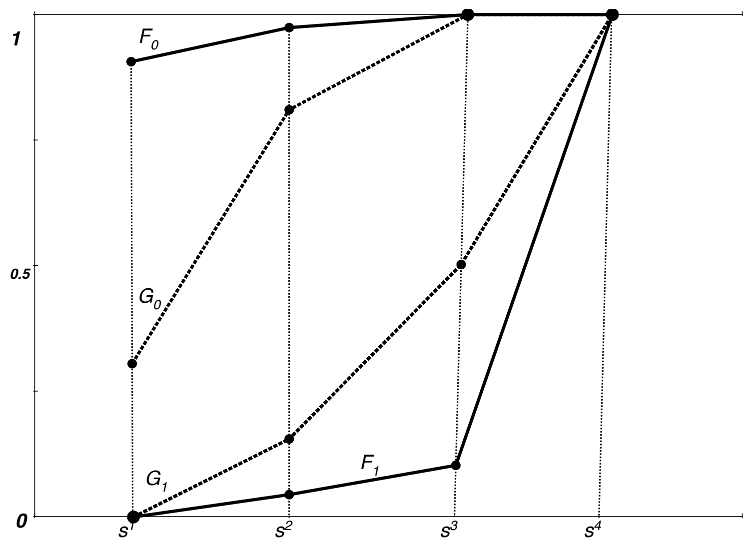

Figure 7: The cdf's for F and G.

It was mentioned already that the cases where a distribution strongly dominates another one are not common. Our next result reflects exactly that.

Proposition 2. *Let F, G be such that $F \gg G$.*

1. *If $\frac{g_1(s)}{g_0(s)} > 0\ \forall s \neq \underline{s}_G$, then there exists $s_F \in S_F : Pr(\theta = 0|s_F) = 1$*

2. *If $\frac{g_0(s)}{g_1(s)} > 0\ \forall s \neq \underline{s}_G$, then there exists $s_F \in S_F : Pr(\theta = 1|s_F) = 1$*

To understand the statement in the proposition it is useful to note which possible forms of G are not included. If $S_G = \mathbb{R}$ and G is such that both $g_0(s_G)$ and $g_1(s_G)$ asymptotically tend to zero in both directions (as s increases or decreases), then Proposition 2 does not apply. If this is not true in either direction and $F \gg G$, then F must be such that the most extreme signal in the respective direction is fully revealing of the state.

As a consequence of Proposition 2, it turns out that the extend to which one can order distributions in terms of strict dominance is limited. Many situations where, intuitively, a certain information structure seems more informative than another are not included in this order.Furthermore, even some comparisons that are possible under the Blackwell criterion are excluded here. This becomes particularly striking when one considers comparisons among binary signals. The following corollary is an application of Proposition 2 to this case.

Corollary 1. If $F \gg G$ and $S_H = \{s'_H, s''_H\}$, for $s'_H < s''_H$ and $H \in \{F, G\}$, then F must be fully revealing: $Pr(\theta = 0|s'_F) = Pr(\theta = 1|s''_F) = 1$.

What is stated here is the following: a fully revealing signal is the only information structure that can strictly dominate a binary signal. This shows that the strict dominance relation is much stronger than any intuitive notion of informativeness one might have. Simply reducing the noise in a signal is not enough.

Why is this so? The best way to understand this is to return to the example from the introduction. The signal from the interview is such that the group's choice is not affected by the signal's realization. The group always hires the candidate after an interview. This means that the group never makes the error of choosing not to hire when it should be hiring. Let's call this a type I error. It may make the error of hiring when it should not be doing it. This would be a type II error. A more precise signal, like the one from the test leads to a collective choice that depends on the realization of the signal. Therefore, after a test the group may commit a type I error with a positive probability. We know from Proposition 1 that under strict dominance both types of errors must be reduced, which is not happening in this case. It turns out that it is always possible to construct such a counterexample with binary signals, unless the dominating signal is perfectly revealing. The same reasoning lies behind the more general result in the proposition.

One conclusion that we can draw at this point is that being able to make comparisons between distributions focusing only on their properties and without any restrictions on the groups preferences or decision rule is closer to an exception than to a rule. This gives rise to the question of whether one can extend this partial order by setting restrictions on either the group's preferences or the voting rule used. The next section deals with the first type of restrictions.

3.2 Restricted preferences

It is not hard to find pairs of distributions that satisfy inequalities (2) and (3) in Proposition 1 for a particular profile of group members' valuations. But these may no longer hold once we consider another group, containing more biased individuals. This happens, when the voting rule is such that one of these biased individuals becomes decisive. It is not hard to understand then, that if inequalities (2) and (3) hold for a particular group, they should also hold for any group where "extreme" biases are reduced. This idea is formalized in the following proposition.

Proposition 3. *Let F, G be two information structures and* $\hat{s}_H(k) = argmax\{\frac{h_1(s_H)}{h_0(s_H)} \leq k\}$ *for* $(H,h) \in \{(F,f),(G,g)\}$*. If:*

$$F_0(\hat{s}_F(k)) \geq G_0(\hat{s}_G(k)) \text{ and } F_1(\hat{s}_F(k)) \leq G_1(\hat{s}_G(k))$$

$$\forall\, k \in \left[\frac{1-\pi}{\pi}\frac{1}{\lambda(\bar{t})}\ ,\ \frac{1-\pi}{\pi}\frac{1}{\lambda(\underline{t})}\right]\ ,\ 0 < \underline{t} < \bar{t}$$

then

$$F \succcurlyeq_I G \quad \forall\, I\,,\ q\,,\ \text{and } T \subseteq [\,\underline{t}\,,\,\bar{t}\,]$$

Up to here, our results concern cases where one distribution is such that a decision can be made by the group in a way that the probability of making the correct decision is higher in any state of the world. We now turn to cases where this may not be true but a particular distribution still offers a higher expected value than another. To obtain such conditions we must make specific restrictions on the group's preferences. For these it is useful to define here a measure of the group's bias.

Definition 5. *The group's bias is:*

$$\Lambda(u,T) = \frac{\int_T [u_i(1,1,t_i) - u_i(0,1,t_i)]\xi(t_i)dt}{\int_T [u_i(0,0,t_i) - u_i(1,0,t_i)]\xi(t_i)dt}$$

Notice that the group bias is not the average of the group's members' biases. For instance, in the example of a LMG in Figure 5 the individual biases of the two group members are $\lambda(t_1) = 1$ and $\lambda(t_2) = \frac{1}{4}$. The average bias is then $\frac{5}{8}$. On the other hand, the group's bias, as defined here, is $\Lambda(u,T) = \frac{1+20}{1+80} = \frac{21}{81}$. That is, the magnitude of individuals' valuations matter.

It turns out that by restricting the group's bias to a particular value that depends on the prior belief about the state, one can find conditions that allow for a comparison between distributions that holds for any voting rule. These are given in the following result.

