INFORMATION ACQUISITION IN COMMITTEES

By

Dino Gerardi and Leeat Yariv

March 2003
Revised: May 2007

COWLES FOUNDATION DISCUSSION PAPER NO. 1411R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Information Acquisition in Committees

DINO GERARDI*AND LEEAT YARIV†‡

May 2007

ABSTRACT. The goal of this paper is to illustrate the significance of information acquisition in mechanism design. We provide a stark example of a mechanism design problem in a collective choice environment with information acquisition. We concentrate on committees that are comprised of agents sharing a common goal and having a joint task. Members of the committee decide whether to acquire costly information or not at the outset and are then asked to report their private information. The designer can choose the size of the committee, as well as the procedure by which it selects the collective choice, i.e., the correspondence between agents' reports and distributions over collective choices. We show that the ex-ante optimal device may be ex-post inefficient, i.e., lead to suboptimal aggregation of information from a statistical point of view. For particular classes of parameters, we describe the full structure of the optimal mechanisms.

JEL classification: D71, D72, and D78.

Keywords: Collective Choice, Mechanism Design, Information Acquisition.

*Department of Economics, Yale University, 30 Hillhouse Avenue, New Haven, CT 06511. e-mail: donato.gerardi@yale.edu

†Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125. e-mail: lyariv@hss.caltech.edu

‡An earlier version of this paper was distributed under the title "Committee Design in the Presence of Communication." We have benefitted from very helpful conversations with Dirk Bergemann, Hongbin Cai, Bernardo Guimaraes, David Levine, Alessandro Lizzeri, Stephen Morris, Nicola Persico, Rann Smorodinsky, and Bill Zame. We are also grateful to an associate editor and two anonymous referees for extremely useful comments and suggestions.

1. Introduction

In many real world situations information is not supplied freely and individuals choose whether to acquire costly information concerning the decision tasks they face. For example, trial jurors need to decide whether or not to attend to testimonies during a trial, referees need to decide whether or not to carefully read a paper under consideration for publication, and top management consultants need to choose the amount of time they invest in learning about different investment opportunities before convening to determine a course of action.

The current paper aims at exploring the potential significance of information acquisition to mechanism design problems. To that effect, we focus on a particular setting in which a committee of homogenous agents, each capable of acquiring costly information, chooses one of two alternatives. The designer chooses the size of the committee and the decision rule in order to maximize the (common) expected utility of the collective decision. Unlike most of the literature on mechanism design in which the distribution of agents' types is exogenously given, the current setup allows for the endogenous determination of agents' types. Specifically, this framework enables us to study mechanism design in situations where there are two forces at play. On the one hand, the mechanism should use the information available as efficiently as possible. On the other hand, the mechanism needs to provide agents with incentives to invest in information (which thereby changes their types).

The analysis of the optimal mechanism yields a few interesting insights. First, in order to provide strong incentives for information acquisition, for a large class of parameters, the optimal device is *ex-post inefficient*, i.e., it does not necessarily utilize all the information that is reported. Second, the optimal ex-post inefficient mechanism is a product of a simple cost-benefit analysis. The designer looks for distortions that maximize the ratio of the (positive) effect on incentives to the (negative) effect on her payoffs. In particular, we can analytically describe the optimal mechanism for extreme values of signal accuracies. Last, the comparative statics of the optimal mechanism exhibit some regularities and irregularities, e.g., the expected social value is monotonic in the cost of information and accuracy of private information, but the optimal committee size is not monotonic in the signals' accuracy.

The paper contributes to the literature on mechanism design with endogenous information. While most of this literature deals with auction and public good models (see, e.g., Auriol and Gary-Bobo [1999], Bergemann and Välimäki [2002], and references therein), there are a few exceptions focusing on collective decision-making. Gersbach [1995] is one of the first papers to study the incentives of committee members to become informed. Gersbach assumes that collective decisions are made according to majority rule and shows that information acquisition is not always socially efficient. Persico [2004] analyzes a problem similar to ours but restricts the designer to threshold voting rules that select one of the alternatives if and only if a certain number of participants support that alternative. In this setup, the optimal threshold rule ends up coinciding with the optimal statistical rule and is, in particular, ex-post efficient. When we allow for a broader class of voting rules, as in the current paper, we see that ex-post efficiency no longer holds.

Li [2001] considers a committee of a fixed size and allows each player to invest in the precision of her private signal. When information is a public good, Li illustrates the optimality of statistical distortions in the decision rule. In Li [2001] investments as well as signals are publicly observed and thereby verifiable. In contrast, in our setup verifiability assures that a non-distortionary rule is optimal when the committee is large enough.

Cai [2003] looks at a continuous framework in which the policy preferences and information structures are captured by normal random variables. Members exert non-verifiable efforts in gathering information, report these preferences to the principal, who then uses the mean decision rule to determine the collective policy. Cai characterizes the optimal committee size in this setting and shows that it is finite. Furthermore, the optimal size is non-monotonic in the variation of preferences of the committee members.

Gershkov and Szentes [2004] consider a problem similar to ours but restrict the set of parameters to be such that the designer is indifferent between the two alternatives when no information is available . They allow the mechanism designer to approach agents sequentially and characterize the optimal stopping rule when the designer is restricted to ex-post efficient outcomes. In their setup, full ex-ante efficiency may yield ex-post inefficiency. In a similar spirit, Smorodinsky and Tennenholtz [2006] allow agents to hold differential costs and characterize environments in which the first best mechanism (corresponding to free information) can be implemented using a potentially sequential mechanism. For some scenarios, they illustrate specific mechanisms implementing the first best solution. The current paper is complementary in that it considers a model germane to situations in which the number of experts is fixed at the outset and communication is not allowed prior to information acquisition. There are many examples that satisfy these restrictions: e.g., a jury in which the size of the jury is transparent but deliberations are allowed only after the testimonies have been presented, a hiring search committee that convenes only after reading candidates' portfolios, etc. In addition, our analysis pertains to a more general environment in that an uninformed decision maker may strictly prefer one alternative over the other.

In broad terms, the current paper adds to the existing literature by introducing a general static mechanism design analysis of problems pertaining to collective choice with information acquisition. Technically, the paper's underlying model is one prevalent in the literature on strategic voting (e.g., Austen-Smith and Banks [1996] and Feddersen and Pesendorfer [1996, 1998]).

The paper is structured as follows. Section 2 presents the design problem. Section 3 illustrates the important features of the optimal mechanism, regarding the way the information agents report is aggregated. Section 4 provides an analysis of the optimal device for extreme signal accuracy levels. It also illustrates various comparative statics results pertaining to the optimal design solution. For the sake of presentation simplicity, throughout most of the paper we assume that agents do not use mixed strategies in the information acquisition stage. In Section 5 we illustrate that the underlying message of the paper, namely that ex-ante optimal mechanisms may be ex-post inefficient, carries through even when agents are allowed to use fully mixed strategies. Section 6 concludes. Propositions' proofs are relegated to the Appendix.

2. The Model

We concentrate on the case replicating the standard committee voting problem (e.g., Feddersen and Pesendorfer [1998]). While our setup is germane to many collective decision environments, the reader may find it useful to trace our modeling choices with a jury metaphor in mind.

There are two states of the world, I (innocent) and G (guilty), with prior distribution $(P(I), P(G))$. The alternatives (or decisions) are A (acquittal), and C (conviction). There is a pool of $N \geqslant 2$ identical agents (the potential jurors). All the agents as well as the mechanism designer share the same preferences which depend on the state of the world and the final decision. Let q be a number in $(0,1)$. The common utility is given by:

$$u(d,\omega) = \begin{cases} -q & \text{if } d = C \text{ and } \omega = I \\ -(1-q) & \text{if } d = A \text{ and } \omega = G \\ 0 & \text{otherwise.} \end{cases}$$

where d and ω denote the collective decision and the state of the world, respectively (using the jury metaphor, preferences are such that jurors prefer to make the right decision and q can be thought of as the threshold of reasonable doubt).

Each agent can purchase a signal of accuracy $p > \frac{1}{2}$. That is, upon paying the cost $c > 0$, the agent receives a signal $s \in \{i, g\}$ satisfying $\Pr(s = i|I) = \Pr(s = g|G) = p$ (each juror has to decide whether to pay attention or not to the testimonies presented during the trial. These testimonies provide a noisy signal concerning the guilt of the defendant).

If more than one agent purchases information, we assume their signals are conditionally independent. As a starting point, we only attend to the case in which an agent can buy at most one signal.

In our environment there are numerous ways to make a collective decision. First, we can have committees of different sizes. Second, for a committee of a given size there is a continuum of ways of aggregating reports into final decisions. Of course, these variables will affect the agents' decisions (whether they acquire information or not, as well as how they report that information) and, therefore, the quality of the final decision. We now analyze the problem of designing the optimal mechanism. To accomplish this, we study the following game.

Stage 1 The designer chooses an *extended mechanism*, i.e., the size of the committee $n \leqslant N$ and a mapping between reports in $\{\emptyset, i, g\}^n$ and probabilities of choosing A or C ($\emptyset$ stands for an agent who does not purchase information, and i or g stand for an agent who purchases information and receives i or g as the realized signals, respectively).

Stage 2 All agents observe the designer's mechanism. Each agent $j = 1, ..., n$ decides whether to purchase a signal. These choices are made simultaneously, and each member of the committee does not observe whether other members have acquired information.[1]

Stage 3 Each agent sends a message in $\{\emptyset, i, g\}$ to the designer, who uses the chosen mechanism to select one of the alternatives.

Stages 2 and 3 constitute an extensive-form game played by the agents $1, ..., n$. For the sake of presentation simplicity, until Section 5 we restrict attention to sequential equilibria in which the players use pure (behavioral) strategies in Stage 2, and are allowed to randomize in Stage 3 (as will become clear shortly, allowing for randomization at Stage 3 does not add complexity to the analysis, hence the apparent asymmetry). A strategy profile of this game determines an outcome (i.e., the probabilities that the correct decision is made in state I and in state G) and therefore, the expected common utility of the decision. The designer chooses the mechanism to maximize her utility (from the decision). In particular, the designer does not take into account the cost c incurred by an agent who purchases a signal. There are different situations in which this assumption is appropriate. The designer may be a CEO who hires a committee of financial advisors. Alternatively, the decision may

[1]Our analysis would, in fact, be tremendously simplified if investments were overt (see Footnote 4). However, in many situations in which agents engage in information acquisition, investment in information is indeed covert and signals are non-verifiable. For example, jurors would have a hard time proving they had attended testimonies, committee members do not check whether their colleagues have gone over the relevant background information before convening, etc.

affect the welfare of every individual in a large society and the designer can be a benevolent planner (e.g., the constitution writers). In this case, any increase in the utility from the decision can compensate for the information costs paid by the agents.

Of course, as will be illustrated formally in what follows, increasing the committee size expands the set of implementable outcomes. In particular, giving the designer the freedom to choose the size of the committee is not crucial. Nonetheless, we view this assumption as appealing in many real world examples in which committee participants are costly (and this cost is independent of the information costs described above). For instance, it is costly to have a juror appear at court instead of at her employment. Suppose that the designer has lexicographic preferences and is willing to pay the additional cost as long as larger committees lead to better decisions. Thus, if two committees of different size lead to the same quality of decisions then the designer strictly prefers the smaller committee. In what follows we assume that for every level of feasible expected payoffs, the designer indeed chooses the smallest committee that generates it.

We denote each agent j's type in Stage 3 by $t_j \in T_j \equiv \{\emptyset, i, g\}$. The designer's problem then constitutes of choosing the size n of the committee and a *device*, a mapping $\gamma : T_1 \times ... \times T_n \to [0, 1]$, where $\gamma(t)$ denotes the probability the designer chooses the alternative C when the vector of reports is t. Each game induced by n and γ generates a set of equilibria. We will use σ_j to denote player j's choice at the information acquisition stage. As already mentioned, in this section we restrict attention to pure strategies in the information acquisition stage. Thus, $\sigma_j \in \{0, 1\}$ where $\sigma_j = 1$ denotes the decision to become informed (in Section 5 we consider the general case $\sigma_j \in [0, 1]$).

Note that, in principle, we could consider an arbitrary set of messages available to each player at Stage 3. Nonetheless, the revelation principle (see Myerson [1991] pages 258-263) assures that it is without loss of generality to restrict attention to mechanisms in which the set of messages of each player coincides with her set of types. Furthermore, the revelation principle assures that it is without loss of generality to assume that players reveal truthfully their types. Intuitively, the mechanism could replicate any garbling of information that would be generated by allowing the agents a larger set of messages or by having them independently alter their reports.

As it turns out, we can put even further restrictions on the mechanisms we consider that are without loss of generality:

1. All players acquire information. Suppose that an outcome is implemented by a pair $(((\sigma_1, \ldots, \sigma_n), \gamma))$ in which only players $1, ..., n'$ acquire the signal (and reveal it truthfully), where $n' < n$. Consider the device γ' defined as follows: for every vector of reports $t_1, ..., t_{n'}$, $\gamma'(t_1, ..., t_{n'}) = \gamma(t_1, ..., t_{n'}, \emptyset, ..., \emptyset)$. Under the original pair $(((\sigma_1, \ldots, \sigma_n), \gamma))$, the first n' players know that players $n'+1, ..., n$ do not purchase the signal and report message $\emptyset$ to the device γ (remember that γ induces truthful revelation). If players $1, ..., n'$ decide to acquire information and be sincere under γ then they have an incentive to do the same under γ'. Therefore, in the remainder of the section we focus on devices that induce all players to acquire information and reveal it sincerely. We call these devices *admissible*. It is important to note that admissible devices are characterized by two classes of incentive compatibility constraints. The first is the already introduced truthful revelation constraint. The second guarantees that each player best responds by acquiring information.

2. Players' message space is binary. Since we consider the case in which all players acquire information we can use the revelation principle to further restrict the set of messages. In fact, it is without loss of generality to constrain messages to be either i or g.

Let $U_j(t_j, t'_j)$ denote the expected utility (from the decision) of player j when her type is $t_j = i, g$, she reports message $t'_j = i, g$, and all her opponents acquire information and are sincere. Let also $\Pr(t_j)$ denotes the probability that agent j will observe signal $t_j = i, g$ if she acquires information. It follows that agent j will purchase the signal if and only if the following *information acquisition* constraints are satisfied:

$$\Pr(g)(U_j(g,g) - U_j(g,i)) \geqslant c, \tag{1}$$

$$\Pr(i)\left(U_j(i,i) - U_j(i,g)\right) \geqslant c. \tag{2}$$

These inequalities guarantee that agent j prefers to buy the signal and be sincere rather than not buy the signal and always report one of $s = i, g$ (i in the first inequality, g in the second one).

Clearly, (1) and (2) also guarantee that a player who purchases a signal has an incentive to reveal it truthfully. We can, therefore, think of an admissible device as a mapping $\gamma : \{i,g\}^n \to [0,1]$ which satisfies conditions (1) and (2).

3. The optimal device is symmetric. Let Λ_n denote the set of permutations on $\{1,...,n\}$. An admissible device γ is *symmetric* if for all $(t_1,...,t_n) \in \{i,g\}^n$ and all $\varphi \in \Lambda_n$, $\gamma(t_1,...,t_n) = \gamma\left(t_{\varphi(1)},...,t_{\varphi(n)}\right)$. In a symmetric device, the probability that the defendant is convicted depends only on the number of messages g (or i) but not on the identity of the players who send g. Suppose that γ is an admissible device. For any $\varphi \in \Lambda_n$ consider the device γ_φ, where $\gamma_\varphi(t_1,...,t_n) = \gamma\left(t_{\varphi(1)},...,t_{\varphi(n)}\right)$ for every $\{t_1,...,t_n\}$ in $\{i,g\}^n$. Since all players are identical and γ is admissible, the device γ_φ is also admissible and outcome equivalent to γ. It follows that the symmetric device $\tilde{\gamma} = \left(\frac{1}{|\Lambda_n|}\right)\sum_{\varphi\in\Lambda_n}\gamma_\varphi = \frac{1}{n!}\sum_{\varphi\in\Lambda_n}\gamma_\varphi$ is admissible and outcome equivalent to the original device γ. We thereby consider only symmetric devices henceforth.

A symmetric device can be represented as a mapping $\gamma : \{0,1,...,n\} \to [0,1]$, where $\gamma(k)$ denotes the probability that the defendant is convicted (alternative C is chosen) when k players report the guilty signal g (each player can report either i or g). For a symmetric device $\gamma : \{0,1,...,n\} \to [0,1]$ conditions (1) and (2) can be expressed as:

$$\sum_{k=0}^{n-1}\binom{n-1}{k} f(k+1;n)\left(\gamma(k+1) - \gamma(k)\right) \geqslant c, \tag{ICi}$$

$$\sum_{k=0}^{n-1}\binom{n-1}{k} f(k;n)\left(\gamma(k) - \gamma(k+1)\right) \geqslant c, \tag{ICg}$$

where $f(\cdot;n):\mathbb{R}\to\mathbb{R}$ is defined by:

$$f(x;n) = -qP(I)(1-p)^x p^{n-x} + (1-q)P(G)p^x(1-p)^{n-x}.$$

In order to get an intuitive sense of each of these formulas, notice that

$$\Pr(\omega = I \mid |\{k \mid s_k = g\}| = x) = \frac{P(I)(1-p)^x p^{n-x}}{P(I)(1-p)^x p^{n-x} + P(G)p^x(1-p)^{n-x}}.$$

Thus, $f(x;n)$ is a proxy for the difference in expected payoff between choosing alternatives A and C when x out of n signals are equal to g. In particular, the designer would like to choose the alternative C whenever $f(x;n) \geqslant 0$ (the designer is, in fact, indifferent when equality holds). For any n, $f(x;n)$ is strictly increasing in x.

The incentive constraints are derived by weighing these differences between the returns to both alternatives by the probabilities of conviction for different constellations of signal realizations (where the binomial coefficients account for the number of ways in which each k guilty reports can be allocated among $n-1$ players).

For each n, we look for the optimal admissible device, i.e., the admissible device that maximizes the expected utility of the decision. This amounts to solving the following linear programming problem P_n :

$$\max_{\gamma:\{0,\ldots,n\}\to[0,1]} -(1-q)P(G) + \sum_{k=0}^{n}\binom{n}{k}f(k;n)\gamma(k)$$

$$\text{s.t. } (ICi),\ (ICg).$$

We denote by $\bar{\gamma}_n$ the solution to the problem P_n (if it exists), and by $V(n)$ the expected utility of the optimal device. If P_n does not have any feasible solution, we set $V(n) = -1$.

The optimal mechanism consists of the optimal size of the committee n^*, and the optimal admissible device $\bar{\gamma}_{n^*}$. n^* is such that $V(n^*) \geqslant V(n)$, for every nonnegative integer $n \leqslant N$.[2] In what follows, we analyze the optimal pair $(n^*, \bar{\gamma}_{n^*})$.

[2]Notice that $V(0) > -1$, and, thus, $n \geqslant 1$ can be the optimal size only if problem P_n admits a feasible solution.

3. Features of The Optimal Extended Mechanism

The extended mechanism the designer chooses is comprised of the size of the committee as well as the aggregation rule and, hence, the incentive scheme it will operate under. In this section we illustrate some fundamental traits of the optimal device the designer would choose. In particular, we show that imperfectly aggregating the available information may induce more players to acquire information, thereby yielding a higher overall expected utility level.

In order to analyze how the optimal device uses the information of the agents, we first consider the case in which the designer makes the final decision after observing n free signals. This will give us an upper bound on what the designer can achieve when she chooses a committee of size n and information is costly. To that effect, we simply need to maximize the objective function of problem P_n (without the constraints). We let γ_n^B denote the solution to this maximization problem. The device γ_n^B, which we term the *Bayesian* device, is of the form:

$$\gamma_n^B(k) = \begin{cases} 0 & \text{if } f(k;n) < 0 \\ 1 & \text{if } f(k;n) \geqslant 0 \end{cases},$$

(in fact, when $f(k;n) = 0$, $\gamma_n^B(k)$ can be any number in the unit interval).

To interpret this result, notice that $f(k;n)$ is positive (negative) if and only if the cost of convicting the innocent q is smaller (greater) than the probability that the defendant is guilty given that k of n signals are g.

Recall that the function $f(\cdot;n)$ is increasing and that $z(n)$ is defined by the equality $f(z(n);n) = 0$. We have:

$$z(n) = \frac{1}{2}\left(n + \frac{\ln\left(\frac{qP(I)}{(1-q)P(G)}\right)}{\ln\left(\frac{p}{1-p}\right)}\right). \tag{3}$$

Let k_n denote the smallest integer greater than or equal to $z(n)$. Another way to express the Bayesian device γ_n^B is:

$$\gamma_n^B(k) = \begin{cases} 0 & \text{if } k < k_n \\ 1 & \text{if } k \geqslant k_n \end{cases}.$$

For small values of n, $z(n)$ can be negative or greater than n. In the first case, the optimal decision is always to convict the defendant. In the latter case, the optimal decision is always to acquit.[3] For large values of n, however, $z(n)$ is positive and smaller than n ($z(n)/n$ converges to $1/2$ as n goes to infinity).

We let $\hat{V}(n)$ denote the expected utility of the Bayesian device:

$$\hat{V}(n) = -(1-q)P(G) + \sum_{k\in\{0,..,n\},\ k>z(n)} \binom{n}{k} f(k;n).$$

The utility $\hat{V}(n)$ is non-decreasing in the number of signals n. Moreover, $\hat{V}(n)$ is strictly greater than $V(0)$, the expected utility of the optimal uninformed decision, if and only if $z(n)$ belongs to $(0,n)$. If $z(n)$ is not in $(0,n)$, then $V(0) = \hat{V}(1) = ... = \hat{V}(n)$.

When n becomes unboundedly large, the Bayesian device uses an infinitely increasing number of i.i.d. signals. The law of large numbers ensures that all uncertainty vanishes asymptotically. In particular, $\hat{V}(n)$ converges to zero, the no uncertainty value, when n goes to infinity.[4]

We now return to the original design problem. In what follows, we will mainly focus on $n > 1$.[5] Clearly, the expected utility of the optimal admissible device $V(n)$ cannot be

[3]These cases arise when the designer is very concerned with a particular mistake (acquitting the guilty ($q \approx 0$) or convicting the innocent ($q \approx 1$)), and the signal is not very accurate, i.e., p is close to $1/2$. In both cases the n signals are of no value.

[4]Note that if information acquisition is overt and $c < 1$, then $\hat{V}(n)$ is implementable (in Nash equilibrium) for sufficiently large $n \leqslant N$. Indeed, assume the designer selects the Bayesian device γ_n^B as long as everyone purchases information, and a device γ that makes a choice contrary to the Bayesian prescription if any agent does not purchase information (i.e., for all k, $\gamma(k) = 1-\gamma_n^B(k)$). The strategy profile under which all players acquire information and are always sincere constitutes a Nash equilibrium. Under this profile, the expected utility of the decision approaches 0. If one player deviates and does not acquire information, she drives the common utility to a level that approaches -1. Finally, no agent has an incentive to lie upon acquiring information.

[5]This is because the case of $n = 1$ is trivial in the sense that either the Bayesian device is admissible and, hence, optimal, or there is no mechanism inducing the single agent to acquire information.

greater than $\hat{V}(n)$. On the other hand, when the Bayesian device γ_n^B is admissible, we have $V(n) = \hat{V}(n)$. In this case the designer is able to give the incentive to the n agents to acquire the signal and, at the same time, to make the best use of the available information. Proposition 1 shows that this can happen if and only if the contribution of the last signal to the utility of a single decision maker is greater than or equal to its cost.

Proposition 1. *For every $n \geqslant 1$, $V(n) = \hat{V}(n)$ if and only if $\hat{V}(n) - \hat{V}(n-1) \geqslant c$.*

The proof (which appears in Gerardi and Yariv [2003]) is straightforward and is thus omitted. However, it is useful to see why the above inequality is necessary. Consider a committee of size n and suppose that players $2, \ldots, n$ acquire the signal and announce it truthfully. Notice that either $k_n = k_{n-1}$ or $k_n = k_{n-1} + 1$. If $k_n = k_{n-1}$ ($k_n = k_{n-1} + 1$) then player 1 can achieve utility $\hat{V}(n-1)$ without acquiring the signal by simply announcing signal i (g). Indeed, the defendant will be convicted only if k_{n-1} opponents of player 1 observe signal g. Thus, player 1 can generate an expected utility of $\hat{V}(n-1)$ by not purchasing and reporting the appropriate signal. As it turns out, when $k_n = k_{n-1}$ ($k_n = k_{n-1} + 1$) the LHS of $IC(i)$ of the Bayesian device is smaller (larger) than the LHS of $IC(g)$.[6] The result then follows.

We assume that there exists at least one integer for which the Bayesian device is admissible. Let n^B denote the greatest such integer. That is, $\hat{V}\left(n^B\right) - \hat{V}\left(n^B - 1\right) \geqslant c$, and $\hat{V}(n) - \hat{V}(n-1) < c$ for every $n > n^B$. The existence of n^B is guaranteed by the fact that the sequence $\left\{\hat{V}(1), ..., \hat{V}(n), ...\right\}$ converges (to zero). The designer can induce more than n^B players to acquire information only if she selects a device that aggregates the available information suboptimally. On the other hand, more information will be available in larger committees. How should the designer solve this trade-off? Is the optimal size of the committee equal to or larger than n^B?

Note that if the optimal size of the committee is n^B, the designer could restrict herself to a simple class of devices characterized by threshold voting rules (the optimal threshold

[6]Note that this observations implies, in particular, that the first constraint to bind as the cost c increases alternates with n.

being determined by the Bayesian device, as in Persico [2004]). On the other hand, if the optimal size is larger than n^B, then mixing, and in particular ex-post inefficiency, may play a very important role in the choice of ex-ante optimal mechanisms. Proposition 2 shows that this is indeed the case (at least when the cost is sufficiently low).[7]

Before formally stating the result, we need to introduce one technical assumption. We say the environment is *regular* if $\frac{\ln\left(\frac{qP(I)}{(1-q)P(G)}\right)}{\ln\left(\frac{p}{1-p}\right)}$ is not an integer. This implies that for all n, $z(n)$, the Bayesian threshold value, is not an integer. In a regular environment, if n is such that $\hat{V}(n) > V(0)$, then for all $n' \geqslant n$, $\hat{V}(n'+1) > \hat{V}(n')$. Note that the environment is, in fact, generically regular.

Proposition 2. *Fix $P(I)$, q and p and assume the environment is regular. Let $n^*(c) \leqslant N$ denote the optimal size of the committee when the cost of acquiring information is c. There exists $\bar{c} > 0$ such that for every $c < \bar{c}$, whenever $\hat{V}(N) - \hat{V}(N-1) < c$, then $V(n^*(c)) < \hat{V}(n^*(c))$.*

The condition $\hat{V}(N) - \hat{V}(N-1) < c$ is equivalent to requiring that $N > n^B$ for the parameters at hand. Note that whenever $\hat{V}(N) - \hat{V}(N-1) \geqslant c$, the Bayesian device with N agents is admissible and clearly optimal.

To understand the intuition of the proposition, it is perhaps easiest to consider a cost c such that $c = \hat{V}(n^B+1) - \hat{V}(n^B) + \varepsilon$ for a small $\varepsilon > 0$, so that the Bayesian device with $n^B + 1$ players is "almost implementable." We now proceed in two steps. First, we explain how small distortions to the Bayesian device with $n^B + 1$ agents can generate a device inducing $n^B + 1$ agents to acquire information. Second, we show that for sufficiently small ε, such distorted devices yield a higher expected utility than the Bayesian device. Indeed, consider the Bayesian device with $n^B + 1$ agents, and for the sake of illustration, assume

[7] Note that there is an alternative way to interpret probabilistic choices. Namely, one could think of a designer who delegates the decision to the committee. The designer chooses a voting rule (e.g., majority or unanimity in the jury setup) and allows participants to deliberate inbetween acquiring information and casting votes. Using Gerardi and Yariv [2007a], the optimal device γ described in our analysis can be interpreted as the optimal communication protocol (see Gerardi and Yariv [2003] for an elaboration).

that $z(n^B+1)$ is very close to (though greater than) $k_{n^B+1}-1=k_{n^B}$ so that the $IC(g)$ constraint is binding.

Consider an agent contemplating the two alternatives: acquiring a signal and reporting an uninformed message g. Under the Bayesian device, the difference arises only from events in which precisely k_{n^B} of the other agents report the message g (so that our agent is pivotal) and signal i is observed by the agent. The difference between reporting i and g in that case can be arbitrarily close to 0 as $z(n^B+1)$ approaches k_{n^B} (formally, recall that this cost is captured by $f(k_{n^B};n^B+1)\approx f(z(n^B+1);n^B+1)\equiv 0)$. Thus, the agent would have an incentive to save the cost c and report an uninformed message g (in particular, this explains why the binding constraint is $IC(g)$).

We now look for a device that makes reporting an uninformed message g more costly. Let us consider the device identified with the threshold k_{n^B}. Again, we compute the difference between acquiring a signal and reporting an uninformed message g. Note that now the difference arises from events in which precisely $k_{n^B}-1$ of the agent's peers report the message g and she observes the signal i. The difference between reporting i and g in that case is bounded away from 0 (formally, recall that $f(x;n)$ is a strictly increasing function of x and so the proxy for the cost of reporting g in lieu of i is given by $f(k_{n^B}-1;n^B+1)<f(k_{n^B};n^B+1)\approx 0)$.

Finally, consider combinations of these two devices, which boil down to acquitting when less than k_{n^B} agents report g, convicting (with probability 1) when $k_{n^B}+1$ or more agents report g, and convicting with some probability α when precisely k_{n^B} agents report g. When ε is sufficiently small, a low value of α assures that both constraints are satisfied.[8]

Assuming that $z(n^B+1)$ is sufficiently close to k_{n^B} allowed us to illustrate the feasibility of the above distorted device regardless of the probabilities of an agent being pivotal under the thresholds k_{n^B} and $k_{n^B}+1$. As it turns out, for arbitrary $z(n^B+1)$, this intuition carries through as long as n^B is large enough, or equivalently, c is sufficiently small.

[8]Indeed, for sufficiently low α, $IC(i)$ is satisfied since it is not binding at $\alpha=0$ (corresponding to the Bayesian device). The comparison above assures that for sufficiently small ε, there exists $\alpha(\varepsilon)$ such that for all $\alpha>\alpha(\varepsilon)$, the $IC(g)$ is satisfied as well and $\alpha(\varepsilon)\to 0$ as $\varepsilon\to 0$.

To summarize, there exists a distorted device that induces the n^B+1 agents to acquire information. The size of the distortion depends on ε (in particular, the distortion is small when ε is close to zero). This implies that for ε small the utility of the above distorted device with n^B+1 agents is almost equal to (but smaller than) $\hat{V}(n^B+1)$, the utility of the Bayesian device with n^B+1 agents. Clearly, $\hat{V}(n^B+1)$ is bounded away from the utility of the Bayesian device with n^B agents. It follows that for ε sufficiently small, when n^B+1 players are available, the designer is strictly better off using the distorted device. In the proof of Proposition 2 we extend this intuition to a wide range of cost values c and to arbitrary $z(n^B+1)$.[9]

Although we do not have a necessary condition for the optimal size to be greater than n^B, notice that the result in Proposition 2 cannot be extended to all values of c. It is possible to construct examples in which the optimal size coincides with the Bayesian size.[10]

With regards to the optimal size of the committee, Proposition 1 in Gerardi and Yariv [2003], as well as Theorem 3 in Al-Najjar and Smorodinsky [2000] imply that for any constellation of parameters, there exists an upper bound $\bar{N}$ such that whenever there are at least $\bar{N}$ available agents (i.e., $N \geqslant \bar{N}$), $n^*(c) = \bar{N}$. This observation stands in contrast to the underlying message of the information aggregation literature (see, e.g., Feddersen and Pesendorfer [1996, 1997]) in which a large pool of agents yields complete aggregation of all of the available information. The contribution of the result in Gerardi and Yariv [2003] lays in the bounds it provides on the maximal n for which P_n has a feasible solution in our setup, which are potentially useful for computational reasons.[11]

[9] A similar intuition holds for sequential mechanisms and indeed Gershkov and Szentes [2004] show that distorted sequential mechanisms may be ex-ante optimal.