Proposition 4. *Let I be a like-minded group such that $\Lambda(u,T) = \frac{1-\pi}{\pi}$. Let F and G be two information structures and let $\hat{s}_H(k) = argmax\{\frac{h_1(s_H)}{h_0(s_H)} \leq k\}$ for $(H,h) \in \{(F,f),(G,g)\}$. Then*

$$F \geqslant_I G \,,\, \forall\, q$$

if and only if,

$$F_0(\hat{s}_F(k)) - F_1(\hat{s}_F(k)) \geq G_0(\hat{s}_G(k)) - G_1(\hat{s}_G(k))$$
$$\forall\, k > 0$$

The intuition behind the condition in proposition 4 is similar to the one that gives the conditions of proposition 1. In that case, F had to such that the probability of an error was reduced in each of the two possible states. Here, the restrictions set on the group's preferences allow us to relax this condition. What needs to hold is that the sum over states of the probability of an error is reduced. Said the over way around, the sume over states of the probability of taking the right action is increased. This is more easily seen if we write the inequality in the following way: $F_0(\hat{s}_F(k)) + [1 - F_1(\hat{s}_F(k))] \geq G_0(\hat{s}_G(k)) + [1 - G_1(\hat{s}_G(k))]$.

Example 3. *Here we give an example of a class of distributions that can be ordered in the way described in Proposition 4, assuming a uniform prior: $\pi = \frac{1}{2}$ and $\Lambda(t) = 1$. Consider the class of distributions where $F_0(s) = 1 - F_1(1-s)$ and $s \in [0,1]$. We call such distributions* symmetric. *We also say that a distribution F* state-wise stochastically dominates *another distribution G when $F_0(s) > G_0(s)$ and $F_1(s) < G_1(s)$.*

Now, let $F(s;\alpha)$ represent a family of symmetric distributions such that $F(s;\alpha)$ state-wise stochastically dominates $F(s;\alpha')$ for any $\alpha > \alpha'$ [4]*. The following graph represents an example of two such distributions.*

[4] As an instance of a family of distributions that satisfies all of the conditions stated in the example, one can think of power distributions with $F_0(s;\alpha) = 1 - (1-s)^\alpha$ and $F_1(s;\alpha) = s^\alpha$, $s \in [0,1]$ and $\alpha \geq 1$.

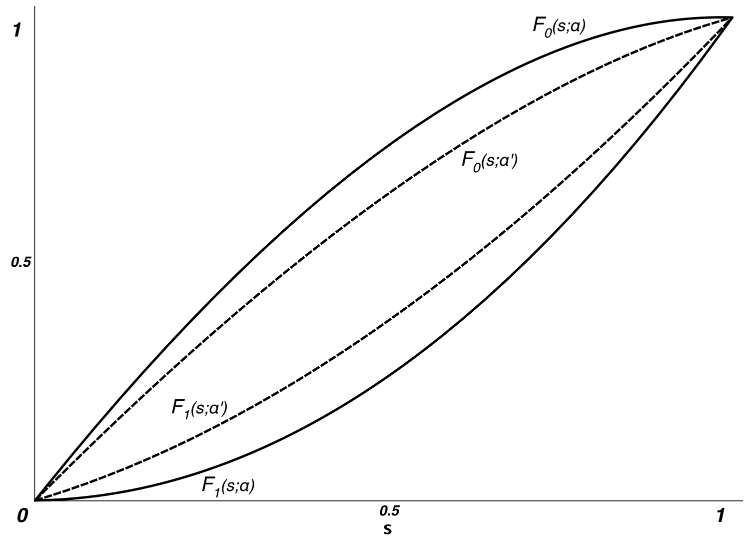

Figure 8: Two distributions from the power distribution family, with $\alpha > \alpha'$.

First note that any symmetric distribution has the following property[5]*:*

$$\frac{f_1(\frac{1}{2};\alpha)}{f_0(\frac{1}{2};\alpha)} = 1$$

Let $\hat{s}(s) = \{s' : \frac{f_1(s;\alpha)}{f_0(s;\alpha)} = \frac{f_1(s';\alpha')}{f_0(s';\alpha')}\}$.
With the help of the graph, it is easy to see that the following is true:

- *First we have* $\hat{s}(\frac{1}{2}) = \frac{1}{2}$*. Then, one can observe from Figure 8 that*

$$F_0\left(\frac{1}{2};\alpha\right) - F_1\left(\frac{1}{2};\alpha\right) > F_0\left(\frac{1}{2};\alpha'\right) - F_1\left(\frac{1}{2};\alpha'\right)$$

- *Then, for* $s \in [0, \frac{1}{2})$ *it must be* $\hat{s}(s) < s$*. Hence, it must be:*

$$F_0(s;\alpha) - F_1(s;\alpha) > F_0\left(\hat{s}(s);\alpha'\right) - F_1\left(\hat{s}(s);\alpha'\right)$$

- *Finally, following the same logic, for* $s \in (\frac{1}{2}, 1]$ *it must be* $\hat{s}(s) > s$ *. Hence, again it must be:*

$$F_0(s;\alpha) - F_1(s;\alpha) > F_0\left(\hat{s}(s);\alpha'\right) - F_1\left(\hat{s}(s);\alpha'\right)$$

[5]Because of symmetry: $F_0(s) = 1 - F_1(1-s)$. Taking the derivative on both sides and substituting for $s = \frac{1}{2}$ gives the property. I am grateful to Dimitris Xefteris for pointing that out.

Summing up, the above means that the condition in Proposition 4 holds and thus, we can say that for a like-minded group with $\Lambda(u, T) = 1$ and a uniform prior, any family of symmetric distributions that can be ordered according to state-wise stochastic dominance, can be ordered according to Proposition 4 in the following way: $F(s; \alpha) \geqslant F(s; \alpha')$ for all $\alpha > \alpha'$.

The previous example demonstrates that when restricting our attention to unbiased groups the scope for ordering distributions, in terms of informativeness for the group, increases. With the following example we demonstrate that, nevertheless, there are limits to the possibility of ordering distributions. The previous example may lead one to believe that state-wise stochastic dominance of the distributions would be enough, but this is not true.

Example 4. *Consider two distributions F and G. The signal may take one of three possible values: $S = \{s^1, s^2, s^3\}$. The table in Figure 9 gives the complete description of the two distributions.*

	$\mathbf{F_0}$	f_0	$\mathbf{F_1}$	f_0	$\frac{f_1}{f_0}$		$\mathbf{G_0}$	g_0	$\mathbf{G_1}$	g_1	$\frac{g_1}{g_0}$
$\mathbf{s^1}$	**0.8**	0.8	**0.05**	0.05	$\frac{1}{16}$		**0.7**	0.7	**0.1**	0.1	$\frac{1}{7}$
$\mathbf{s^2}$	**0.95**	0.15	**0.4**	0.35	$\frac{7}{3}$		**0.85**	0.15	**0.5**	0.4	$\frac{8}{3}$
$\mathbf{s^3}$	**1**	0.5	**1**	0.6	12		**1**	0.15	**1**	0.5	$\frac{10}{3}$