In Li [2001] the agents invest in the accuracy of signals and signals are publicly observed. Suppose that uninformed agents prefer to convict. An aggregation rule that requires stronger evidence in favor of conviction induces agents to invest more in information potentially supporting their preferred alternative and is thus beneficial for incentives. This is the driving intuition behind Li's results. Our underlying setup is very different as are our results. In particular, the direction of the distortion (in favor of acquittal or conviction) depends on which constraint is binding.

[10] E.g., consider the environment $P(I) = P(G)$, $p = 0.85$, $q = 0.52$, and $c = 0.035$ (which is regular). Whenever $N \geqslant 3$, the optimal size coincides with the Bayesian size and equals 3.

[11] Indeed, it is interesting to note that the fact that $P_{\tilde{n}}$ does not have any feasible solution does not imply that P_n does not have any feasible solution for all $n > \tilde{n}$. For example, for $P(I) = P(G) = \frac{1}{2}$, $p = 0.8$,

4. Comparative Statics

In this section we analyze how the optimal extended mechanism and the quality of the decision depend on the primitives of the model. We first provide some analytical results on the structure of the optimal device for extreme values of information accuracy p. We then look at the impact that changes in the information cost c and in the accuracy of the signal p have on the expected utility of the designer and on the optimal size of the committee.

The Optimal Device

As was illustrated in Proposition 2, for certain parameters, the optimal device does not coincide with the corresponding Bayesian device. The underlying intuition is that distortions may increase the agents' incentives to acquire information. Proposition 2 illustrates that this effect is stronger than the statistical efficiency loss.

As it turns out, for extreme values of signal accuracy, we can actually describe the shape of the distortions precisely.

We start with the case of very accurate signals. For any committee of size $n \leqslant N$ we look for the optimal mechanism in which all n agents acquire. Recall that under the Bayesian device the defendant is convicted if and only if at least k_n agents observe a guilty signal g, where k_n is the smallest integer greater than or equal to $z(n)$ (defined in equation 3). Note that when p is sufficiently large, $k_n \in \left\{\frac{n}{2}, \frac{n+1}{2}, \frac{n}{2}+1\right\}$ and the Bayesian device essentially takes the form of a majority rule.

When the cost c is sufficiently small the Bayesian device is admissible and, clearly, optimal. As we increase c, we reach a threshold in which (generically) one of the constraints binds and distortions need to be introduced. As it turns out, the optimal device distorts away from the middle and is weakly increasing between 0 and $k_n - 1$ and between k_n and n. Formally, we have the following proposition:

Proposition 3. *Fix $n \geqslant 2$, $P(I)$ and q, and assume that the environment is regular. There exists $\bar{p}_n \in \left(\frac{1}{2}, 1\right)$ such that for each $p > \bar{p}_n$ the following holds. For any $c > 0$, the optimal*

$q = 0.7$, and $c = 0.032$, P_6 and P_7 have feasible solutions, while P_5 does not.

device in which all players acquire (if it exists) takes the form:

$$\begin{aligned}&\bar{\gamma}_n(0)=\ldots=\bar{\gamma}_n(k'-1)=0, \quad \bar{\gamma}_n(k')=\alpha, \quad \bar{\gamma}_n(k'+1)=\ldots=\bar{\gamma}_n(k_n-1)=1,\\ &\bar{\gamma}_n(k_n)=\ldots=\bar{\gamma}_n(k''-1)=0, \quad \bar{\gamma}_n(k'')=\beta, \quad \bar{\gamma}_n(k''+1)=\ldots=\bar{\gamma}_n(n)=1,\end{aligned}$$

where $\alpha,\beta\in[0,1]$, *and* $0<k'<k_n\leqslant k''<n$.[12]

The proof is essentially a cost-benefit analysis of all potential distortions. Indeed, we look for the distortion that maximizes the ratio of "effect on binding constraint(s)" to "effect on designer's payoff." The proof of Proposition 3 therefore requires the comparison of a large set of potential distortions and may be found in a technical addendum, Gerardi and Yariv [2007b].

For the sake of brevity, in the Appendix, we prove Proposition 3 for the case in which c is larger but sufficiently close to $\hat{V}(n)-\hat{V}(n-1)$ (so that the Bayesian device with n players is "almost" implementable). In that case, say the binding constraint is $IC(i)$, so that $k_n=k_{n-1}$. We show that the optimal device entails only distortions at k_n. Intuitively, when p is sufficiently large, it is optimal to make the ex-post inefficient decision when the number of agents who report signal i is almost identical to the number of agents who report g. Recall, in fact, that when p is large k_n is close to $\frac{n}{2}$. Clearly, when p is sufficiently close to one it is very unlikely that each signal is observed by half of the agents. This suggests that the effect of a distortion at k_n on both the designer's payoff as well as on the $IC\,(i)$ constraint converges to zero as p goes to one. However, in the Appendix we show that the convergence of the effect on the designer's payoff is much faster. Similar considerations hold for the more general case pertaining to Proposition 3 when we compare distortions at arbitrary points k and k'.

So far, we have considered the optimal device corresponding to a given size n. Consider now the extended mechanism design problem in which the committee size is a choice parameter. First, for any fixed N, for p sufficiently large (namely, $p>max\{\bar{p}_2,...,\bar{p}_N\}$), it

[12]Of course, the case $\bar{\gamma}_n(k'+1)=\ldots=\bar{\gamma}_n(k_n-1)=1$ is relevant only when $k'+1\leqslant k_n-1$. The same observation applies to the adjacent interval on which $\bar{\gamma}_n$ is constant.

follows that the optimal solution to the *extended* mechanism design problem takes the shape of the devices described in Proposition 3. Second, while this in itself does not rule out the possibility of the optimal device being non-distortionary, Proposition 2 can in fact be readily extended to the case in which N is fixed and p is sufficiently close to 1, under mild conditions. Namely, assume that $qP(I) > 2(1-q)P(G)$ or that $qP(I) < \frac{1}{2}(1-q)P(G)$. There exists $\tilde{p}$ such that for any $p > \tilde{p}$, the following holds. If the Bayesian device with n agents, $2 \leqslant n < N$, is admissible then there exists an admissible distorted device with $n+1$ players that yields expected utility greater than $\hat{V}(n)$ (see proof in Gerardi and Yariv [2007b]).

To summarize, fix N and assume that the above conditions are satisfied, and pick p sufficiently large. Finally, assume that $c > \hat{V}(N) - \hat{V}(N-1)$ and that there exists $n = 2, ..., N-1$ for which $c < \hat{V}(n) - \hat{V}(n-1)$, so that the Bayesian device is not admissible with the entire set of N accessible agents, but admissible for some non-trivial subset. The solution to the extended mechanism design problem: a. entails a distortionary device; and b. takes the form described in Proposition 3.

We now turn to the other extreme case of inaccurate signals. Admittedly, this case is perhaps of less interest than the previous one since as signals become extremely uninformative, small committees generate precisely the same expected utility as the fully uninformed choice. Consequently, the design problem boils down to a comparison between the utility generated by large groups of agents and that generated by the uninformed decision.

Specifically, consider the case in which the optimal uninformed decision is A (the analysis of the complementary case is analogous). Thus, we assume that

$$y \equiv \frac{qP(I)}{(1-q)P(G)} > 1.$$

Clearly, if p is sufficiently close to $\frac{1}{2}$, the aggregate information is outweighed by the prior and the designer would have no desire to induce the agents to acquire information. Notice that if $p = \frac{y^{1/N}}{1+y^{1/N}} > \frac{1}{2}$ then $z(N) = k_N = N$. When p is slightly above $\frac{y^{1/N}}{1+y^{1/N}}$, then $N-1 < z(N) < N$ and the Bayesian device convicts the defendant only when all N agents observe the guilty signal g. In that case, utilizing a group of $n < N$ agents cannot generate a higher expected utility than the uninformed decision. In what follows we assume that p is small but larger than $\frac{y^{1/N}}{1+y^{1/N}}$. Thus, we focus on the extreme case in which the designer faces a trade-off between implementing the uninformed decision and using the entire set of accessible agents.

Of course, if $c \leqslant f(N;N)$ the Bayesian device with N agents can be implemented and would constitute the optimal solution. As c increases above the critical level $f(N;N)$ the first constraint to bind for the Bayesian device is $IC(i)$. Our next proposition shows that when c is slightly above $f(N;N)$ it is optimal to introduce a distortion at $k=0$. It then follows from continuity that for sufficiently low cost levels, utilizing N agents by ways of a distorted device would indeed yield greater expected utility than the uninformed decision.

Proposition 4. *Fix $P(I)$ and q with $y > 1$. There exists $\underline{p} \in \left(\frac{y^{1/N}}{1+y^{1/N}}, 1\right)$ such that for each p, $\frac{y^{1/N}}{1+y^{1/N}} < p < \underline{p}$, the following holds. There exists $\varepsilon > 0$ such that for $c \in (f(N;N), f(N;N)+\varepsilon)$ the optimal device with N agents takes the form:*

$$\bar{\gamma}_N(k) = \begin{cases} \alpha & \text{for } k=0 \\ \gamma_N^B(k)=0 & \text{for } k=1,\ldots,N-1 \\ \gamma_N^B(k)=1 & \text{for } k=N \end{cases}$$

where $\alpha \in (0,1]$ is such that the $IC(i)$ constraint is satisfied with equality.

The proof of Proposition 4 follows the same logic as the proof of Proposition 3. We show that when p is sufficiently small the ratio of "effect on $IC(i)$" to "effect on designer's payoff" is maximized at $k=0$.[13]

[13] In fact, much in the spirit of Proposition 3, it is also possible to show that when p is within the range of Proposition 4, increasing further the cost c yields an optimal mechanism of the following structure:

$$\begin{aligned} \bar{\gamma}_N(0) = \ldots = \bar{\gamma}_N(k'-1) = 1, \quad & \bar{\gamma}_N(k') = \alpha, \quad \bar{\gamma}_N(k'+1) = \ldots = \bar{\gamma}_N(k''-1) = 0, \\ & \bar{\gamma}_N(k'') = \beta, \quad \bar{\gamma}_N(k''+1) = \ldots = \bar{\gamma}_N(N) = 1, \end{aligned}$$

where $\alpha, \beta \in [0,1]$, and $0 \leqslant k' < k'' \leqslant N-1$.

The Cost of Information

The first, obvious, result is that the expected utility of the optimal device is decreasing in the cost of information acquisition. Indeed, for any given size of the committee, if a device is admissible when the cost is c, then the device is also admissible when the cost is lower than c.

We now consider how the optimal size is affected by a change in the information cost. Clearly, given any cost c with optimal size $n^*(c) < N$, we can always find another cost c', sufficiently lower than c, such that $n^*(c')$ is greater than $n^*(c)$ (it is enough to choose c' such that the Bayesian device γ_n^B is admissible for some n greater than $n^*(c)$). Unfortunately, it is less straightforward to perform the comparative statics for small changes of the information cost. In all the examples we have constructed, the optimal size decreases when the information cost increases. However, we have not been able to prove that this is a general result. To illustrate why it is difficult to obtain analytical results, consider two committees of size n and $n+1$. For any cost c, consider the difference between the utility of the optimal device at $n+1$ and the utility of the optimal device at n. It is possible to construct examples such that this difference is positive for low and high values of the cost, but is negative for intermediate values (in a sense, the utility does not exhibit a single crossing property). In other words, suppose we start with a level of the cost such that size $n+1$ is better than size n. In general, we cannot conclude that this relation holds when the cost becomes smaller. Therefore, it remains an open question whether the optimal size is indeed always decreasing in the cost or not.

The Signals' Accuracy

For a committee of a given size n, it is straightforward to show that the utility of the optimal device increases when the quality of the signal improves. Intuitively, when signals are more accurate, the designer can always garble the reported information and replicate the optimal device corresponding to less accurate signals.[14]

[14] See the exact construction in Gerardi and Yariv [2003].

Note that while in our model the designer always benefits from a more informative signal, this is not necessarily the case when restricting the aggregation devices to be ex-post efficient. In that case, the designer can induce n agents to acquire information if and only if $\hat{V}(n) - \hat{V}(n-1)$ is greater than the information cost. When the signal is very informative, $\hat{V}(n)$ is close to zero for n relatively small, and therefore it is impossible to induce many players to acquire information. In contrast, when the signal is less accurate, the difference $\hat{V}(n) - \hat{V}(n-1)$ can be larger than the cost for large values of n. It is possible for the designer to prefer having many uninformative signals over a few very accurate ones.[15]

As far as the optimal size of the committee is concerned, several examples indicate that it is not monotonic in p.[16]

5. Mixed Strategies

So far, we have assumed that agents' decisions in Stage 2, the information acquisition stage, are pure binary and cannot be random. Of course, allowing agents to mix expands the set of equilibria corresponding to each extended mechanism. Consequently, the maximal expected payoffs generated by ex-post efficient mechanisms could conceivably be higher when random acquisition is allowed. It is therefore important to check whether the results regarding ex-post inefficiency of ex-ante optimal mechanisms continue to hold true when agents' strategies are unrestricted. In this section we show that the result of Proposition 2 extends to the case of mixed strategies.

We consider our extended mechanism problem (with $n \leqslant N$ players) and allow each agent j to acquire the signal with probability $\sigma_j \in (0,1]$. As explained in Section 2 we can disregard the case $\sigma_j = 0$ because equilibria in which some players do not acquire the signal can be "replicated" by committees of smaller size in which all agents become informed with positive probability.

[15] To give a concrete example, let us assume that $P(I) = \frac{1}{2}$, $q = 0.82$, and $c = 0.0013$. When $p = 0.85$, the Bayesian size is $n^B = 10$ while for $p = 0.95$, the Bayesian size is $n^B = 4$. Furthermore, the expected utility corresponding to the optimal device when $p = 0.85$ is higher than that corresponding to $p = 0.95$.

[16] Consider, for instance, the case $P(I) = \frac{1}{2}$, $q = 0.62$, and $c = 0.004$. For any $N \geqslant 24$, the optimal size is 13 for $p = 0.55$, 24 for $p = 0.65$, and 15 for $p = 0.75$.

In Section 2 we also pointed out that if a player j acquires the signal with probability $\sigma_j = 1$ then we can, without loss of generality, assume that the set of messages available to j is $\{i, g\}$. Of course, this does not hold for a player j who randomizes with probability $\sigma_j < 1$. In this case, each type $t_j \in \{\emptyset, i, g\}$ has positive probability and the set of messages of j must be $\{\emptyset, i, g\}$. Therefore, it is enough to restrict attention to pairs (σ, γ) of the following form. $\sigma = (\sigma_1, \ldots, \sigma_n)$ denotes the players' strategies at the information acquisition stage (since all players are ex-ante identical, without loss of generality, we restrict attention to profiles in which $\sigma_1 \leqslant \ldots \leqslant \sigma_n$). γ is a mapping from $M = M_1 \times \ldots \times M_n$ into $[0, 1]$ where $M_j = \{i, g\}$ if $\sigma_j = 1$ and $M_j = \{\emptyset, i, g\}$ if $\sigma_j < 1$ ($\gamma(m)$ denotes the probability of conviction if the players announce the profile of messages $m \in M$).

We say that a pair (σ, γ) is *admissible* if the following is true. Consider the game in which the players decide whether to become informed or not and then send a message to the designer who, in turn, makes a (random) decision according to γ. Then this game admits an equilibrium in which each player j acquires the signal with probability σ_j and reports it truthfully. We let $V(\sigma, \gamma; n)$ denote the designer's expected payoff associated with this equilibrium.

We say that an admissible pair (σ, γ) is *ex-post efficient* if for every $m \in M$, $\gamma(m)$ assigns probability one to the alternative that is optimal given the number of signals i and g contained in m. If this condition is violated we also say that the admissible pair (σ, γ) is ex-post inefficient. Finally, we say that the strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is *weakly symmetric* if $\sigma_1 = \ldots = \sigma_k = \rho < 1$ and $\sigma_{k+1} = \ldots = \sigma_n = 1$ where $k = 0, \ldots, n$. In a weakly symmetric strategy profile there are at most two probabilities with which the players acquire information.

In general, there exist ex-post efficient admissible pairs (σ, γ) such that σ is not weakly symmetric. However, this can occur only in non-generic cases. In fact, it is possible to show that for generic values of $P(I)$, p, q, and c, if (σ, γ) is admissible and ex-post efficient then σ is weakly symmetric. Moreover, each player j with $\sigma_j = 1$ has a strict incentive to acquire information.[17] In what follows we restrict attention to generic cases (this also means that we continue to assume that the environment is regular).

We are now ready to state our final result. We assume that $N > 2$ and that $n \leqslant N$.

Proposition 5. *Fix $n \geqslant 2$ and an ex-post efficient admissible pair (σ, γ) with σ weakly symmetric.*[18] *Suppose that $V(\sigma, \gamma; n) > \hat{V}\left(n^B\right)$. Then there exists an ex-post inefficient admissible pair (σ', γ') (with σ' weakly symmetric) such that $V(\sigma', \gamma'; n) > V(\sigma, \gamma; n)$.*

Consider an ex-post efficient admissible pair (σ, γ) in which $\bar{n}$ players acquire the signal with probability $\hat{\rho} \in (0,1)$ and $n - \bar{n}$ acquire with probability one. It is immediate to see that $V(\sigma, \gamma; n) > \hat{V}\left(n^B\right)$ implies $\bar{n} \geqslant 2$. The proof of Proposition 5 considers ex-post inefficient devices γ' of the following form. Fix $\varepsilon > 0$ and $k \in \{0, \bar{n}\}$. If $n - \bar{n} + k'$ (where $k' = 0, \ldots, \bar{n}$ and $k' \neq k$) players acquire information then the designer makes the ex-post efficient decision. If precisely $n - \bar{n} + k$ players acquire information, then the designer makes the ex-post efficient decision with probability $1 - \varepsilon$ and chooses the uninformed (i.e., relying only on the prior) *inferior* alternative with probability ε.

It turns out that there is always a value of $k \in \{0, \bar{n}\}$ such that for sufficiently small ε, two things occur: (i) there exists a weakly symmetric equilibrium in which the first $\bar{n}$ players acquire the signal with probability $\rho(\varepsilon) > \hat{\rho}$ and the last $n - \bar{n}$ players acquire it with probability one; and (ii) the designer gets overall higher utility under this equilibrium than under the original pair (σ, γ).[19]

To summarize, the main message of this section is that for any fixed number of experts, for generic parameter constellations, the designer cannot be optimizing by choosing an ex-post efficient mechanism and having the agents use a mixed equilibrium.

[17] For brevity, we omit the proofs of these technical results.

[18] As mentioned above, we also assume that each player j with $\sigma_j = 1$ has a strict incentive to acquire information and that the environment is regular.

[19] Note that there are two forces at play: on the one hand, the distorted mechanism makes the players more likely to acquire the signal and, therefore, allows a higher expected accuracy of aggregate information; on the other hand, the mechanism introduces "mistakes." As it turns out, for the appropriate choice of k and ε the former prevails.

6. Conclusions

The current paper analyzed a mechanism design problem pertaining to collective choice with information acquisition.

Our analysis yielded three key insights. First, the optimal incentive scheme in such an environment balances a trade-off between inducing players to acquire information and extracting the maximal amount of information from them. In particular, the optimal device ex-ante may be inefficient ex-post (i.e., aggregate information suboptimally from a statistical point of view). Second, the optimal distortions to the ex-post efficient rule depend crucially on the accuracy of the (costly) signals, and can be described analytically for extreme accuracy levels. Third, the comparative statics of the optimal mechanism exhibit some regularities and irregularities, e.g., the expected social value is monotonic in the cost of information and accuracy of private information, but the optimal committee size is not monotonic in the signals' accuracy.

There are many directions this framework suggests pursuing. For example, adding heterogeneity amongst agents, in the form of differential preferences, may affect the optimal design. Indeed, in our model, both the designer and all of the players share the same utility parameter q. However, in many situations it is conceivable that agents have heterogenous preferences (e.g., jurors with different conviction thresholds, political advisors of differing political agendas, department members in different fields, etc.). One could then study the extended mechanism design problem in which, at stage 1, the designer chooses the distribution of preference parameters of the committee members, in addition to choosing the committee's size and the aggregation rule. An analysis of such a scenario would entail defining carefully the goal of the designer (maximizing her own preferences, as characterized by one given q, or implementing a point in the Pareto frontier of the equilibria set). It would be especially interesting to gain insights into the optimal composition of the committee. In particular, would the designer choose a committee comprised of agents with preferences coinciding with her own or would she choose agents with diverging tastes?

Another interesting complication arises when considering environments with more than two alternatives and a richer set of signals. When signals are binary, an agent who acquires information automatically has incentives to reveal her information truthfully. In contrast, this is not necessarily the case when signals have more than two realizations. The analysis of such setups is qualitatively different and is left for future research.

Appendix

Proof of Proposition 2 For every c, let $n^B(c)$ denote the largest integer for which the Bayesian device is admissible. We show that if $n^B(c)$ is sufficiently large then $V\left(n^B(c)+1\right) > \hat{V}\left(n^B(c)\right)$. This will complete the proof of Proposition 2 since $n^B(c)$ is decreasing in c.

We now fix c and write n for $n^B(c)$. We assume $0 < k_n - z(n) < \frac{1}{2}$ (the proof for the case $\frac{1}{2} < k_n - z(n) < 1$ is similar and is therefore omitted).[20] The Bayesian device γ_n^B is admissible, and so the following two (Bayesian) constraints hold:

$$\binom{n-1}{k_n-1} f(k_n; n) \geqslant c, \tag{4}$$

$$-\binom{n-1}{k_n-1} f(k_n-1; n) \geqslant c. \tag{5}$$

Consider now a committee of size $n+1$. For every α in the unit interval, let the device $\gamma_\alpha : \{0, ..., n+1\}$ be defined by:

$$\gamma_\alpha(k) = \begin{cases} 0 & \text{if } k = 0, ..., k_n - 1 \\ \alpha & \text{if } k = k_n \\ 1 & \text{if } k = k_n + 1, ..., n+1 \end{cases}.$$

The constraints that the device γ_α has to satisfy to be admissible can be expressed as:

$$F(\alpha) = \binom{n}{k_n} f(k_n+1; n+1) + \alpha\left[\binom{n}{k_n-1} f(k_n; n+1) - \binom{n}{k_n} f(k_n+1; n+1)\right] \geqslant c, \tag{6}$$

$$L(\alpha) = -\binom{n}{k_n} f(k_n; n+1) + \alpha\left[\binom{n}{k_n} f(k_n; n+1) - \binom{n}{k_n-1} f(k_n-1; n+1)\right] \geqslant c. \tag{7}$$

[20] Since the environment is regular, $k_n - z(n) \neq \frac{1}{2}, 1$.

The function F is decreasing in α. We now assume that n is sufficiently large, so that $k_n - 1 \geqslant n(1-p)$ and $k_n \leqslant np$. This implies:

$$F(0) = \binom{n}{k_n} f(k_n+1; n+1) > \binom{n-1}{k_n-1} f(k_n; n) \geqslant c,$$

and that L is an increasing function that satisfies:

$$L(1) = -\binom{n}{k_n-1} f(k_n-1; n+1) > -\binom{n-1}{k_n-1} f(k_n-1; n) \geqslant c.$$

We let $\hat{\alpha}_1$ denote the greatest value of α for which the device γ_α satisfies constraint (6). Similarly, we let $\hat{\alpha}_2$ denote the smallest value of α for which the device γ_α satisfies constraint (7). Notice that $-f(k_n-1; n) \geqslant f(k_n; n)$ since we are assuming that $k_n - z(n)$ is in $\left(0, \frac{1}{2}\right)$. Thus, the cost c can be at most $\binom{n-1}{k_n-1} f(k_n; n)$ since the Bayesian device γ_n^B is admissible. It follows that:

$$\hat{\alpha}_1 \geqslant \frac{\binom{n}{k_n} f(k_n+1; n+1) - \binom{n-1}{k_n-1} f(k_n; n)}{\binom{n}{k_n} f(k_n+1; n+1) - \binom{n}{k_n-1} f(k_n; n+1)} \equiv \alpha_1,$$

$$\hat{\alpha}_2 \leqslant \frac{\binom{n-1}{k_n-1} f(k_n; n) + \binom{n}{k_n} f(k_n; n+1)}{\binom{n}{k_n} f(k_n; n+1) - \binom{n}{k_n-1} f(k_n-1; n+1)} \equiv \alpha_2.$$

With a slight abuse of notation we let $V(\alpha)$ denote the expected utility of the device γ_α :

$$V(\alpha) = -(1-q) P(G) + \alpha \binom{n+1}{k_n} f(k_n; n+1) + \sum_{k=k_n+1}^{n+1} \binom{n+1}{k} f(k; n+1).$$

The difference between $V(\alpha)$ and $\hat{V}(n)$ is equal to:

$$V(\alpha) - \hat{V}(n) = \binom{n}{k_n} f(k_n; n+1) \left(\frac{n+1}{n-k_n+1} \alpha - 1 \right).$$

Let $\alpha^* = \frac{n-k_n+1}{n+1}$. Then $V(\alpha)$ is greater than $\hat{V}(n)$ if and only if $\alpha < \alpha^*$.

It remains to be shown that $\alpha_2 < \alpha^*$ and $\alpha_2 < \alpha_1$ for sufficiently large values of n. Let us start with the first inequality. We need to show:

$$(n-k_n+1)\left[\binom{n}{k_n}f(k_n;n+1)-\binom{n}{k_n-1}f(k_n-1;n+1)\right]>$$
$$(n+1)\left[\binom{n-1}{k_n-1}f(k_n;n)+\binom{n}{k_n}f(k_n;n+1)\right],$$

which can be rewritten as:

$$-\binom{n-1}{k_n-1}\left(nf(k_n;n+1)+nf(k_n-1;n+1)+(n+1)f(k_n;n)\right)>0.$$

We divide the above the expression by $\binom{n-1}{k_n-1}$, and notice that

$$f(k_n;n+1)+f(k_n-1;n+1)=f(k_n-1;n).$$

We obtain:

$$-nf(k_n-1;n)-(n+1)f(k_n;n)>0.$$

After dividing both sides by $qP(I)(1-p)^{z(n)}p^{n-z(n)}$ and rearranging terms we have:

$$\left(\frac{p}{1-p}\right)^{1-2\lambda}>\frac{n+p}{n+1-p},$$

where $\lambda=k_n-z(n)$. The left hand side is greater than 1 since λ belongs to $\left(0,\frac{1}{2}\right)$, while the right hand side is decreasing in n, and converges to 1 as n goes to infinity.

We now compare α_1 and α_2. We divide both the numerator and the denominator of α_1 by $\binom{n-1}{k_n-1}qP(I)(1-p)^{z(n)}p^{n-z(n)}$, rearrange terms, and obtain:

$$\alpha_1=\frac{\left((1-p)^{\lambda}p^{-\lambda}-p^{\lambda}(1-p)^{-\lambda}\right)+\left(\frac{n}{k_n}\right)\left(-(1-p)^{1+\lambda}p^{-\lambda}+p^{1+\lambda}(1-p)^{-\lambda}\right)}{\left(\frac{n}{k_n}\right)\left(-(1-p)^{1+\lambda}p^{-\lambda}+p^{1+\lambda}(1-p)^{-\lambda}\right)+\left(\frac{n}{n-k_n+1}\right)\left((1-p)^{\lambda}p^{1-\lambda}-p^{\lambda}(1-p)^{1-\lambda}\right)}.$$

We now take the limit of α_1 as n goes to infinity. Both $\left(\frac{n}{k_n}\right)$ and $\left(\frac{n}{n-k_n+1}\right)$ converge to

2 as n grows large. Thus, we have:

$$\bar{\alpha}_1 = \lim_{n\to\infty} \alpha_1 = \frac{\frac{1}{2}\left((1-p)^{\lambda} p^{-\lambda} - p^{\lambda}(1-p)^{-\lambda}\right) - (1-p)^{1+\lambda} p^{-\lambda} + p^{1+\lambda}(1-p)^{-\lambda}}{-(1-p)^{1+\lambda} p^{-\lambda} + p^{1+\lambda}(1-p)^{-\lambda} + (1-p)^{\lambda} p^{1-\lambda} - p^{\lambda}(1-p)^{1-\lambda}}.$$

In a similar way we derive:

$$\bar{\alpha}_2 = \lim_{n\to\infty} \alpha_2 = \frac{\frac{1}{2}\left(-(1-p)^{\lambda} p^{-\lambda} + p^{\lambda}(1-p)^{-\lambda}\right) - (1-p)^{\lambda} p^{1-\lambda} + p^{\lambda}(1-p)^{1-\lambda}}{-(1-p)^{\lambda} p^{1-\lambda} + p^{\lambda}(1-p)^{1-\lambda} + (1-p)^{\lambda-1} p^{2-\lambda} - p^{\lambda-1}(1-p)^{2-\lambda}}.$$

It is tedious but simple to show that $\bar{\alpha}_2 < \bar{\alpha}_1$ for all p in $\left(\frac{1}{2},1\right)$ and all λ in $\left(0,\frac{1}{2}\right)$. ■

Proof of Proposition 3 As mentioned in the body of the paper, we focus on a special case of Proposition 3. Namely, we restrict attention to the following:

Proposition 3*. *Fix $n \geqslant 2$, $P(I)$ and q. There exists $\bar{p} \in \left(\frac{1}{2},1\right)$ such that for each $p > \bar{p}$ the following holds. If the environment is regular and $k_n = k_{n-1}$ then there exists $\varepsilon > 0$ such that for $c \in (\hat{c}, \hat{c}+\varepsilon)$ the optimal device with n agents takes the form:*

$$\bar{\gamma}_n(k) = \begin{cases} \gamma_n^B(k) = 0 & \text{for } k = 0,\ldots,k_n - 1 \\ \alpha & \text{for } k = k_n \\ \gamma_n^B(k) = 1 & \text{for } k = k_n + 1,\ldots,n \end{cases}$$

where $\hat{c} = \hat{V}(n) - \hat{V}(n-1) = \binom{n-1}{k_n-1} f(k_n;n)$ and $\alpha \in [0,1)$ is such that the $IC(i)$ constraint is satisfied with equality.[21]

Proof of Proposition 3* For a given number $n \geqslant 2$ of agents who acquire information, the designer's problem can be written as:

$$\max_{\gamma(0),\ldots,\gamma(n)\in[0,1]} -(1-q)P(G) + \sum_{k=0}^{n} v(k)\gamma(k)$$

[21] In a similar way, we can show that when signals are very accurate, $k_n = k_{n-1}+1$ (i.e., $IC(g)$ is the first constraint to bind with the Bayesian device), and the cost c is slightly above $-\binom{n-1}{k_n-1} f(k_n-1,n)$ then it is optimal to distort the Bayesian device at $k_n - 1$.

$$\text{s.t. } \sum_{k=0}^{n} a(k)\gamma(k) \geqslant c$$
$$\sum_{k=0}^{n} b(k)\gamma(k) \geqslant c,$$

where

$$v(k) = \binom{n}{k} f(k;n),$$

$$a(k) = \binom{n-1}{k-1} f(k;n) - \binom{n-1}{k} f(k+1;n)$$

is the coefficient of $\gamma(k)$ in $IC(i)$, and

$$b(k) = \binom{n-1}{k} f(k;n) - \binom{n-1}{k-1} f(k-1;n)$$

is the coefficient of $\gamma(k)$ in $IC(g)$. We use the convention $\binom{n-1}{-1} = \binom{n-1}{n} = 0$.