Figure 9: These distributions can not be ordered, even when $\Lambda(u, T) = 1$

These two distributions can not be compared in the terms posited in Proposition 4. Suppose the group uses a voting rule such that $\lambda(t_d(q)) \in (\frac{3}{8}, \frac{3}{7})$. Then $\tilde{s}_F(t_d(q)) = s^2$ and $\tilde{s}_G(t_d(q)) = s^1$. But in that case, $F_0(s^2) - F_1(s^2) = 0.55$ while $G_0(s^1) - G_1(s^1) = 0.6$, and the condition of Proposition 4 does not hold. The following graph depicts the cumulative distribution functions for both F and G:

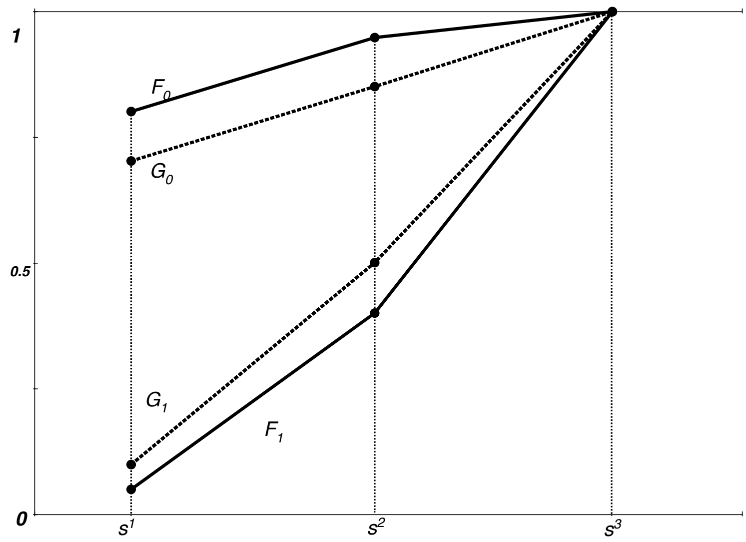

Figure 10: F and G can not be ordered for the group.

As can be seen from the graph, G_0 stochastically dominates F_0 and F_1 stochastically dominates G_1. Yet, this is not enough to consider F more informative than G for any voting rule.

Up to this point we have examined cases where an ordering of distributions is possible for any possible voting rule. In the remainder of this section we examine the possibility of ordering distributions when a particular voting rule is applied, and its role in the group members' demand for information.

3.3 The optimal voting rule

Our results show that it is possible to compare distributions without paying attention to the specific voting rule. Still, different voting rules affect the final outcome of the collective choice process. It makes sense then to ask whether there exists an "optimal voting rule", and if the answer is affirmative, what institutional elements determine it.

Our optimality criterion, in accordance with previous sections, is aggregate expected value maximization. As we show, it is possible for such an "optimal" rule to exist. Furthermore, it depends only on the profile of the group members' valuations.

Lemma 5. *The group's value is maximized for* $q^* = \{q : \lambda\left(t_d(q^*)\right) = \Lambda\left(u, T\right)\}$

According to Lemma 5 an optimal rule must be such that the bias of the decisive voter coincides with the group's bias. In other words, the group's expected value is maximized if it makes decisions in the same way as an individual decision maker with the same bias would make them.

Whether such a member exists within the group is not guaranteed by any means. If group members form a continuous in the type space, then there should exist an individual with the required bias. On the other hand, when the set of agents is discrete, or generally non-convex, there may not exist such a representative agent. The example of Anne and Bob in the introduction is such an instance. The group's bias is $\Lambda(u,T) = \frac{\frac{1}{4}+\frac{2}{3}}{1+1} = \frac{11}{24}$ and does not coincide with any of the two agents' biases because: $\lambda(t_{Anne}) = \frac{1}{4}$ and $\lambda(t_{Bob}) = \frac{2}{3}$. As a consequence, it seems easier to approach an optimal voting rule in large groups than in small ones.

3.4 The optimal voting rule and comparisons of information structures

Having defined the optimal voting rule in such a way we apply it in the task of comparing distributions. This is done in the following result.

Proposition 5. *Let F and G be two information structures. Let $q = q^*$. If $F \geqslant_d G$, then $F \geqslant_I G$.*

According to this result, if the optimal voting rule exists and the decisive voter it defines is such that one distribution is more informative than the other for this individual, then the same is true for the group. The importance of this result lies in the fact that the use of the optimal rule not only maximizes the group's expected value. It also allows to use notions of informativeness established for individual decision problems to the group's problem. This means for instance that any set of distributions that can be ranked according to Blackwell's criterion for individuals, can be ranked in the same way for a group that uses an optimal voting rule. But Blackwell's criterion is not the only useful one. Our decision problem is such that when the group is comprised of a single individual, it falls within the class of monotone decision problems. Lehman (1988) gives criteria according to which distributions can be compared in the context of monotone decision problems. Under these same criteria it is possible to rank distributions that are not comparable according to Blackwell's notion of informativeness.

Our next example illustrates a situation where an optimal rule does not exist. Combining this with our previous results gives a situation where the axiom "more information is better" fails.

Example 5. *Consider a group made up of n individuals. Each individual's valuation is given by $u(x, \theta, t_i) = -|x - \theta| - |x - t_i|$, where $t \in T = (0, 1)$. It follows that each individual's bias function is given by $\lambda(t_i) = \frac{t_i}{1-t_i}$. Let us assume that n is even and individuals are distributed over T. That is, for any individual i with $t_i = t'$ there exists another individual j with $t_j = 1 - t'$. Given this symmetry, the group is unbiased: $\Lambda(u, T) = 1$.*

Suppose that it is possible for each individual in the group to obtain information regarding the state θ. Furthermore, suppose that a mechanism exists that fully aggregates such information and makes it available to the group. This is of course a strong assumption, especially if the information is not verifiable[6]*. What we want to examine here is whether is is better for the group that all individuals obtain information or some abstain from doing so. Following the "condorcet jury" tradition, we assume that the available information comes in the form of a binary signal that truthfully reveals the state with probability p: $Pr(s_i = \theta) = p$. Since private signals are fully aggregated, the public signal for the group amounts to some $s \in \{0, 1, ..., m + 1\}$, where $m \leq n$ is the number of individuals that obtained a private signal, and it follows a binomial distribution with parameters m and either p or $1 - p$, depending on the state.*

The optimal voting rule here according to lemma 5 would in this case be one that sets $\lambda(t_d) = 1$, that is $t_d = \frac{1}{2}$. Alas, since we assume that n is even, no optimal rule exists in this case. Let us then assume that simple majority is needed for $x = 1$. This means that if one orders individuals according to their types, such that $t_i < t_j$ for $i < j$, then $t_d = t_{\frac{n}{2}}$.

In what follows we shall put specific numerical values to the different parameters of the model and show that it is better for the group to obtain "less" information.