It is immediate to see that when p is sufficiently large (and the environment is regular) the following holds: $a(0) = -f(1;n) > 0$, $a(n) = f(n;n) > 0$, $b(0) = f(0;n) < 0$ and $b(n) = -f(n-1;n) < 0$.

Notice that for $k = 1, \ldots, n-1$, we can rewrite $a(k)$ and $b(k)$ as

$$a(k) = \binom{n-1}{k-1}\frac{1}{k}\left[(n(1-p)-k)\,qP(I)(1-p)^{k}p^{n-k-1} + (k-np)(1-q)P(G)p^{k}(1-p)^{n-k-1}\right],$$

$$b(k) = \binom{n-1}{k-1}\frac{1}{k}\left[(k-n(1-p))\,qP(I)(1-p)^{k-1}p^{n-k} + (np-k)(1-q)P(G)p^{k-1}(1-p)^{n-k}\right].$$

Clearly, $a(k) < 0$ and $b(k) > 0$ for any $k \in [n(1-p), np]$.

For p sufficiently close to one, $n(1-p) < 1$ and $np > n-1$ and so the signs of $\{a(k)\}_{k=0}^{n}$ and $\{b(k)\}_{k=0}^{n}$ are determined. Notice also that for p sufficiently large k_n is equal to $\frac{n+1}{2}$ when n is odd and equal either to $\frac{n}{2}$ or $\frac{n}{2}+1$ when n is even.

Recall that we consider the case $k_n = k_{n-1}$. Thus, $\hat{c} = \binom{n-1}{k_n-1} f(k_n;n)$ denotes the largest value of c for which the Bayesian device is admissible ($\hat{c}$ is the LHS of $IC(i)$ of the Bayesian device). Suppose now that the cost c is slightly above $\hat{c}$. Clearly, when using the Bayesian device the $IC(i)$ constraint is violated (while the $IC(g)$ is not binding). There are two ways to satisfy $IC(i)$ at c : (i) decrease $\gamma(k)$ for some $k = k_n, k_n+1, \ldots, n-1$ (recall $n-1 < np$); or (ii) increase $\gamma(0)$ (recall $n(1-p) < 1$).[22] Of course, it is optimal to choose k which maximizes the function

$$\left|\frac{a(k)}{v(k)}\right|.$$

In what follows we show that

$$\left|\frac{a(k_n)}{v(k_n)}\right| > \left|\frac{a(k_n+1)}{v(k_n+1)}\right| > \ldots \left|\frac{a(n-1)}{v(n-1)}\right| > \left|\frac{a(0)}{v(0)}\right|.$$

First, we show that

$$\left|\frac{a(k)}{v(k)}\right| > \left|\frac{a(k')}{v(k')}\right| \tag{8}$$

for any k, k' with $k_n \leqslant k < k' \leqslant n-1$.

Consider $k = k_n, \ldots, n-2$. Notice that $a(k)v(k) < 0$ and $a(k')v(k') < 0$.[23] Thus, inequality (8) is equivalent to

$$\frac{a(k)}{v(k)} < \frac{a(k')}{v(k')}.$$

Notice that

$$\frac{a(k)}{v(k)} = \frac{\binom{n-1}{k-1} f(k;n) - \binom{n-1}{k} f(k+1;n)}{\binom{n}{k} f(k;n)} = \frac{k}{n} - \frac{n-k}{n}\frac{f(k+1;n)}{f(k;n)}.$$

Thus, it is enough to show that

$$k - (n-k)\frac{f(k+1;n)}{f(k;n)} < k' - (n-k')\frac{f(k'+1;n)}{f(k';n)},$$

or equivalently (notice that $f(k;n) > 0$, $f(k';n) > 0$)

$$(kf(k;n) - (n-k) f(k+1;n)) f(k';n) < (k' f(k';n) - (n-k') f(k'+1;n)) f(k;n).$$

[22] It is also obvious that it is optimal to satisfy the $IC(i)$ constraint with equality.

[23] Recall that the environment is regular and, thus, $v(k_n) > 0$.

By using the definition of the function f we get

$$(1-p)^{2n-k-k'-1} \begin{array}{l}\Big[-kqP(I)(1-p)^{2k-n+1}p^{n-k} + k(1-q)P(G)p^{k}(1-p) + \\ (n-k)qP(I)(1-p)^{2k-n+2} - (n-k)(1-q)P(G)p^{k+1}\Big]\end{array}$$
$$\left(-qP(I)(1-p)^{2k'-n}p^{n-k'} + (1-q)P(G)p^{k'}\right) <$$

$$< (1-p)^{2n-k-k'-1} \begin{array}{l}\Big[-k'qP(I)(1-p)^{2k'-n+1}p^{n-k'} + k'(1-q)P(G)p^{k'}(1-p) + \\ (n-k')qP(I)(1-p)^{2k'-n+2} - (n-k')(1-q)P(G)p^{k'+1}\Big]\end{array}$$
$$\left(-qP(I)(1-p)^{2k-n}p^{n-k} + (1-q)P(G)p^{k}\right).$$

We divide both sides by $(1-p)^{2n-k-k'-1}$ and take the limit as p becomes close to one.

The LHS converges to $(k-n)((1-q)P(G))^2$. The RHS converges to $(k'-n)((1-q)P(G))^2$ if $k > \frac{n}{2}$ and to $(k'-n)((1-q)P(G))^2 + (n-k')q(1-q)P(I)P(G)$ if $k = k_n = \frac{n}{2}$. In any case the strict inequality is satisfied.

Similarly, it is easy to show that for p sufficiently close to one

$$\left|\frac{a(n-1)}{v(n-1)}\right| > \left|\frac{a(0)}{v(0)}\right|.$$

The result then follows. ■

Proof of Proposition 4 Let $y > 1$ be given. Suppose that p is slightly larger than $\frac{y^{1/N}}{1+y^{1/N}}$. Clearly, the Bayesian device is $\gamma_N^B(0) = \ldots = \gamma_N^B(N-1) = 0$ and $\gamma_N^B(N) = 1$. This device is admissible as long as the cost is smaller than or equal to $f(N;N)$ (this is the LHS of $IC(i)$ of the Bayesian device).

It is easy to show that there exists $\hat{k} \in [0, N(1-p))$ such that $a(k)$ is negative if $k = \hat{k}+1, \ldots, N-1$ and positive if $k = 0, \ldots, \hat{k}$ or if $k = N$.

Suppose now that the cost is slightly above $f(N;N)$. Clearly, there is only one way to satisfy the incentive constraint $IC(i)$: increase $\gamma(k)$ for some $k = 0, \ldots, \hat{k}$.

Notice that if $N = 2$ then $\hat{k} = 0$ and the proof is complete. We now assume $N > 2$ and show that

$$\left|\frac{a(0)}{v(0)}\right| > \ldots > \left|\frac{a(\hat{k})}{v(\hat{k})}\right|.$$

This implies that the optimal distortion is at $k = 0$.

Notice that $a(k)\,v(k) < 0$ and, thus, it suffices to show that the function $\frac{a(k)}{v(k)}$ is increasing in the interval $[0, N(1-p))$. We now illustrate that $\frac{a(k)}{v(k)}$ is increasing in k when $p = \frac{y^{1/N}}{1+y^{1/N}}$. By continuity, the result holds if p is sufficiently close to $\frac{y^{1/N}}{1+y^{1/N}}$. Let us fix $p = \frac{y^{1/N}}{1+y^{1/N}}$. Using previous algebraic manipulations, we have

$$\frac{a(k)}{v(k)} = \frac{k}{N} - \frac{N-k}{N}\frac{-qP(I)(1-p)^{k+1}p^{N-k-1} + (1-q)P(G)p^{k+1}(1-p)^{N-k-1}}{-qP(I)(1-p)^{k}p^{N-k} + (1-q)P(G)p^{k}(1-p)^{N-k}} =$$

$$\frac{k}{N} + \frac{N-k}{N}\frac{qP(I)\left(\frac{1-p}{p}\right)^{k+1}p^{N} - (1-q)P(G)\left(\frac{p}{1-p}\right)^{k+1}(1-p)^{N}}{-qP(I)\left(\frac{1-p}{p}\right)^{k}p^{N} + (1-q)P(G)\left(\frac{p}{1-p}\right)^{k}(1-p)^{N}}.$$

We use the definition of p and divide both the numerator and the denominator of the RHS of the above equality by $(1-q)P(G)\left(1+y^{1/N}\right)^{N}$. We get:

$$\frac{a(k)}{v(k)} = \frac{k}{N} + \frac{N-k}{N}\frac{y^{\frac{2N-k-1}{N}} - y^{\frac{k+1}{N}}}{-y^{\frac{2N-k}{N}} + y^{\frac{k}{N}}}.$$

Remember that $y > 1$ and $k < N(1-p) < N-1$. Thus,

$$\frac{y^{\frac{2N-k-1}{N}} - y^{\frac{k+1}{N}}}{-y^{\frac{2N-k}{N}} + y^{\frac{k}{N}}} < 0.$$

We compute the derivative of $\frac{a(k)}{v(k)}$ with respect to k and obtain

$$\frac{1}{N} - \frac{1}{N}\frac{y^{\frac{2N-k-1}{N}} - y^{\frac{k+1}{N}}}{-y^{\frac{2N-k}{N}} + y^{\frac{k}{N}}} + \left(\frac{N-k}{N}\right)\left(\frac{2\ln y}{N}\right)\frac{y^{\frac{2N+1}{N}} - y^{\frac{2N-1}{N}}}{\left(-y^{\frac{2N-k}{N}} + y^{\frac{k}{N}}\right)^{2}} > 0,$$

and the result follows. ■

Proof of Proposition 5 Consider a cost level c, a committee of size $n \geqslant 2$ and an ex-post efficient admissible pair (σ, γ) with σ weakly symmetric and such that each player

j with $\sigma_j = 1$ has a strict incentive to acquire information. Let $\bar{n} \leqslant n$ be such that $\sigma_1 = \ldots = \sigma_{\bar{n}} = \hat{\rho} < 1$ and $\sigma_{\bar{n}+1} = \ldots = \sigma_n = 1$. It follows from $V(\sigma, \gamma; n) > \hat{V}\left(n^B\right)$ that $\bar{n} \geqslant 2$. The fact that the designer always makes the ex-post efficient decision and the fact that each player has at least one opponent who randomizes imply that each player $j = 1, \ldots, n$ has a strict incentive to reveal truthfully her type after making the information acquisition decision (here we use the assumption that the environment is regular).

We now turn to the incentives at the information acquisition stage of a player who randomizes, say player 1. For each $k = 0, \ldots, \bar{n}$, let $w(k) = \hat{V}(n - \bar{n} + k)$ denote the expected utility of the Bayesian device with $n - \bar{n} + k$ players. Suppose that players $2, \ldots, \bar{n}$ acquire the signal with probability ρ, players $\bar{n}+1, \ldots, n$ acquire the signal with probability one and that each player $j = 2, \ldots, n$ reveals her type truthfully. Suppose also that the designer uses the device γ (i.e., she makes the ex-post optimal decision). Let $W^1(\rho, \gamma)$ denote the expected utility (from the decision) for player 1 when she acquires the signal and is sincere. Similarly, let $W^0(\rho, \gamma)$ denote the utility of player 1 when she does not acquire the signal and is sincere. We have:

$$W^1(\rho, \gamma) = \sum_{k=0}^{\bar{n}-1} \tbinom{n-1}{k} \rho^k (1-\rho)^{\bar{n}-1-k} w(k+1),$$
$$W^0(\rho, \gamma) = \sum_{k=0}^{\bar{n}-1} \tbinom{n-1}{k} \rho^k (1-\rho)^{\bar{n}-1-k} w(k).$$

We also let $H(\rho)$ denote the difference between the utility of becoming informed and the utility of remaining uninformed:

$$H(\rho) = W^1(\rho, \gamma) - W^0(\rho, \gamma) = \sum_{k=0}^{\bar{n}-1} \binom{n-1}{k} \rho^k (1-\rho)^{\bar{n}-1-k} \left[w(k+1) - w(k)\right].$$

Of course, the function $H(\cdot)$ is continuous and $H(\hat{\rho}) = c$ (so that $W^0(\hat{\rho}, \gamma) = W^1(\hat{\rho}, \gamma) - c$). Finally, notice that the designer's utility $V(\sigma, \gamma; n)$ associated with the ex-post efficient admissible pair (σ, γ) can be expressed as

$$V(\sigma, \gamma; n) = \hat{\rho} W^1(\hat{\rho}, \gamma) + (1 - \hat{\rho}) W^0(\hat{\rho}, \gamma) = W^1(\hat{\rho}, \gamma) - (1 - \hat{\rho}) c.$$

We now construct an ex-post inefficient admissible pair that guarantees the designer an expected payoff higher than $V(\sigma, \gamma; n)$. We need to distinguish between two cases depending on whether $H(\rho)$ approaches $H(\hat{\rho})$ from below or from above as ρ approaches $\hat{\rho}$ from above.

Case 1: $\lim_{\rho \searrow \hat{\rho}} \frac{H(\rho) - H(\hat{\rho})}{|H(\rho) - H(\hat{\rho})|} = -1$

Consider the following class of devices. The set of messages available to players $1, \ldots, \bar{n}$ is $\{\emptyset, i, g\}$ while the set of messages available to players $\bar{n}+1, \ldots, n$ is $\{i, g\}$. The designer uses the following decision rule γ^ε. If at least one of the first $\bar{n}$ players acquire the signal (and report i or g) then the designer makes the ex-post efficient decision. If nobody of the first $\bar{n}$ acquires the signal then the designer makes the ex-post efficient decision with probability $1 - \varepsilon$ and chooses the uninformed (i.e., relying only on the prior) inferior alternative with probability $\varepsilon > 0$.

Consider player 1 and suppose that players $j = 2, \ldots, \bar{n}$ acquire the signal with probability ρ, players $j = \bar{n}+1, \ldots, n$ acquire the signal with probability one, and that each player $j = 2, \ldots, n$ reveals her type truthfully. Let $W^1(\rho, \gamma^\varepsilon)$ and $W^0(\rho, \gamma^\varepsilon)$ denote the expected payoff (from the decision) of player 1 if she acquires and if she does not acquire the signal, respectively (in both cases we assume that player 1 is sincere). Of course, player 1 is indifferent between the two courses of actions if and only if ρ is such that:

$$W^1(\rho, \gamma^\varepsilon) - W^0(\rho, \gamma^\varepsilon) = H(\rho) + \varepsilon (1 - \rho)^{\bar{n}-1} (w(0) - \underline{v}) = c,$$

where $\underline{v} < w(0)$ denotes the designer's expected payoff from making the wrong uninformed decision. Note that for ε sufficiently small, there exist $\rho_1 > \hat{\rho}$ and $\rho_2 < \hat{\rho}$ such that

$$H(\rho_1) + \varepsilon (1 - \rho_1)^{\bar{n}-1} (w(0) - \underline{v}) < c,$$
$$H(\rho_2) + \varepsilon (1 - \rho_2)^{\bar{n}-1} (w(0) - \underline{v}) > c.$$

Therefore, by continuity for $\varepsilon > 0$ there exists $\rho(\varepsilon) \in (\rho_1, \rho_2)$ for which

$$H(\rho(\varepsilon)) + \varepsilon (1-\rho(\varepsilon))^{\bar{n}-1} (w(0) - \underline{v}) = c.$$

Moreover, it follows from the implicit function theorem that $\rho(\varepsilon) > \hat{\rho}$ for $\varepsilon > 0$ (of course, $\lim_{\varepsilon \searrow 0} \rho(\varepsilon) = \hat{\rho}$). Again, from the above equality, $W^0(\rho(\varepsilon), \gamma^\varepsilon) = W^1(\rho(\varepsilon), \gamma^\varepsilon) - c$.

Consider the pair $(\sigma^\varepsilon, \gamma^\varepsilon)$ where $\sigma_1^\varepsilon = \ldots = \sigma_{\bar{n}}^\varepsilon = \rho(\varepsilon)$ and $\sigma_{\bar{n}+1}^\varepsilon = \ldots = \sigma_n^\varepsilon = 1$. Recall that with the original pair (σ, γ) only the information acquisition constraints of the first $\bar{n}$ players are satisfied with equality. All the other constraints hold with strict inequality. It follows that for ε sufficiently small the ex-post inefficient pair $(\sigma^\varepsilon, \gamma^\varepsilon)$ is admissible.

Notice that

$$W^1(\rho(\varepsilon), \gamma^\varepsilon) = \sum_{k=0}^{\bar{n}-1} \binom{n-1}{k} \rho(\varepsilon)^k (1-\rho(\varepsilon))^{\bar{n}-1-k} w(k+1) > W^1(\hat{\rho}, \gamma),$$

where the inequality follows from $\rho(\varepsilon) > \hat{\rho}$.

The designer's expected utility associated with $(\sigma^\varepsilon, \gamma^\varepsilon)$ is equal to

$$\begin{aligned} V(\sigma^\varepsilon, \gamma^\varepsilon; n) &= \rho(\varepsilon) W^1(\rho(\varepsilon), \gamma^\varepsilon) + (1-\rho(\varepsilon)) W^0(\rho(\varepsilon), \gamma^\varepsilon) = \\ &= W^1(\rho(\varepsilon), \gamma^\varepsilon) - (1-\rho(\varepsilon)) c. \end{aligned}$$

Hence, $\rho(\varepsilon) > \hat{\rho}$ and $W^1(\rho(\varepsilon), \gamma^\varepsilon) > W^1(\hat{\rho}, \gamma)$ imply $V(\sigma^\varepsilon, \gamma^\varepsilon; n) > V(\sigma, \gamma; n)$.

Case 2: $\lim\limits_{\rho \searrow \hat{\rho}} \frac{H(\rho) - H(\hat{\rho})}{|H(\rho) - H(\hat{\rho})|} = 1$

This case is analogous to the previous one. We consider a class of devices in which the first $\bar{n}$ can announce a message in $\{\emptyset, i, g\}$ and the last $n - \bar{n}$ have to report a message in $\{i, g\}$. The designer uses a device γ^ε which makes the ex-post efficient decision when at least one of the first $\bar{n}$ players does not acquire the signal. If the first $\bar{n}$ players all acquire the signal then the designer makes the right decision with probability $1 - \varepsilon$ and chooses the uninformed inferior alternative with probability ε.

Proceeding as in Case 1 it is possible to show that for ε sufficiently small there exists an ex-post inefficient admissible pair $(\sigma^\varepsilon, \gamma^\varepsilon)$ with $\sigma_1^\varepsilon = \ldots = \sigma_{\bar{n}}^\varepsilon = \rho(\varepsilon) > \hat{\rho}$ and $\sigma_{\bar{n}+1}^\varepsilon = \ldots =$

$\sigma_n^{\varepsilon} = 1$.

Notice that $\rho(\varepsilon) > \hat{\rho}$ implies that

$$W^0(\rho(\varepsilon), \gamma^{\varepsilon}) = \sum_{k=0}^{\bar{n}-1} \binom{n-1}{k} \rho(\varepsilon)^k (1-\rho(\varepsilon))^{\bar{n}-1-k} w(k) > W^0(\hat{\rho}, \gamma).$$

Finally, it follows from the two inequalities above that $V(\sigma^{\varepsilon}, \gamma^{\varepsilon}; n) > V(\sigma, \gamma; n)$. ■

References

[1] Al-Najjar, N. and Smorodinsky, R. [2000], "Pivotal Players and the Characterization of Influence," *Journal of Economic Theory*, Volume 92, pages 318-342.

[2] Auriol, E. and Gary-Bobo, R. J. [1999], "On The Optimal Number of Representatives," CMS-EMS Discussion Paper 1286.

[3] Austen-Smith, D. and Banks, J. S. [1996], "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review*, Volume 90(1), pages 34-45.

[4] Bergemann, D. and Välimäki, J. [2002], "Information Acquisition and Efficient Mechanism Design," *Econometrica*, Volume 70(3), pages 1007-1033.

[5] Cai, H. [2003], "Costly Participation and Heterogeneous Preferences in Informational Committees," mimeo.

[6] Feddersen, T. J. and Pesendorfer, W. [1996], "The Swing Voter's Curse," *American Economic Review*, Volume 86(3), pages 408-424.

[7] Feddersen, T. J. and Pesendorfer, W. [1997], "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, Volume 65(5), pages 1029-58.

[8] Feddersen, T. J. and Pesendorfer, W. [1998], "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review*, Volume 92(1), pages 23-35.

[9] Gerardi, D. and Yariv, L. [2003], "Committee Design in the Presence of Communication," Cowles Foundation Discussion Paper Number 1411.

[10] Gerardi, D. and Yariv, L. [2007a], "Deliberative Voting," *Journal of Economic Theory*, Volume 134(1), pages 317-338.

[11] Gerardi, D. and Yariv, L. [2007b], "Information Acquisition in Committees: Technical Addendum," mimeo.

[12] Gersbach, H. [1995], "Information Efficiency and Majority Decisions," *Social Choice and Welfare*, Volume 12, pages 363-370.

[13] Gershkov, A. and Szentes, B. [2004], "Optimal Voting Scheme with Costly Information Acquisition," mimeo.

[14] Li, H. [2001], "A Theory of Conservatism," *Journal of Political Economy*, Volume 109(3), pages 617-636.

[15] Myerson, R. B. [1991], *Game Theory: Analysis of Conflict.* Cambridge, MA: Harvard University Press.

[16] Persico, N. [2004], "Committee Design with Endogenous Information," *The Review of Economic Studies*, Volume 71(1), pages 165-94.

[17] Smorodinsky, R. and Tennenholtz, M. [2006], "Overcoming Free Riding in Multi-Party Computations - The Anonymous Case," *Games and Economic Behavior*, Volume 55(2), pages 385-406.

Information Acquisition in Committees: Technical Addendum

Dino Gerardi*and Leeat Yariv†

May 2007

1 Proof of Proposition 3

Consider a committee of size n. We look for the optimal mechanism under the restriction that all n players acquire information.

The problem is:

$$\max_{\gamma(0),\ldots,\gamma(n)\in[0,1]} -(1-q)\,P(G)+\sum_{k=0}^{n} v(k)\,\gamma(k)$$
$$\text{s.t. } \sum_{k=0}^{n} a(k)\,\gamma(k) \geqslant c$$
$$\sum_{k=0}^{n} b(k)\,\gamma(k) \geqslant c,$$

where

$$v(k)=\binom{n}{k} f(k,n),$$

$$a(k)=\binom{n-1}{k-1} f(k,n)-\binom{n-1}{k} f(k+1,n)$$

is the coefficient of $\gamma(k)$ in $IC(i)$, and

$$b(k)=\binom{n-1}{k} f(k,n)-\binom{n-1}{k-1} f(k-1,n)$$

is the coefficient of $\gamma(k)$ in $IC(g)$. We use the convention $\binom{n-1}{-1}=\binom{n-1}{n}=0$.

*Department of Economics, Yale University, 30 Hillhouse Avenue, New Haven, CT 06511. e-mail: donato.gerardi@yale.edu

†Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125. e-mail: lyariv@hss.caltech.edu

Our optimization problem falls under the class of problems known as parametric linear programs. In particular, notice that the solution is continuous in the cost c (see, for instance, Zhang and Liu [1990]).

The goal is to show that when p is sufficiently close to one the optimal mechanism takes the form:

$$\bar{\gamma}_n(0) = \ldots = \bar{\gamma}_n\left(\hat{k}-1\right) = 0, \quad \bar{\gamma}_n\left(\hat{k}\right) = \alpha, \quad \bar{\gamma}_n\left(\hat{k}+1\right) = \ldots = \bar{\gamma}_n(k_n - 1) = 1,$$
$$\bar{\gamma}_n(k_n) = \ldots = \bar{\gamma}_n\left(\bar{k}-1\right) = 0, \quad \bar{\gamma}_n\left(\bar{k}\right) = \beta, \quad \bar{\gamma}_n\left(\bar{k}+1\right) = \ldots = \bar{\gamma}_n(n) = 1,$$

where $\alpha, \beta \in [0,1]$, and $0 < \hat{k} < k_n \leqslant \bar{k} < n$.

We assume that p is sufficiently large. Of course, $a(0) = -f(1,n) > 0$, $a(n) = f(n,n) > 0$, $b(0) = f(0,n) < 0$ and $b(n) = -f(n-1,n) < 0$.

Notice that for $k = 1, \ldots, n-1$, we can rewrite $a(k)$ and $b(k)$ as

$$a(k) = \binom{n-1}{k-1}\frac{1}{k}\left[(n(1-p)-k)\,qP(I)(1-p)^k p^{n-k-1} + (k-np)(1-q)P(G)p^k(1-p)^{n-k-1}\right],$$

$$b(k) = \binom{n-1}{k-1}\frac{1}{k}\left[(k-n(1-p))\,qP(I)(1-p)^{k-1} p^{n-k} + (np-k)(1-q)P(G)p^{k-1}(1-p)^{n-k}\right].$$

Clearly, $a(k) < 0$ and $b(k) > 0$ for any $k \in [n(1-p), np]$.

Since p is close to one, $n(1-p) < 1$ and $np > n-1$ and, therefore, $a(k) < 0$ and $b(k) > 0$ for every $k = 1, \ldots, n-1$.

Throughout, we assume that n is odd and that $k_n = k_{n-1}$ (so that $IC(i)$ is the first constraint to bind when the device is Bayesian). In this case, k_n is equal to $\frac{n+1}{2}$.[1]

We know that when the cost is $\hat{c} = \binom{n-1}{k_n-1} f(k_n, n)$, the Bayesian device satisfies the $IC(i)$ constraint with equality. For costs above $\hat{c}$ we need to introduce distortions in order to induce all n players to acquire information. We also know from Proposition 3* in the Appendix of the paper that for c sufficiently close to $\hat{c}$ it is optimal to distort the mechanism at $k_n = \frac{n+1}{2}$ and set $\gamma(k_n)$ smaller than one. As c increases, $\gamma(k_n)$ decreases. Notice, however, that there exists a critical value of the cost $\bar{c} > \hat{c}$ such that at $\bar{c}$ the optimal mechanism is $\gamma(0) = \ldots = \gamma\left(\frac{n-1}{2}\right) = 0$, $\gamma\left(\frac{n+1}{2}\right) \in (0,1)$, $\gamma\left(\frac{n+3}{2}\right) = \ldots = \gamma(n) = 1$, and the value of $\gamma\left(\frac{n+1}{2}\right)$ is such that both constraints are satisfied with equality. To see this, note that if $\gamma(0) = \ldots = \gamma\left(\frac{n+1}{2}\right) = 0$ and $\gamma\left(\frac{n+3}{2}\right) = \ldots = \gamma(n) = 1$, then the LHS of the $IC(g)$ constraint is equal to $-\binom{n-1}{k_n} f(k_n; n) < 0$.

We now show that as the cost increases above $\bar{c}$ it is optimal to continue decreasing the value of $\gamma\left(\frac{n+1}{2}\right)$ and to start increasing the value of $\gamma\left(\frac{n-1}{2}\right)$. More generally, we prove the following.

[1] The cases in which n is even and/or $k_n = k_{n-1} + 1$ can be analyzed in a similar way.

Claim 1 *Assume that we are at a point $c > \bar{c}$ where the optimal mechanism is*

$$\begin{aligned}
&\bar{\gamma}_n(0) = \ldots = \bar{\gamma}_n\left(\hat{k}\right) = 0, && \bar{\gamma}_n\left(\hat{k}+1\right) = \ldots = \bar{\gamma}_n(k_n - 1) = 1, \\
&\bar{\gamma}_n(k_n) = \ldots = \bar{\gamma}_n\left(\bar{k}-1\right) = 0, \quad \bar{\gamma}_n\left(\bar{k}\right) = \beta, && \bar{\gamma}_n\left(\bar{k}+1\right) = \ldots = \bar{\gamma}_n(n) = 1,
\end{aligned} \tag{1}$$

$\beta \in (0,1)$, and $0 < \hat{k} < k_n \leqslant \bar{k} < n$. Suppose that the cost increases. Then it is optimal to continue decreasing $\bar{\gamma}_n\left(\bar{k}\right)$ and to start increasing $\bar{\gamma}_n\left(\hat{k}\right)$.

In what follows, we provide a proof for Claim 1. A symmetric claim also holds:

Claim 2 *Assume that we are at a cost $c > \bar{c}$ where the optimal mechanism is*

$$\begin{aligned}
&\bar{\gamma}_n(0) = \ldots = \bar{\gamma}_n\left(\hat{k}-1\right) = 0, \quad \bar{\gamma}_n\left(\hat{k}\right) = \alpha, && \bar{\gamma}_n\left(\hat{k}+1\right) = \ldots = \bar{\gamma}_n(k_n - 1) = 1, \\
&\bar{\gamma}_n(k_n) = \ldots = \bar{\gamma}_n\left(\bar{k}-1\right) = 0, && \bar{\gamma}_n\left(\bar{k}\right) = \bar{\gamma}_n\left(\bar{k}+1\right) = \ldots = \bar{\gamma}_n(n) = 1,
\end{aligned}$$

where $\alpha \in (0,1)$, and $0 < \hat{k} < k_n \leqslant \bar{k} < n$. Suppose that the cost increases. Then it is optimal to continue increasing $\bar{\gamma}_n\left(\hat{k}\right)$ and to start decreasing $\bar{\gamma}_n\left(\bar{k}\right)$.

The proof of Claim 2 is identical to that of Claim 1 and is thus omitted. The combination of these two claims (together with Remark 3 below) provide the proof of Proposition 3.[2]

Proof of Claim 1

Note that the optimal device is the solution to a linear programming problem with two constraints, $IC(i)$ and $IC(g)$, and the additional constraints that every $\gamma(k)$ belongs to $[0,1]$. It follows that there will be at most two values of k at which $\gamma(k)$ is different from 0 or 1 (see, e.g., Luenberger [1965], Chapter 3). Clearly, the optimal mechanism is continuous in c. Thus, if we start from the device (1) and increase c by a small amount, the optimal mechanism is such that the value of $\bar{\gamma}_n\left(\bar{k}\right)$ is close to β. Therefore, if we start from (1) and increase c, one change must pertain to $\bar{\gamma}_n\left(\bar{k}\right)$.

In principle, there are different ways to satisfy the constraints when c increases:

1. Decrease the value of $\gamma\left(\bar{k}\right)$ and increase the value of $\gamma(k)$ for some $k = 1, \ldots, \hat{k}$;
2. Decrease the value of $\gamma\left(\bar{k}\right)$ and increase the value of $\gamma(k)$ for some $k = k_n\left(= \frac{n+1}{2}\right), \ldots, \bar{k}-1$;
3. Increase the value of $\gamma\left(\bar{k}\right)$ and decrease the value of $\gamma(k)$ for some $k = \bar{k} + 1, \ldots, n-1$;

[2]Note that in generic environments the optimal distortionary device entails randomization for at least one profile of reports. Our proof does, however, extend to non-generic cases in which for some cost levels, the optimal distortionary device entails no randomization.