Let $n = 6$ and $p = .8$. We shall compare the case where all 6 individuals obtain information ($m = 6$) to a case where all but one do so ($m = 5$). The following tables give the likelihood ratios and difference in cdf's, for each of these two cases.

s_F	0	1	2	3	4	5	6
$\frac{f_1(s_F)}{f_0(s_F)} \approx$	0.0002	0.004	0.06	1	16	256	4096
$F_0(s_F) - F_1(s_F) \approx$	0.26	0.65	0.88	0.88	0.65	0.26	0

Figure 11: Six available signals: $F_0(s) = \sum_{i=0}^{s} \binom{6}{i} .2^i (1 - .2)^{6-i}$ and $F_1(s) = \sum_{i=0}^{s} \binom{6}{i} .8^i (1 - .8)^{6-i}$

[6] For more on this issue in the case of verifiable information see Schulte (2008) [23] and for the cases of non-verifiable information Doraszelski et al (2003) [7] and Austen-Smith and Feddersen (2005) [3]

s_G	0	1	2	3	4	5
$\frac{g_1(s_G)}{g_0(s_G)} \approx$	0.001	0.016	0.25	4	64	1024
$G_0(s_G) - G_1(s_G) \approx$	0.33	0.73	0.88	0.73	0.33	0

Figure 12: Five available signals: $G_0(s) = \sum_{i=0}^{s} \binom{5}{i}.2^i(1-.2)^{5-i}$ and $G_1(s) = \sum_{i=0}^{s} \binom{5}{i}.8^i(1-.8)^{5-i}$

It can easily be seen that proposition 4 does not hold: take for instance $k \in (16, 64)$. Then $(\hat{s}_F(k), \hat{s}_G(k)) = (4, 3)$ and $F_0(4) - F_1(4) < G_0(3) - G_1(3)$. A similar contradiction occurs when $k \in (256, 1024)$.

In terms of our example, this means that since $t_d = t_3$, whenever $t_3 \in (\frac{1}{65}, \frac{1}{17}] \cup (\frac{1}{1025}, \frac{1}{257}]$ the expected value of the group is higher under G than under F.

The reason this is so is the following. The additional private signal allows for a richer set of public signals. For an individual decision maker this would suffice to make F more valuable than G. But the additional signal also means a different distribution of signals, and this is relevant in our setup of collective decision making. In particular, given a decisive voter, the signal space gets partitioned in two subsets: one for lower signals that lead to $x = 0$ and one for higher signals that lead to $x = 1$. The public signal comes from one of these subsets. Fixing a voting rule aslo fixes a distribution over these subsets in each state. The additional signal leads to a different distribution over these two subsets in each state. When the parameters are the ones mentioned above, this new distribution leads to more frequent mistakes.

This example demostrates how in certain situations full information aggregation may not be desirable from the point of view of the group. Notice that besides assuming an even number of subjects in the group, which leads to the inexistence of an optimal rule, assumptions about the group's mebers preferences are quite regular. The counterintuitive result is derived from the statistical properties of the available information structures and the way these interact with the collective decision making process.

3.5 Group members' demand for information

We have looked at the comparison of information structures from two points of view: the group's and the individual's as a decision maker. Speaking in the terms of our introductory example, the first refers to the value of information to the group formed by Anne and Bob. The second refers to the value of information to either Anne or Bob in a case where one of them is deciding on his or her own whether or not to hire the candidate. The first is the main object of this paper. The second serves as a yardstick that allows us to measure the degree to which the notion of informativeness for the group departs from the notion of informativeness for individuals.

There is a third point of view though that may be useful when applying the concepts of this paper in economic models of collective choice: the value of information to individuals within the group. Again in terms of the example, this refers to the value of information to Anne when she must make a decision together with Bob. The next section considers an application where this value is relevant. In general it will may be relevant in cases where group members must take actions that influence the choice of the information structure that is chosen to make the final decision. An example of this would be if group members are to vote on what source of information should be used.

Our next result characterizes this demand for information within the group. It should be no surprise that this demand depends on the voting rule that is used.

Let $\Gamma_q(F,G) = \{t : E_F[u(x(q,s_F),\theta,t)] \geq E_G[u(x(q,s_G),\theta,t)]\}$. That is, $\Gamma_q(F,G)$ is the set of the group's members that prefer distribution F over G or, in other words, all $i \in I$ such that $F \succcurlyeq_i G$. The following lemma characterizes this set.

Lemma 6. *Let* $\{\underline{t},\bar{t}\} = \{min\{T\}, max\{T\}\}$ *and* $\hat{t}(q) = \left\{t \;:\; \lambda(t) = \frac{1-\pi}{\pi}\frac{F_0(\tilde{s}_F(t_d(q)))-G_0(\tilde{s}_G(t_d(q)))}{F_1(\tilde{s}_F(t_d(q)))-G_1(\tilde{s}_G(t_d(q)))}\right\}$.

$$\Gamma_q(F,G) = \begin{cases} [\underline{t},\bar{t}]\,, & F_0(\tilde{s}_F(t_d(q))) \geq G_0(\tilde{s}_G(t_d(q))) \;\; and \;\; F_1(\tilde{s}_F(t_d(q))) \leq G_1(\tilde{s}_G(t_d(q))) \\ [\underline{t},\hat{t}(q)]\,, & F_0(\tilde{s}_F(t_d(q))) \geq G_0(\tilde{s}_G(t_d(q))) \;\; and \;\; F_1(\tilde{s}_F(t_d(q))) \geq G_1(\tilde{s}_G(t_d(q))) \\ [\hat{t}(q),\bar{t}]\,, & F_0(\tilde{s}_F(t_d(q))) \leq G_0(\tilde{s}_G(t_d(q))) \;\; and \;\; F_1(\tilde{s}_F(t_d(q))) \leq G_1(\tilde{s}_G(t_d(q))) \\ \varnothing\,, & otherwise \end{cases}$$

First, if F and G are such that given the voting rule there is a higher probability of taking the right decision in both states of the world under F, then all group members prefer the decision to be taken under F. Second, if the distributions are such that under F it is more likely to take the right decision in the low state but less likely in the high state, then there exists a type $\hat{t}$ that is indifferent between the two distributions, and all individuals to his left prefer F over G. This situation is reversed if under F it is more likely to take the right decision in the high state but less likely to do so in the low state. Finally, if a right decision in any state is more likely under G, then all agents prefer that distribution over F.

So, according to this result, group members are split: the ones of a lower type prefer one distribution while others of a higher type prefer the other one. Notice though that this division depends directly on the voting rule. Changing the voting rule not only moves the line of division. It can also lead to a switch of preferences for some group members.

This concludes our analysis of the possibility of comparing different distribution structures. In the following section we apply some of our results to a more structured collective decision problem. We do this to demonstrate the applicability of the tools we introduce in the study of collective decision making and the design of relevant institutions.

4 A model with endogenous information acquisition.

We now turn to an example of an application where we can make use of the ranking resulting from a comparison of information structures. In the model we consider, the group must make a decision. More information arrives if the group is willing to wait for it. Group members that want the group to obtain more information will try to prolong the waiting time. Others will press for an immediate decision. The more individuals press for either option, the more likely it is to happen. This is similar to the models of Gersbach (1992) [12] and Messner and Polborn (2012) [19] but with one big difference: there is no vote to decide whether or not to wait for more information. It is determined stochastically, depending on the proportion of group members that support this option.