4. Increase the value of $\gamma\left(\bar{k}\right)$ and decrease the value of $\gamma\left(k\right)$ for some $k=\hat{k}+1,\ldots,k_n-1\left(=\frac{n-1}{2}\right)$;

5. Increase the value of $\gamma\left(\bar{k}\right)$ and increase the value of $\gamma\left(0\right)$;

6. Decrease the value of $\gamma\left(\bar{k}\right)$ and decrease the value of $\gamma\left(n\right)$.

In all cases, the optimal thing to do is to satisfy both constraints with equality. Recall that we start at a point where both constraints are binding and the mechanism is not Bayesian. If we end up with a mechanism under which one constraint is not binding, the mechanism cannot be optimal.[3]

Below we prove the following facts:

A In case 1, the optimal distortion is to use $\hat{k}$, the largest k available.

B Any change in which we increase $\gamma(k')$ and decrease $\gamma(k'')$, where $k'=k_n,...,n-2$ and $k''=k'+1,...,n-1$ has a negative effect on the designer's expected utility (the objective function). Furthermore, this change is worse than any change in which we decrease $\gamma(k')$ and increase $\gamma(k)$, where $k=1,...,\hat{k}$.

C Case 4 is not feasible.

D Case 5 is not feasible.

E Case 6 is not feasible.

Note that the distortions mentioned in Fact A certainly generate a decrease in the expected value of the designer's objective function. Fact B implies that case 3 cannot be optimal directly. In fact, it implies that distortions of the type specified in case 3 generate lower expected values to the designer than distortions of the type specified in case 1. In particular, the former yield a decrease in the designer's expected value as well. Fact B also implies that case 2 cannot be optimal. Indeed, suppose we end up with a device in which $\gamma\left(\bar{k}\right)\in(0,1)$ and $\gamma\left(k\right)\in(0,1)$ for some $k=k_n,\ldots,\bar{k}-1$. Then consider the following deviation. Decrease the value of $\gamma\left(k\right)$ and increase the value of $\gamma\left(\bar{k}\right)$ so that the LHS of both constraints decreases by the same (small) amount δ. It follows from the first part of Fact B that this change will *increase* the value of the objective function by some amount $\Delta>0$.[4] Now, decrease the value of

[3]The proof of this fact depends on which case -1 through 6- we are considering. In each case, it is straightforward to identify a deviation that does not violate either constraint and improves the utility. For the sake of brevity, we do not include the relevant calculations.

[4]We know from Fact B that if we increase $\gamma\left(k'\right)$ and decrease $\gamma\left(k''\right)$, where $k'=k_n,...,n-2$ and $k''=k'+1,...,n-1$, then the expected utility decreases. Notice that $\bar{k}\leqslant n-1$. Therefore, if we decrease the value of $\gamma\left(k\right)$ for some $k=k_n,\ldots,\bar{k}-1$ and increase the value of $\gamma\left(\bar{k}\right)$ (i.e., we take a "mirror image" of the type of changes described in Fact B), then the expected utility must increase.

$\gamma(k)$ and increase the value of $\gamma\left(\tilde{k}\right)$, for some $\tilde{k} = 1, \ldots, \hat{k}$, so that the LHS of both constraints increases by δ given above. This will *decrease* the value of the objective function by $\Delta' > 0$. The second part of Fact B implies that $\Delta > \Delta'$ and so the the combination of the two changes is feasible and strictly beneficial.

Proof of Fact A

The goal of this section is as follows. Fix $k' = k_n \left(= \frac{n+1}{2}\right), \ldots, n-1$ and $k = 1, \ldots, \frac{n+1}{2} - 2 \left(= \frac{n-3}{2}\right)$. Suppose that we decrease $\gamma(k')$ by $\eta > 0$ and increase the value of $\gamma(k)$ by $\varepsilon > 0$ to increase the LHS of both constraints by the same (small) number $\delta > 0$ (we will show that this is possible). Let $Z(k)$ denote the change of the value of the objective function. We show that $Z(k) < Z(k+1) < 0$.

Consider k. To find ε and η, we need to solve

$$a(k)\varepsilon - a(k')\eta = \delta,$$
$$b(k)\varepsilon - b(k')\eta = \delta.$$

The solution to this system is

$$\varepsilon = \frac{a(k')-b(k')}{b(k)a(k')-b(k')a(k)}\delta,$$
$$\eta = \frac{a(k)}{a(k')}\frac{a(k')-b(k')}{b(k)a(k')-b(k')a(k)}\delta - \frac{1}{a(k')}\delta.$$

Notice that $a(k') - b(k') < 0$ and $a(k') < 0$. Thus, to show that $\varepsilon > 0$ and $\eta > 0$, it is necessary and sufficient that

$$b(k)a(k') - b(k')a(k) < 0.$$

To simplify the notation we define:

$$a_1(k) = \binom{n-1}{k-1}\frac{1}{k}(n(1-p)-k)qP(I)p^{n-k-1}$$
$$a_2(k) = \binom{n-1}{k-1}\frac{1}{k}(k-np)(1-q)P(G)p^{k}$$

so that

$$a(k) = a_1(k)(1-p)^k + a_2(k)(1-p)^{n-k-1}.$$

Similarly, define

$$b_1(k) = \binom{n-1}{k-1}\frac{1}{k}(k-n(1-p))qP(I)p^{n-k}$$
$$b_2(k) = \binom{n-1}{k-1}\frac{1}{k}(np-k)(1-q)P(G)p^{k-1}$$

so that

$$b(k) = b_1(k)(1-p)^{k-1} + b_2(k)(1-p)^{n-k}$$

Notice that

$$b_1(k)\,a_1(k') = b_1(k')\,a_1(k)$$
$$b_2(k)\,a_2(k') = b_2(k')\,a_2(k)$$

and so

$$\begin{gathered} b(k)\,a(k') - b(k')\,a(k) = \\ b_1(k)\,a_2(k')\,(1-p)^{n-k'+k-2} + b_2(k)\,a_1(k')\,(1-p)^{n-k+k'} \\ -b_1(k')\,a_2(k)\,(1-p)^{n-k+k'-2} - b_2(k')\,a_1(k)\,(1-p)^{n-k'+k}. \end{gathered}$$

Note that the smallest power of the term $(1-p)$ in the expression above is $n-k'+k-2$. Therefore, for p close to 1 the sign of $b(k)\,a(k') - b(k')\,a(k)$ coincides with the sign of $b_1(k)\,a_2(k')$, which is negative.

The total effect $Z(k)$ on the utility is then

$$\begin{gathered} Z(k) = v(k)\,\varepsilon - v(k')\,\eta = \\ \left(v(k) - v(k')\,\tfrac{a(k)}{a(k')}\right) \tfrac{a(k')-b(k')}{b(k)a(k')-b(k')a(k)}\delta + \tfrac{v(k')}{a(k')}\delta, \end{gathered}$$

which is negative. In a similar way, for $k+1$ we get

$$Z(k+1) = \left(v(k+1) - v(k')\,\frac{a(k+1)}{a(k')}\right)\frac{a(k')-b(k')}{b(k+1)\,a(k')-b(k')\,a(k+1)}\delta + \frac{v(k')}{a(k')}\delta.$$

Recall that we need to show that $Z(k+1) > Z(k)$. We subtract $\frac{v(k')}{a(k')}\delta$ from $Z(k)$ and $Z(k+1)$. We then multiply both terms by the positive quantity (recall $a(k') < 0$ and $b(k') > 0$)

$$\frac{a(k')}{a(k')-b(k')}\frac{1}{\delta}.$$

We need to show

$$\frac{v(k+1)\,a(k') - v(k')\,a(k+1)}{b(k+1)\,a(k') - b(k')\,a(k+1)} > \frac{v(k)\,a(k') - v(k')\,a(k)}{b(k)\,a(k') - b(k')\,a(k)}.$$

We multiply both sides by

$$[b(k+1)\,a(k') - b(k')\,a(k+1)]\,[b(k)\,a(k') - b(k')\,a(k)] > 0$$

and obtain

$$\begin{gathered} [v(k+1)\,a(k') - v(k')\,a(k+1)]\,[b(k)\,a(k') - b(k')\,a(k)] > \\ [v(k)\,a(k') - v(k')\,a(k)]\,[b(k+1)\,a(k') - b(k')\,a(k+1)]. \end{gathered} \tag{2}$$

Each side of the inequality contains several terms. However, as p approaches 1, it suffices to consider the terms with the smallest power of $(1-p)$ to determine whether the inequality is satisfied or not.

We now write

$$v(k) = v_1(k)(1-p)^k + v_2(k)(1-p)^{n-k},$$

where we define

$$v_1(k) = -\binom{n}{k} qP(I)\, p^{n-k},$$
$$v_2(k) = \binom{n}{k}(1-q)\, P(G)\, p^k.$$

Then,

$$v(k)\,a(k') - v(k')\,a(k) = v_1(k)\,a_1(k')(1-p)^{k+k'} + v_1(k)\,a_2(k')(1-p)^{k+n-k'-1} +$$
$$v_2(k)\,a_1(k')(1-p)^{n-k+k'} + v_2(k)\,a_2(k')(1-p)^{2n-k-k'-1} - v_1(k')\,a_1(k)(1-p)^{k+k'}$$
$$-v_1(k')\,a_2(k)(1-p)^{k'+n-k-1} - v_2(k')\,a_1(k)(1-p)^{n-k'+k} - v_2(k')\,a_2(k)(1-p)^{2n-k-k'-1}.$$

The smallest power of $(1-p)$ is $k+n-k'-1$ (similarly, if we switch k with $k+1$, the smallest power would be $k+n-k'$).

Consider now the LHS of inequality (2):

$$[v(k+1)\,a(k') - v(k')\,a(k+1)]\,[b(k)\,a(k') - b(k')\,a(k)].$$

The term with the smallest power of $(1-p)$ is $v_1(k+1)\,a_2(k')\,b_1(k)\,a_2(k')$ and that power is $2(n-k'-1+k)$.

Consider the RHS of inequality (2):

$$[v(k)\,a(k') - v(k')\,a(k)]\,[b(k+1)\,a(k') - b(k')\,a(k+1)].$$

The term with the smallest power of $(1-p)$ is $v_1(k)\,a_2(k')\,b_1(k+1)\,a_2(k')$ and that power is $2(n-k'-1+k)$.

Thus, the two sides have the same powers and we have to show that

$$v_1(k+1)\,b_1(k)\,(a_2(k'))^2 > v_1(k)\,b_1(k+1)\,(a_2(k'))^2.$$

We divide both sides by $(a_2(k'))^2$ and compute the value of

$$v_1(k+1)\,b_1(k) - v_1(k)\,b_1(k+1)$$

when $p=1$ (by continuity, the sign of the expression extends to p close to 1).

When $p=1$,

$$v_1(k+1)\,b_1(k) - v_1(k)\,b_1(k+1) =$$
$$(qP(I))^2\left[-\binom{n}{k+1}\binom{n-1}{k-1} + \binom{n}{k}\binom{n-1}{k}\right] =$$

$$(qP(I))^2\left[-\tfrac{n!}{(k+1)!(n-k-1)!}\tfrac{(n-1)!}{(k-1)!(n-k)!}+\tfrac{n!}{k!(n-k)!}\tfrac{(n-1)!}{k!(n-k-1)!}\right]=$$
$$(qP(I))^2\tfrac{n!(n-1)!}{(n-k-1)!(n-k)!(k!)^2}\left(-\tfrac{k}{k+1}+1\right)>0.$$

This concludes the proof of Fact A.

Proof of Fact B

In this section we will prove the following. Consider $k'=k_n\left(=\frac{n+1}{2}\right),\ldots,n-2$, $k''=k'+1,\ldots,n-1$ and $k=1,\ldots,\frac{n-1}{2}$. Consider two different courses of action. In the first one, we decrease $\gamma(k')$ by $\eta>0$ and increase the value of $\gamma(k)$ by $\varepsilon>0$ to increase the LHS of both constraints by the same (small) number $\delta>0$. Let $Z(k)$ denote the corresponding change of the value of the objective function (this is the case analyzed in the previous section). In the second course of action, we increase $\gamma(k')$ by $\eta>0$ and decrease the value of $\gamma(k'')$ by $\varepsilon>0$ to increase the LHS of both constraints by the same (small) number $\delta>0$. This will change the value of the objective function by $\bar{Z}(k'')$. We want to show that $\bar{Z}(k'')<Z(k)$. (Recall that $Z(k)<0$. Thus, the inequality $\bar{Z}(k'')<Z(k)$ will also prove the first part of Fact B.)

Consider the second course of action. We need to solve the following system of equations:

$$-a(k'')\varepsilon+a(k')\eta=\delta,$$
$$-b(k'')\varepsilon+b(k')\eta=\delta.$$

The solution is

$$\varepsilon=\tfrac{a(k')-b(k')}{b(k')a(k'')-b(k'')a(k')}\delta,$$
$$\eta=\tfrac{a(k'')}{a(k')}\tfrac{a(k')-b(k')}{b(k')a(k'')-b(k'')a(k')}\delta+\tfrac{1}{a(k')}\delta.$$

It is simple to check that when p is close to 1 both ε and η are positive. Notice also that the denominator of ε is negative.

The total effect on the objective function $\bar{Z}(k'')$ is equal to

$$\bar{Z}(k'')=v(k')\eta-v(k'')\varepsilon=$$
$$\left(v(k')\tfrac{a(k'')}{a(k')}-v(k'')\right)\tfrac{a(k')-b(k')}{b(k')a(k'')-b(k'')a(k')}\delta+\tfrac{v(k')}{a(k')}\delta.$$

Recall that $Z(k)$ is equal to

$$Z(k)=\left(v(k)-v(k')\frac{a(k)}{a(k')}\right)\frac{a(k')-b(k')}{b(k)a(k')-b(k')a(k)}\delta+\frac{v(k')}{a(k')}\delta.$$

We subtract $\frac{v(k')}{a(k')}\delta$ from both $\bar{Z}(k'')$ and $Z(k)$ and multiply both by $\delta\frac{a(k')}{a(k')-b(k')}>0$. It remains to show that

$$\frac{v(k)a(k')-v(k')a(k)}{b(k)a(k')-b(k')a(k)}>\frac{v(k')a(k'')-v(k'')a(k')}{b(k')a(k'')-b(k'')a(k')}.$$

We multiply both sides by $[b(k)a(k')-b(k')a(k)][b(k')a(k'')-b(k'')a(k')]>0$ and get

$$\begin{gathered}[v(k)a(k')-v(k')a(k)][b(k')a(k'')-b(k'')a(k')]>\\ [v(k')a(k'')-v(k'')a(k')][b(k)a(k')-b(k')a(k)].\end{gathered}\tag{3}$$

For each term inside the square brackets we now identify the element with the smallest power of $(1-p)$.

We already know from the previous section that for $[b(k)a(k')-b(k')a(k)]$ we select $b_1(k)a_2(k')(1-p)^{n-k'+k-2}$.

In a similar way, for $[b(k')a(k'')-b(k'')a(k')]$ we select $b_1(k')a_2(k'')(1-p)^{n-k''+k'-2}$.

Consider now $[v(k)a(k')-v(k')a(k)]$. We select $v_1(k)a_2(k')(1-p)^{k+n-k'-1}$.

Finally, consider $[v(k')a(k'')-v(k'')a(k')]$. We select

$$[v_2(k')a_2(k'')-v_2(k'')a_2(k')](1-p)^{2n-k'-k''-1}.$$

Thus for p close to 1, inequality (3) is satisfied if and only if the following inequality is satisfied:

$$v_1(k)a_2(k')b_1(k')a_2(k'')(1-p)^{2n-k''+k-3}>$$

$$[v_2(k')a_2(k'')-v_2(k'')a_2(k')]b_1(k)a_2(k')(1-p)^{3n+k-2k'-k''-3}.$$

The exponent of the RHS is strictly smaller than the exponent of the LHS. Thus, it suffices to show

$$[v_2(k')a_2(k'')-v_2(k'')a_2(k')]b_1(k)a_2(k')<0.$$

Notice that for p close to 1, $b_1(k)a_2(k')<0$. We now evaluate the difference $v_2(k')a_2(k'')-v_2(k'')a_2(k')$ at $p=1$ and show that it is positive. By continuity, the above inequality will be satisfied when p is close to 1.