As we know from our previous analysis, the demand for information in the group depends on the voting rule used. First we characterize the optimal voting rule for this setting. This differs from the optimal rule we described in the previous section where information was considered entirely exogenous. That is because now its effect on the demand for information, and hence on the likelihood of obtaining more information, must be taken in to account. The trade-off is the following. Positioning the voting rule away from q^* increases the demand for information within the group and therefore the likelihood to obtain more information before takingthe final decision. On the other hand, the final decision is not taken optimally any longer.

The optimal voting rule in this setting depends on the information structure that may provide the public signal. For some information structures that can be ranked according to Proposition 4, more informative distributions are associated with optimal rules for endogenous information that are further away from the optimal rule for exogenous information.

4.1 The model with endogenous information

To simplify the analysis, in this section we assume that I is a continuum distributed uniformly in the unit interval: $\xi(t) = 1$, $t \in [0, 1]$. Given this assumption we can economize on notation. In particular, note that now $t_d(q) = q$. This comes without any particular loss in generality, since we have not imposed any restrictions on $\lambda(t)$. We further assume that the group is unbiased, in terms of Proposition 4 : $\Lambda(u, T) = \frac{1-\pi}{\pi}$. This assumption is quite restrictive and not necessary in order to perform our analysis. Still, it lets us focus on the role of information when setting the optimal voting rule without having to bother about the group's bias or priors.

As mentioned before, the group faces two alternative scenarios: to have some additional information or none. Let the public signal be $\sigma \in \{\emptyset, s\}$, with $s \sim F(s)$. Again, in order to keep analysis tractable we focus on distributions with a continuous domain: $s \in [\underline{s}, \bar{s}]$. The following lemma characterizes the group's decision rule under no information:

Lemma 7. *The group's decision is given by*

$$x(q, \emptyset) = \begin{cases} 1, & q > q^* \\ 0, & \textit{otherwise} \end{cases}$$

where

$$q^* = \left\{ q \ : \ \lambda(q) = \frac{1-\pi}{\pi} \right\}$$

Let $\gamma(q) = \int_{t \in \Gamma_q} \xi(t) dt$ be the fraction of group members that prefer that the group receives public information before making a decision. According to Lemma 6, Γ_q takes the following form:

$$\Gamma_q(F, \emptyset) = \begin{cases} [o, \hat{t}(q)] & , \ q > q^* \\ (\hat{t}(q), 1] & , \ q \le q^* \end{cases}$$

where $\hat{t}(q) = \left\{ t \ : \ \lambda(t) \frac{\pi}{1-\pi} \frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))} = 1 \right\}$. From this it follows that

$$\gamma(q) = \begin{cases} \hat{t}(q) & , \ q > q^* \\ 1 - \hat{t}(q) & , \ q \le q^* \end{cases}$$

It is important to note that given the definition of $\hat{t}(q)$ and the behavior of $\tilde{s}(q)$ with respect to q, the further away from q^* is q, the higher is the demand for information by group member's, captured by $\gamma(q)$.

Now we assume that whether or not the group receives a public signal depends on $\gamma(q)$. In particular we assume that there exists a function $\mu : [0, 1] \rightarrow [0, 1]$ such that $\mu(\gamma(q)) = Pr(\sigma = s)$. We assume μ is increasing in it's argument.

One interpretation of this is that a contest ensues between group members: some exert effort to keep the group from deciding before information arrives, while others exert effort to accomplish the opposite. In this case, $\mu(\gamma)$ is the contest success function.

In such scenario, the aggregate utility of the group is given by:

$$W(q) = \mu\left(\gamma(q)\right) \mathcal{V}(\phi_d, F, \pi, q) + \left[1 - \mu\left(\gamma(q)\right)\right] \mathcal{V}(\phi_d, \emptyset, \pi, q)$$

This is actually a piecewise function and depends on whether the voting rule q is higher or lower than q^*. We have: For $q \geq q^*$:

$$\begin{aligned} W(q) =& \pi \int_T u(1,1,t)dt + (1-\pi) \int_T u(1,0,t)dt \\ &+ \mu(\gamma(q))(1-\pi)F_0(\tilde{s}(q)) \int_T [u(0,0,t) - u(1,0,t)]dt \\ &- \mu(\gamma(q))\pi F_1(\tilde{s}(q)) \int_T [u(1,1,t) - u(0,1,t)]dt \end{aligned}$$

While for $q < q^*$:

$$\begin{aligned} W(q) =& \pi \int_T u(0,1,t)dt + (1-\pi) \int_T u(0,0,t)dt \\ &+ \mu(\gamma(q))\pi[1 - F_1(\tilde{s}(q))] \int_T [u(1,1,t) - u(0,1,t)]dt \\ &- \mu(\gamma(q))(1-\pi)[1 - F_0(\tilde{s}(q))] \int_T [u(0,0,t) - u(1,0,t)]dt \end{aligned}$$

To find the optimal voting rule one must solve the first order condition.[7] Again, given the piecewise nature of the function, we get two equations.

[7] We do not check the second order condition. Uniqueness depends on particular functional forms but should in general not be an issue given the various monotonicity assumptions made so far. In any case, the first order condition characterizes the optimal rule, unless we have a corner solution which in this case would be unanimity. An interior solution is assumed in what follows.

For $\mathbf{q} > \mathbf{q}^*$ we have the following[8]:

$$\frac{\partial \mu(\gamma(q))}{\partial \gamma}\frac{\partial \gamma(q)}{\partial q}F_1(\tilde{s}(q)) + \mu(\gamma(q))f_1(\tilde{s}(q))\frac{\partial \tilde{s}(q)}{\partial q} = \frac{\partial \mu(\gamma(q))}{\partial \gamma}\frac{\partial \gamma(q)}{\partial q}F_0(\tilde{s}(q)) + \mu(\gamma(q))f_0(\tilde{s}(q))\frac{\partial \tilde{s}(q)}{\partial q}$$

Notice that for $q \geq q^*$ we have $\gamma(q) = \hat{t}(q)$. That gives:

$$\mu(\gamma(q))f_0(\tilde{s}(q))\frac{\partial \tilde{s}(q)}{\partial q} + \frac{\partial \mu(\gamma(q))}{\partial \gamma}\frac{\partial \hat{t}(q)}{\partial q}F_0(\tilde{s}(q)) = \mu(\gamma(q))f_1(\tilde{s}(q))\frac{\partial \tilde{s}(q)}{\partial q} + \frac{\partial \mu(\gamma(q))}{\partial \gamma}\frac{\partial \hat{t}(q)}{\partial q}F_1(\tilde{s}(q))$$

By implicit differentiation we get:

$$\frac{\partial \hat{t}(q)}{\partial q} = -\frac{\lambda(\hat{t}(q))}{\frac{\partial \lambda(\hat{t}(q))}{\partial t}}\frac{\partial \tilde{s}(q)}{\partial q}\frac{\left(\frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))}\right)'_s}{\frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))}}$$

Plugging this in to the last expression and simplifying we obtain the following:

$$\frac{f_0(\tilde{s}(q)) - f_1(\tilde{s}(q))}{F_0(\tilde{s}(q)) - F_1(\tilde{s}(q))}\frac{\frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))}}{\left(\frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))}\right)'_s} = \frac{\frac{\partial \mu(\gamma(q))}{\partial \gamma}}{\mu(\gamma(q))}\frac{\lambda(\hat{t}(q))}{\frac{\partial \lambda(\hat{t}(q))}{\partial \hat{t}}} \tag{4}$$

Let

$$\begin{aligned}\Delta(q) =& F_0(\tilde{s}(q)) - F_1(\tilde{s}(q)) \\ =& F_0(\tilde{s}(q)) - [1 - F_1(\tilde{s}(q))] - 1 \\ \propto& Pr(x = \theta|\sigma = s) - Pr(x = \theta|\sigma = \varnothing)\end{aligned}$$

That is, $\Delta(q)$ represents the improvement in the likelihood of making the correct decision after receiving a public signal, given q. Then 4 can be written as:

$$\frac{\Delta'_s(q)}{\Delta(q)}\frac{\frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))}}{\left(\frac{F_1(\tilde{s}(q))}{F_0(\tilde{s}(q))}\right)'_s} = \frac{\frac{\partial \mu(\gamma(q))}{\partial \gamma}}{\mu(\gamma(q))}\frac{\lambda(\hat{t}(q))}{\frac{\partial \lambda(\hat{t}(q))}{\partial \hat{t}}} \tag{5}$$

Expression (5) characterizes the optimal voting rule in this setup. It captures the

[8]For $\mathbf{q} < \mathbf{q}^*$ we obtain a similar expression as the one we obtain here. The only difference is a negative sign in the RHS. Which of the two solutions represents the optimal voting rule depends on the specific functional forms for $\lambda(t)$ and μ. What interests us more here are the comparative statics that are of a similar nature in either case. We therefore focus the analysis on this expression keeping in mind that the actual optimal voting rule may be characterized by the symmetric expression.

trade-off we have described between optimal decision making and increased demand for information. The first is captured by the first term in the LHS which represents the proportional change of $\Delta(q)$. The other terms capture the change in demand for information. In particular, the RHS together with the second term of the LHS can be thought as the elasticity of the likelihood of receiving the public signal with respect to the the proportion of group members that want the public signal. This proportion depends on the individuals' biases, captured by the last term in the RHS and the form of the available information structure, captured by the second term in the LHS.

First note that for $q = q^*$ as defined in lemma 5, $\frac{\Delta_s'(q)}{\Delta(q)}$ must be zero,[9] since $\Lambda(u, T) = \frac{1-\pi}{\pi}$. Given that the terms in the RHS are strictly positive by definition, q^* is not a solution to the equation in (5). Remember that $\tilde{s}$ is decreasing in t and therefore also in q. This implies, the solution in this case must be some $q^{**} > q^*$. This assures that $f_0(s) > f_1(s)$ which makes the LHS positive.

The optimal rule in this model with endogenous information is more conservative than the optimal rule when information is exogenous. To understand this, notice that q^* is such that the decisive voter must have a bias $\lambda(t_d) = \frac{1-\pi}{\pi}$ which is decreasing in π. This means that the decisive voter under q^* is biased towards the action that matches the state that is a priori less likely. What we show here is that since $q^{**} > q^*$ it must be that

$$\lambda\left(t_d(q^{**})\right) > \lambda\left(t_d(q^*)\right)$$

Hence, the bias of the decisive voter according to q^{**} is closer to the state favored by the prior, compared to the decisive voter under q^*. The case of a uniform prior ($\pi = \frac{1}{2}$) offers a good illustration of this. In that case, $q^* = \frac{1}{2}$, which is simple majority. Then, the optimal rule in this model with endogenous information is some kind of super-majority.

4.2 Comparative statics with respect to information

How does the optimal voting rule here change for more informative public signals? It is at this point where Proposition 4 is useful. It allows us to determine whether a particular information structure is more informative than another for any voting rule. Thus, if distributions that satisfy the conditions in this proposition are considered, informativeness remains exogenous to the model and comparative statics with respect to information make sense.

We do this here using the family of power distributions defined in example 3. As we show, this family satisfies the conditions of Proposition 4. In particular,

[9]Rember the definition of $\tilde{s}$ given in lemma 1

$F(s;\alpha)$ becomes more informative for higher values of α. The question is then, how does q^{**} change as α increases?

Remember that the family of distributions we consider is such that $F_0(s;\alpha) = 1-(1-s)^\alpha$ and $F_1(s;\alpha) = s^\alpha$. Plugging this in to the LHS of 5 gives:

$$\frac{s^\alpha}{1-(1-s)^\alpha} \frac{\frac{s^\alpha}{1-(1-s)^\alpha}}{\frac{s^{\alpha-1}\alpha}{1-(1-s)^\alpha} - \frac{s^\alpha(1-\alpha)^{\alpha-1}\alpha}{(1-(1-s)^\alpha)^2}}$$

This can be shown to be decreasing in α. It follows that q^{**} must be increasing in α. In other words, The better the possibly available information, the more conservative should the voting rule be.

One obtains similar results performing this exercise with other families of distributions that can be ordered according to Proposition 4. Whether there is a more general relation between informativeness and the comparative statics on the voting rule in this model remains as a research question for the future.

5 Conclusions

We ask the question of whether an ordering of information structures is possible for a group of like-minded individuals. We answer by saying that only a partial, and very limited, order is possible in general. It can be extended if more restrictions are put on the group's profile of preferences.

What is important to understand is that the notion of "better information" as we understand it for individuals, even in its most restrictive form, as formalized by Blackwell, can not be applied to groups. Even if group members are like-minded, in that they agree on what should be done in a given state of the world. The reason is that uncertainty introduces disagreement. More precise information does not guarantee more agreement. This is why it may be better for the group to be "less informed".

There is an aspect of information we ignore throughout the analysis: it usually comes at some cost. In particular, more precise information is usually more costly. The omission of the costs from acquiring information in our analysis is intentional. It aims at emphasizing how the collective decision making process generates frictions that can reduce the value of information even when there are no other costs to pay. Having established that, the question of how should a group proceed to acquire public information when it is costly, becomes particularly interesting. Public information has some of the characteristics of public goods. But as we show here it may also be a public "bad" for a subset of group members. What mechanism should be used to elicit individuals' valuations and choose among different sources of information and finance it's acquisition? What is the role of the voting rule in such a mechanism? These questions are left as future research.