When $p=1$,

$$\begin{gathered}v_2(k')a_2(k'')-v_2(k'')a_2(k')=\\ ((1-q)P(G))^2\left[\binom{n}{k'}\binom{n-1}{k''-1}\frac{k''-n}{k''}-\binom{n}{k''}\binom{n-1}{k'-1}\frac{k'-n}{k'}\right]=\\ ((1-q)P(G))^2\frac{n!(n-1)!}{(n-k')!(n-k'')!k'!k''!}(k''-k')>0.\end{gathered}$$

This concludes the proof of Fact B.

Proof of Fact C

Consider $k=1,\ldots,k_n-1\left(=\frac{n-1}{2}\right)$ and $k'=k_n\left(=\frac{n+1}{2}\right),\ldots,n-1$. Suppose that we want to decrease the value of $\gamma(k)$ and increase the value of $\gamma(k')$ to increase the

LHS of both constraints by the same positive amount δ. We now show that this is impossible.

If the change described above is possible then there exist $\varepsilon > 0$ and $\eta > 0$ that solve the following system

$$-a(k)\varepsilon + a(k')\eta = \delta,$$
$$-b(k)\varepsilon + b(k')\eta = \delta.$$

The solution is

$$\varepsilon = \tfrac{a(k')-b(k')}{b(k')a(k)-b(k)a(k')}\delta,$$
$$\eta = \tfrac{a(k)}{a(k')}\tfrac{a(k')-b(k')}{b(k')a(k)-b(k)a(k')}\delta + \tfrac{1}{a(k')}\delta.$$

Notice that $a(k') - b(k') < 0$. Moreover, we know from the analysis above that for p close to 1 the sign of

$$b(k')a(k) - b(k)a(k')$$

coincides with the sign of $-b_1(k)a_2(k')$, which is positive. Thus, ε and η must be negative.

Proof of Fact D

Consider $k = k_n, \ldots, n-1$. Suppose that we want to increase both the value of $\gamma(k)$ and the value of $\gamma(0)$ to increase the LHS of both constraints by the same positive amount δ. We now show that this is impossible.

If the change described above is possible then there exist $\varepsilon > 0$ and $\eta > 0$ that solve the following system

$$a(0)\varepsilon + a(k)\eta = \delta,$$
$$b(0)\varepsilon + b(k)\eta = \delta.$$

The solution is

$$\varepsilon = \tfrac{a(k)-b(k)}{b(0)a(k)-b(k)a(0)}\delta,$$
$$\eta = -\tfrac{a(0)}{a(k)}\tfrac{a(k)-b(k)}{b(0)a(k)-b(k)a(0)}\delta + \tfrac{1}{a(k)}\delta.$$

Notice that $a(k) - b(k) < 0$. We now show that $b(0)a(k) - b(k)a(0)$ is positive, which implies that ε is negative.

Recall that

$$a(0) = -f(1;n) = qP(I)p^{n-1}(1-p) - (1-q)P(G)p(1-p)^{n-1}$$

and that

$$b(0) = f(0;n) = -qP(I)p^n + (1-q)P(G)(1-p)^n.$$

For p close to 1 the sign of $b(0)a(k) - b(k)a(0)$ coincides with the sign of $-qP(I)a_2(k)$ which is positive.

Proof of Fact E

Consider $k = k_n, \ldots, n-1$. Suppose that we want to decrease both the value of $\gamma(k)$ and the value of $\gamma(n)$ to increase the LHS of both constraints by the same positive amount δ. We now show that this is impossible.

If the change described above is possible then there exist $\varepsilon > 0$ and $\eta > 0$ that solve the following system

$$-a(k)\,\varepsilon - a(n)\,\eta = \delta,$$
$$-b(k)\,\varepsilon - b(n)\,\eta = \delta.$$

The solution is

$$\varepsilon = \tfrac{a(n)-b(n)}{b(n)a(k)-b(k)a(n)}\delta,$$
$$\eta = -\tfrac{a(k)}{a(n)}\tfrac{a(n)-b(n)}{b(n)a(k)-b(k)a(n)}\delta - \tfrac{1}{a(n)}\delta.$$

Recall that

$$a(n) = f(n;n) = -qP(I)(1-p)^n + (1-q)P(G)p^n$$

and that

$$b(n) = -f(n-1;n) = qP(I)\,p(1-p)^{n-1} - (1-q)P(G)p^{n-1}(1-p).$$

Define $a_1(n) = -qP(I)$ and $a_2(n) = (1-q)P(G)p^n$. Also, define $b_1(n) = qP(I)\,p$ and $b_2(n) = -(1-q)P(G)p^{n-1}$.

The numerator of ε is positive. We now show that the denominator of ε is negative.

We have to show $b(n)a(k) - b(k)a(n) < 0$ for p large. Notice that (after some simplifications)

$$b(n)a(k) - b(k)a(n) = b_1(n)a_2(k)(1-p)^{2n-k-2} + b_2(n)a_1(k)(1-p)^{k+1}$$
$$-b_1(k)a_2(n)(1-p)^{k-1} - b_2(k)a_1(n)(1-p)^{2n-k}.$$

The smallest power of $(1-p)$ is $k-1$, and thus for p close to 1 the sign of $b(n)a(k) - b(k)a(n)$ coincides with the sign of $-b_1(k)a_2(n)$ which is negative.

Remark 3 *Suppose that there exists a cost c' such that the optimal device takes the form*

$$\begin{aligned} &\gamma(0) = 0, \quad \gamma(1) = \ldots = \gamma(k_n - 1) = 1, \quad \gamma(k_n) = \ldots = \gamma(k'-1) = 0 \\ &\gamma(k') = \alpha \quad \gamma(k'+1) = \ldots = \gamma(n) = 1 \end{aligned} \tag{4}$$

then $k' = n-1$ and $\alpha < 1$.

Similarly, suppose that there exists a cost c'' such that the optimal device takes the form

$$\bar{\gamma}_n(0) = \ldots = \bar{\gamma}_n(k''-1) = 0, \quad \bar{\gamma}_n(k'') = \beta, \quad \bar{\gamma}_n(k''+1) = \ldots = \bar{\gamma}_n(k_n-1) = 1,$$
$$\bar{\gamma}_n(k_n) = \ldots = \bar{\gamma}_n(n-1) = 0, \quad \bar{\gamma}_n(n) = 1,$$

then $k'' = 1$ and $\beta > 0$.

An implication of the first part of the remark is the following. Suppose k' were smaller than $n-1$, and consider a cost c above c'. To satisfy the constraints, we could increase the value of $\gamma(k')$ and decrease the value of $\gamma(k)$ for some $k = k'+1, \ldots, n-1$. On the other hand, if $k' = n-1$ as claimed then it is impossible to modify the mechanism in order to satisfy both constraints. A similar implication follows from the second part of the remark and therefore the optimal device must take the form specified in Proposition 3.

Proof of Remark 3

We provide the proof for the first claim. The proof for the second claim is analogous.

To see that $k' = n-1$ when p is close to 1, consider the device described in (4). Both constraints are satisfied with equality. Thus,

$$f(1;n) - \binom{n-1}{\frac{n-1}{2}} f\left(\frac{n+1}{2}\right) + \alpha \binom{n-1}{k'-1} f(k';n) + (1-\alpha)\binom{n-1}{k'} f(k'+1;n) =$$
$$-f(0;n) + \binom{n-1}{\frac{n-1}{2}} f\left(\frac{n-1}{2}\right) - \alpha \binom{n-1}{k'-1} f(k'-1;n) - (1-\alpha)\binom{n-1}{k'} f(k';n)$$

(and both sides are equal to c'). Notice that as p approaches 1 the RHS of the equality converges to $qP(I)$ (since $-f(0;n)$ contains the term $qP(I)p^n$ and every other term contains $(1-p)^r$ for some $r > 0$). If $k' < n-1$, the LHS converges to zero (since each term contains $(1-p)^r$ for some $r > 0$) and the equality cannot be satisfied.

2 Distortionary Mechanisms when N is Fixed and p is Close to 1

In Proposition 2 we fix q, $P(I)$, p and let N go to infinity. In Proposition 3 and the notes above, we fix N and let p approach 1. The following Proposition extends Proposition 2 and provides conditions for the optimal extended mechanism to involve distortions when, indeed, N is fixed and p is large.

Proposition 2^* *Fix N, q and $P(I)$ and assume that either $qP(I) > 2(1-q)P(G)$ or $qP(I) < \frac{1}{2}(1-q)P(G)$. There exists $\tilde{p} < 1$ such that for every $p > \tilde{p}$ the following holds. For any $n = 2, ..., N$, suppose that the Bayesian device with n agents is admissible. Then there exists an admissible distortionary device with $n+1$ agents that yields greater expected utility than $\hat{V}(n)$.*

Proof of Proposition 2*

To simplify the notation, we define $D \equiv qP(I)$ and $E \equiv (1-q)P(G)$. The proof depends on which of the two cases specified in the proposition holds and on whether n is even or odd. We present the proof for the case $D > 2E$ and n odd (so that $n \geqslant 3$). The other three cases follow analogously.

When p is close to 1, and n is odd, then $k_n = \frac{n+1}{2}$. Moreover, $z(n)$ is strictly larger than $\frac{n}{2}$ but very close to $\frac{n}{2}$. In particular, $k_n - z(n) < \frac{1}{2}$.

We now adapt the proof of Proposition 2. Clearly, when p is close to 1, the inequalities used in the proof of Proposition 2: $k_n - 1 \geqslant n(1-p)$ and $k_n \leqslant np$, are satisfied. As in the proof of Proposition 2 we need to show that $\alpha_2 < \alpha^*$ and $\alpha_2 < \alpha_1$, where

$$\alpha_1 = \frac{\binom{n}{k_n} f(k_n+1; n+1) - \binom{n-1}{k_n-1} f(k_n; n)}{\binom{n}{k_n} f(k_n+1; n+1) - \binom{n}{k_n-1} f(k_n; n+1)},$$

$$\alpha_2 = \frac{\binom{n-1}{k_n-1} f(k_n; n) + \binom{n}{k_n} f(k_n; n+1)}{\binom{n}{k_n} f(k_n; n+1) - \binom{n}{k_n-1} f(k_n-1; n+1)},$$

and

$$\alpha^* = \frac{n-k_n+1}{n+1}.$$

The denominators of α_1 and α_2 are positive. We begin with the inequality $\alpha^* > \alpha_2$. We need to show

$$\left(n - \tfrac{n+1}{2} + 1\right)\left[\binom{n}{\frac{n+1}{2}} f\left(\tfrac{n+1}{2}; n+1\right) - \binom{n}{\frac{n+1}{2}-1} f\left(\tfrac{n+1}{2} - 1; n+1\right)\right] >$$
$$(n+1)\left[\binom{n-1}{\frac{n+1}{2}-1} f\left(\tfrac{n+1}{2}; n\right) + \binom{n}{\frac{n+1}{2}} f\left(\tfrac{n+1}{2}; n+1\right)\right].$$

The easiest way to show that the inequality is satisfied for p close to 1 is to identify, for each term $f(k'; n')$, the component with the smallest power of $(1-p)$.

For $f\left(\frac{n+1}{2}; n+1\right)$ we select $-D(1-p)^{\frac{n+1}{2}} p^{\frac{n+1}{2}} + E(1-p)^{\frac{n+1}{2}} p^{\frac{n+1}{2}}$.

For $f\left(\frac{n-1}{2}; n+1\right)$ we select $-D(1-p)^{\frac{n-1}{2}} p^{\frac{n+3}{2}}$.

For $f\left(\frac{n+1}{2}; n\right)$ we select $E(1-p)^{\frac{n-1}{2}} p^{\frac{n+1}{2}}$.

Thus, when p is sufficiently close to 1, the above inequality is satisfied if and only if

$$\frac{n+1}{2}\binom{n}{\frac{n-1}{2}} D > (n+1)\binom{n-1}{\frac{n-1}{2}} E$$

which is equivalent to

$$\frac{n}{n+1} D > E.$$

Clearly, if $D > 2E$ then the inequality is satisfied for every $n \geqslant 3$.

Consider now the inequality $\alpha_1 > \alpha_2$. We need to show (recall the denominators are positive):

$$\left[\binom{n}{\frac{n+1}{2}} f\left(\tfrac{n+3}{2}; n+1\right) - \binom{n-1}{\frac{n-1}{2}} f\left(\tfrac{n+1}{2}; n\right)\right] \left[\binom{n}{\frac{n+1}{2}} f\left(\tfrac{n+1}{2}; n+1\right) - \binom{n}{\frac{n-1}{2}} f\left(\tfrac{n-1}{2}; n+1\right)\right] >$$
$$\left[\binom{n-1}{\frac{n-1}{2}} f\left(\tfrac{n+1}{2}; n\right) + \binom{n}{\frac{n+1}{2}} f\left(\tfrac{n+1}{2}; n+1\right)\right] \left[\binom{n}{\frac{n+1}{2}} f\left(\tfrac{n+3}{2}; n+1\right) - \binom{n}{\frac{n-1}{2}} f\left(\tfrac{n+1}{2}; n+1\right)\right]$$

We proceed as above and identify the components with the smallest power of $(1-p)$.

For $f\left(\frac{n+3}{2}; n+1\right)$ we select $E\left(1-p\right)^{\frac{n-1}{2}} p^{\frac{n+3}{2}}$.

For $f\left(\frac{n+1}{2}; n+1\right)$ we select $-D\left(1-p\right)^{\frac{n+1}{2}} p^{\frac{n+1}{2}} + E\left(1-p\right)^{\frac{n+1}{2}} p^{\frac{n+1}{2}}$.

For $f\left(\frac{n-1}{2}; n+1\right)$ we select $-D\left(1-p\right)^{\frac{n-1}{2}} p^{\frac{n+3}{2}}$.

For $f\left(\frac{n+1}{2}; n\right)$ we select $E\left(1-p\right)^{\frac{n-1}{2}} p^{\frac{n+1}{2}}$.

Thus, we need to show

$$E\left[\binom{n}{\frac{n+1}{2}} - \binom{n-1}{\frac{n-1}{2}}\right] (1-p)^{\frac{n-1}{2}}\, D\binom{n}{\frac{n-1}{2}} (1-p)^{\frac{n-1}{2}} >$$
$$E\binom{n-1}{\frac{n-1}{2}} (1-p)^{\frac{n-1}{2}}\, E\binom{n}{\frac{n+1}{2}} (1-p)^{\frac{n-1}{2}}.$$

We divide both sides by $E\left(1-p\right)^{n-1}$ yielding

$$D\left[\binom{n}{\frac{n+1}{2}} - \binom{n-1}{\frac{n-1}{2}}\right]\binom{n}{\frac{n-1}{2}} > E\binom{n-1}{\frac{n-1}{2}}\binom{n}{\frac{n+1}{2}}$$

Notice that $\binom{n}{\frac{n+1}{2}} = \frac{n}{\frac{n+1}{2}}\binom{n-1}{\frac{n-1}{2}}$ and that $\binom{n}{\frac{n+1}{2}} = \binom{n}{\frac{n-1}{2}}$. Therefore, the inequality above translates into

$$D\binom{n-1}{\frac{n-1}{2}}\left[\frac{n}{\frac{n+1}{2}} - 1\right]\binom{n}{\frac{n-1}{2}} > E\binom{n-1}{\frac{n-1}{2}}\binom{n}{\frac{n+1}{2}}.$$

We divide both sides by $\binom{n-1}{\frac{n-1}{2}}\binom{n}{\frac{n-1}{2}}$ and get

$$D\left(\frac{2n}{n+1} - 1\right) > E$$

The inequality is satisfied for every odd $n \geqslant 3$ provided that $D > 2E$.

References

[1] Luenberger, D. G. [1965], *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, Reading, Massachusetts.

[2] Zhang, X. and Liu, D. [1990], "A Note on the Continuity of Solutions of Parametric Linear Programs," *Mathematical Programming*, Volume 47, pages 143-153.