A Proofs

Proof of Lemma 1.
Agent i chooses $x = 1$ if:

$$\frac{f_1(s)\pi}{f(s;\pi)}u(1,1,t_i) + \frac{f_0(s)(1-\pi)}{f(s;\pi)}u(1,0,t_i) > \frac{f_1(s)\pi}{f(s;\pi)}u(0,1,t_i) + \frac{f_0(s)(1-\pi)}{f(s;\pi)}u(0,0,t_i)$$

$$\frac{f_1(s)}{f_0(s)}\frac{\pi}{1-\pi}\lambda(t_i) > 1 \tag{6}$$

Given the MLRP there must either be a threshold value $\tilde{s}_i(t_i)$ such that the inequality holds for $s > \tilde{s}_i(t_i)$ proving the first part of the lemma, either it always or never holds, and $\hat{\phi}_i(s, t_i)$ is constant. □

Proof of Lemma 2.
This follows directly from inequality 6 in the proof of lemma 1 and the monotonicity assumption on $\lambda(t)$ □

Proof of Lemma 3.
It follows from the monotonicity of $\hat{\phi}(s, t)$ with respect to t. □

Proof of Proposition 1.
From the definition of $F \geqslant_I G$ we get the following inequality:

$$F_0\left(\tilde{s}_F(t_d)\right) - G_0\left(\tilde{s}_G(t_d)\right) \geq \frac{\pi}{1-\pi}\Lambda(u_i, \xi(t))\left[F_1\left(\tilde{s}_F(t_d)\right) - G_1\left(\tilde{s}_G(t_d)\right)\right] \tag{7}$$

where $\Lambda(u_i, \xi(t)) = \frac{\int_T [u_i(1,1,t_i) - u_i(0,1,t_i)]\xi(t_i)dt}{\int_T [u_i(0,0,t_i) - u_i(1,0,t_i)]\xi(t_i)dt}$ represents the group's bias.

Also, from the definition of $\tilde{s}(t)$ it follows that:

$$\frac{h_1(\tilde{s}(t))}{h_0(\tilde{s}(t))} \leq \frac{1-\pi}{\pi}\frac{1}{\lambda(t)}, \ \forall h \in \{f, g\}$$

The RHS of this inequality is always positive and decreasing in t. Then, for $\hat{t}(k) = \{t : \frac{1-\pi}{\pi}\frac{1}{\lambda(t)} = k\}$ it must be that $\tilde{s}(\hat{t}(k)) = \hat{s}(k)$. Thus, if the inequalities stated in the second part of the proposition hold for any k, they must also hold for any $\tilde{s}(\hat{t}(k))$. In such case, then the LHS in (7) is positive and the RHS is negative. This proves the sufficiency of these conditions for the informativeness relationship stated in the proposition.

We prove the "only if" part of the proposition by showing that if the inequalities do not hold, one can always two instances of groups and voting rules such that the informativeness ranking between F and G is different in each case. Consider a group $I = \{1, ..., n\}$ and $q = 0$. Suppose $t_1 < t_2 < ... < t_n$. Then we have $t_d(q = 0) = t_1$. Individuals' preferences are as follows:

$$u_i(1, 1, t_i) = t_i$$
$$u_i(0, 0, t_i) = 1$$
$$u_i(1, 0, t_i) = u_i(0, 1, t_i) = 0$$

for all $i \in I$. If $F \succcurlyeq_I G$, then it must be that:

$$F_0\left(\tilde{s}_F(t_1)\right) - G_0\left(\tilde{s}_G(t_1)\right) \geq \frac{\pi}{1-\pi}\frac{\sum_I t_i}{n}\left[F_1\left(\tilde{s}_F(t_i)\right) - G_1\left(\tilde{s}_G(t_i)\right)\right] \tag{8}$$

Suppose $F_0\left(\tilde{s}_F(t_1)\right) \geq G_0\left(\tilde{s}_G(t_1)\right)$ and $F_1\left(\tilde{s}_F(t_i)\right) \geq G_1\left(\tilde{s}_G(t_i)\right)$. It is clear that inequality(8) may hold for $\sum_{i\neq 1} t_i$ low enough, but not for $\sum_{i\neq 1} t_i$ above a certain threshold. The same argument can be made for $F_0\left(\tilde{s}_F(t_1)\right) \leq G_0\left(\tilde{s}_G(t_1)\right)$ and $F_1\left(\tilde{s}_F(t_i)\right) \leq G_1\left(\tilde{s}_G(t_i)\right)$. This proves the necessity of the inequality conditions stated in the proposition in order for $\succcurlyeq_I$ to be valid for any I and any q. □

Proof of Proposition 2*.*
We show the result for point 1. Point 2 follows from a symmetric argument. We show that given $F \gg G$ and $\frac{g_1(s)}{g_0(s)} > 0$, $\forall s \neq \underline{s}_G$, it must be $\frac{f_1(s)}{f_0(s)} = 0$ for some $s \neq \underline{s}_F$. Suppose not. Then $\inf\{\frac{f_1(s)}{f_0(s)}\} > 0$. There are two cases to consider:
<u>Case 1:</u> $\inf\{\frac{f_1(s)}{f_0(s)}\} > \inf\{\frac{g_1(s)}{g_0(s)}\}$.
Then there exists I such that for some q', $\frac{\pi}{1-\pi}\frac{1}{\lambda(t_d(q'))} \in \left(\inf\{\frac{g_1(s)}{g_0(s)}\}, \inf\{\frac{f_1(s)}{f_0(s)}\}\right)$. From the definition of $\tilde{s}(t)$, this means that:
$\tilde{s}_F(t_d(q')) = \underline{s}_F$ and $arg\inf\{\frac{g_1(s)}{g_0(s)}\} \leq \tilde{s}_G(t_d(q')) \geq \sigma$. This in turn means that $F_0(\tilde{s}_F(t_d(q'))) = F_1(\tilde{s}_F(t_d(q'))) = 0$ and $G_0(\tilde{s}_G(t_d(q'))) > 0$. This inequality holds because given our assumptions, $g_0(\tilde{s}_G(t_d(q')$ must be positive. If s_G is discrete in $[arg\inf\{\frac{g_1(s)}{g_0(s)}\}, \sigma]$ the inequality follows directly. If it is continuous in some interval then it must be true as well for the supremum of that interval. Notice then that $F_0(\tilde{s}_F(t_d(q'))) = 0 < G_0(\tilde{s}_G(t_d(q'))) > 0$ which violates $F \gg G$.
<u>Case 2:</u> $\inf\{\frac{f_1(s)}{f_0(s)}\} < \inf\{\frac{g_1(s)}{g_0(s)}\}$.
Then there exists I such that for some q', $\frac{\pi}{1-\pi}\frac{1}{\lambda(t_d(q'))} \in \left(\inf\{\frac{f_1(s)}{f_0(s)}\}, \inf\{\frac{g_1(s)}{g_0(s)}\}\right)$. From the definition of $\tilde{s}(t)$, this means that:

$\tilde{s}_G(t_d(q')) = \underline{s}_G$ and $arg\inf\{\frac{f_1(s)}{f_0(s)}\} \leq \tilde{s}_F(t_d(q')) \geq \sigma'$. This in turn means that $G_0(\tilde{s}_G(t_d(q'))) = G_1(\tilde{s}_G(t_d(q'))) = 0$ and $F_1(\tilde{s}_F(t_d(q'))) > 0$, for the same reason as with G_0 in case 1. Again this violates $F \gg G$. Thus for $F \gg G$ to be true when $\frac{g_1(s)}{g_0(s)} > 0\,,\ \forall s \neq \underline{s}_G$ it must be $\frac{f_1(s)}{f_0(s)} = 0$ for some $s \neq \underline{s}_F$. This means that for that s $Pr(\theta = 0|s_F) = 1$. □

***Proof of Proposition 3**.*
Given the conditions in the proposition, inequality 7 must hold for any q, as long as $t_d \in [\underline{t}, \bar{t}]$. The restriction on T makes sure of that and thus the result holds. □

***Proof of Proposition 4**.*
The result follows directly by rearranging inequality (7). □

***Proof of Lemma 5**.*
Solving the FOC for the aggregate utility we obtain:
$\frac{f_1(\tilde{s}(t_d(q^*)))}{f_0(\tilde{s}(t_d(q^*)))} = \frac{1-\pi}{\pi}\frac{1}{\Lambda(u_i,\xi(t))}$.
Combining this with inequality 6 in the proof of lemma 1 gives the definition of q^*. Monotonicity with respect to Λ follows from monotonicity of $t_d(q)$ and the MLRP. □

***Proof of Proposition 5**.*
From Lemma 5 we have $\lambda\left(t_d(q^*)\right) = \Lambda\left(u_i, \xi(t)\right)$. Plugging this into inequality (7) gives:
$F_0\left(\tilde{s}_F(t_d(q^*))\right) - G_0\left(\tilde{s}_G(t_d(q^*))\right) \geq \frac{\pi}{1-\pi}\lambda\left(t_d(q^*)\right)\left[F_1\left(\tilde{s}_F(t_d(q^*))\right) - G_1\left(\tilde{s}_G(t_d(q^*))\right)\right]$.
Since we assume $F \geqslant_i G\,,\ \forall\ i \in I$, this inequality must hold, proving the point. □

***Proof of Lemma 6**.*
Note that:

$$E_F[u(x(q,s_F),\theta,t)] \geq E_G[u(x(q,s_G),\theta,t)]$$
$$F_0(\tilde{s}_F(t_d(q)) - G_0(\tilde{s}_G(t_d(q)) \geq \frac{\pi}{1-\pi}\lambda(t)[F_1(\tilde{s}_F(t_d(q)) - G_1(\tilde{s}_G(t_d(q))]$$

For $t = \hat{t}(q)$ this expression holds with equality, i.e. the individual of type $\hat{t}(q)$ is the one that is indifferent between the two distributions. The definition of $\Gamma_q(F,G)$ follows from the monotonicity of $\lambda(t)$ with respect to t. □

Proof of Lemma 7.
Note that $t_d(q) = q$. We know that $x(q, \emptyset) = 1$ if

$$\begin{aligned} E[u_d(1, \theta, q)] &> E[u_d(0, \theta, q)] \\ \pi u_d(1, 1, q) + (1 - \pi)u_d(1, 0, q) &> \pi u_d(0, 1, q) + (1 - \pi)u_d(0, 0, q) \\ \lambda(q) &> \frac{1 - \pi}{\pi} \end{aligned}$$

Remember that q^* is such that $\lambda(q^*) = \Lambda(u, T)$, and by assumption $\Lambda(u, T) = \frac{1-\pi}{\pi}$. This prooves the lemma □

References

[1] Athey, S., and Levin, J. The value of information in monotone decision problems. *MIT Working Papers 98-24* (Dec 2001).

[2] Austen-Smith, D., and Banks, J. Information aggregation, rationality, and the condorcet jury theorem. *The American Political Science Review 90*, 1 (Mar 1996), 34–45.

[3] Austen-smith, D., and Feddersen, T. Deliberation and Voting Rules. In *Social choice and strategic decisions: Essays in honor of Jeffrey S. Banks*, D. Austen-Smith and J. Duggan, Eds. Springer, Berlin, 2005.

[4] Bergemann, D., and Välimäki, J. Information in mechanism design. *Cowles Foundation Discussion Papers*, 1532 (Jan 2005).

[5] Blackwell, D. Comparison of experiments. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability. July 31-August 12, 1950. Statistical Laboratory of the University of California, Berkeley. Berkeley, Calif.: University of California Press, 1951. 666 pp. Editor: Jerzy Neyman* (Jan 1951), 93–102.

[6] Cabrales, A., Gossner, O., and Roberto, S. Entropy and the value of information for investors. *American Economic Review forthcoming* (Jan 2012).

[7] Doraszelski, U., Gerardi, D., and Squintani, F. Communication and Voting with Double-Sided Information. *The B.E. Journal of Theoretical Economics 3*, 1 (2003).

[8] Feddersen, T., and Pesendorfer, W. Voting behavior and information aggregation in elections with private information. *Econometrica 65*, 5 (Sep 1997), 1029–1058.

[9] GANUZA..., J. Signal orderings based on dispersion and the supply of private information in auctions. *Econometrica 78*, 3 (May 2010), 1007–1030.

[10] GERARDI, D., AND YARIV, L. Information acquisition in committees. *Games and Economic Behavior 62*, 2 (Mar 2008), 436–459.

[11] GERSBACH, H. The value of public information in majority decisions. *Economics Letters 35*, 3 (Mar 1991), 239–242.

[12] GERSBACH, H. Allocation of information by majority decisions. *Journal of Public Economics 48*, 2 (Jul 1992), 259–268.

[13] GERSBACH, H. The private and social value of information in majority decisions. *Political Behavior 15*, 1 (Mar 1993), 15–23.

[14] GERSBACH, H. Public information and social choice. *Social Choice and Welfare 17*, 1 (Jan 2000), 25–31.

[15] GERSHKOV, A., AND SZENTES, B. Optimal voting schemes with costly information acquisition. *Journal of Economic Theory 144*, 1 (Jan 2008), 36–68.

[16] JEWITT, I. Information order in decision and agency problems. *Mimeo* (Jan 2007).

[17] LEHMANN, E. Comparing location experiments. *The Annals of Statistics 16*, 2 (Jun 1988), 521–533.

[18] MARSCHAK, J., AND RADNER, R. Economic theory of teams. *Cowles Foundation for Research in Economics at Yale University, Monograph 22. Yale University Press, New Haven and London* (1972).

[19] MESSNER, M., AND POLBORN, M. The option to wait in collective decisions. *Journal of Public Economics 96*, 5-6 (Jun 2012), 524–540.

[20] MORRIS, S., AND SHIN, H. Social value of public information. *The American economic review 92*, 5 (Dec 2002), 1521–1534 (14).

[21] PERSICO, N. Information acquisition in auctions. *Econometrica 68*, 1 (Jan 2000), 135–148.

[22] PERSICO, N. Committee design with endogenous information. *Review of Economic Studies 71*, 1 (Jan 2004), 165–191.

[23] SCHULTE, E. Information aggregation and preference heterogeneity in committees. *Theory and Decision 69*, 1 (Sept. 2008), 97–118.

[24] Strulovici, B. Learning while voting: Determinants of collective experimentation. *Econometrica 78*, 3 (May 2010), 933–971.