Theoretical Economics 4 (2009), 227–252 1555-7561/20090227

A resurrection of the Condorcet Jury Theorem

Yukio Koriyama
Département d'Économie, École Polytechnique, Palaiseau

Balázs Szentes
Department of Economics, University College London

This paper analyzes the optimal size of a deliberating committee where (i) there is no conflict of interest among individuals and (ii) information acquisition is costly. The committee members simultaneously decide whether to acquire information, and then make the ex-post efficient decision. The optimal committee size, k^*, is shown to be bounded. The main result of this paper is that any arbitrarily large committee aggregates the decentralized information more efficiently than the committee of size k^*-2. This result implies that oversized committees generate only small inefficiencies.

Keywords. Voting, information aggregation, costly information.

JEL classification. D72, D81.

1. Introduction

The classical Condorcet Jury Theorem (CJT) states that large committees can aggregate decentralized information more efficiently than small ones. Its origin can be traced to the dawn of the French Revolution, when Marie-Jean-Antoine-Nicolas de Caritat, marquis de Condorcet (1785) investigated the decision-making processes in societies.[1] Recent literature on committee design has pointed out that if information acquisition is costly, the CJT fails to hold. The reasoning is that if the size of a committee is large, a committee member realizes that the probability that she can influence the final decision is small compared to the cost of information acquisition. As a result, she might prefer to avoid this cost and free-ride on the information of others. Therefore, large committees might generate lower social welfare than small ones. These results suggest that in the presence of costly information acquisition, optimally choosing the size of a committee is both an important and delicate issue. In this paper, we characterize the

Yukio Koriyama: yukio.koriyama@polytechnique.edu
Balázs Szentes: b.szentes@ucl.ac.uk

The authors are grateful to seminar participants at the University of Chicago, Berkeley, University of Rochester, École Polytechnique, Arizona State University, Maastricht University, University of Hawaii, and University College London for helpful comments. Financial support from the NSF is gratefully acknowledged.

[1]Summaries of the history of the CJT can be found in, for example, Grofman and Owen (1986), Miller (1986), and Gerling et al. (2005).

welfare loss associated with oversized committees and show that this loss is surprisingly small in certain environments. Therefore, as long as the committee size is large enough, the careful design of a committee might not be as important as it was originally thought to be. In fact, if either the information structure is ambiguous, or the committee has to make decisions in various informational environments, it might be optimal to design the committee to be as large as possible.

Committee design receives considerable attention from economists because, in many situations, groups rather than individuals make decisions. Information about the desirability of the possible decisions is often decentralized: individual group members must separately acquire costly information about the alternatives. A classical example is a jury trial where a jury has to decide whether a defendant is guilty or innocent. Each juror individually obtains some information about the defendant, at some cost of effort (paying attention to the trial, investigating evidence, etc.). Likewise, when a firm is facing the decision whether to implement a project, each member of the executive committee can collect information about the profitability of the project (by spending time and exerting effort). Yet another example is the decision of an academic department to hire a new member. Each member of the recruiting committee must review the applications individually before making a collective decision. What these examples have in common is the fact that information acquisition is costly and often unobservable.

The exact setup analyzed in this paper is described as follows. A group of individuals has to make a binary decision. There is no conflict of interest among the group members, but they have imperfect information about which decision is best. First, k individuals are asked to serve on a committee. Then the committee members simultaneously decide whether to invest in an informative signal. Finally, the committee makes the optimal decision given the acquired information. We do not explicitly model how the committee members communicate and aggregate information. Instead, we simply assume that they end up making the ex post efficient decision.[2] The only strategic choice an individual must make in our model is whether to acquire a signal upon being selected to serve on the committee.

The central question of our paper is the following: how does the committee size, k, affect social welfare? First, for each k, we fully characterize the set of equilibria (including asymmetric and mixed-strategy equilibria). We show that there exists k^P ($\in \mathbb{N}$) such that whenever $k \leq k^P$, there is a unique equilibrium in which each committee member invests in information with probability one. Furthermore, the social welfare generated by these equilibria is an increasing function of k. If $k > k^P$, then there are multiple equilibria, many of which involve randomization by the members. We show also that the social welfare generated by the *worst* equilibrium in the game, where the committee size is k, is decreasing in k if $k > k^P$. The optimal committee size k^* is defined as the smallest committee size such that there is an equilibrium in the committee with k^* members that maximizes social welfare among all equilibria in any committee. We

[2] Since there is no conflict of interest among the individuals, it is easy to design a mechanism that is incentive compatible and efficiently aggregates the signals. Alternatively, one can assume that the collected information is hard.

prove that the optimal committee size is either k^P or k^P+1. This implies that the CJT fails to hold: large committees can generate smaller social welfare than small committees. We show, nevertheless, that if the size of the committee is larger than k^*, even the worst equilibrium generates higher social welfare than the unique equilibrium in the committee of size k^*-2. That is, the welfare loss due to an oversized committee is quite small. For our results to hold, we need an assumption on the distribution of the signals. This assumption, stated formally in the next section, requires the marginal benefit from a signal to decrease fast.

Literature Review

Although the Condorcet Jury Theorem provides important support for the basis of democratic decision making, many of the premises of the theorem have been criticized. Perhaps most importantly, Condorcet assumes sincere voting. That is, each individual votes as if she were the only voter in the society. This means that an individual votes for the alternative that is best, conditional on her signal. Austen-Smith and Banks (1996) show that in general, sincere voting is not consistent with equilibrium behavior. This is because a rational individual votes conditional not only on her signal, but also on her being pivotal. Feddersen and Pesendorfer (1998) show that as the jury size increases, the probability of convicting an innocent can actually increase under unanimity rule.

A variety of papers show, however, that even if the voters are strategic, in certain environments the outcome of voting converges to the efficient outcome as the number of voters goes to infinity. Feddersen and Pesendorfer (1997) investigate a model in which preferences are heterogeneous and each voter has a private signal concerning which alternative is best. They construct an equilibrium for each population size, such that the equilibrium outcome converges to the *full information* outcome as the number of voters goes to infinity. The full information outcome is defined as the result of a voting game where all information is public. Myerson (1998) shows that asymptotic efficiency can be achieved even if there is population uncertainty; that is, if a voter does not know how many other voters there are.

In contrast, the Condorcet Jury Theorem might fail to hold if the information acquisition is costly. Mukhopadhaya (2003) considers a model similar to ours where voters have identical preferences, but information acquisition is costly. He shows by numerical examples that mixed-strategy equilibria in large committees may generate lower expected welfare than pure-strategy equilibria in small committees.[3]

Martinelli (2006) also introduces a cost of information acquisition into a model where signals are binary. He allows the precision of the signals to depend continuously on the amount of investment and proves that if the cost and the marginal cost of the precision are zero at the zero level of precision, then the decision is asymptotically efficient. More precisely, if the size of the committee converges to infinity, then there is

[3]The results are quite different if the voting, rather than the information acquisition, is costly: see e.g. Börgers (2004).

a sequence of symmetric equilibria in which each member invests only a little, and the probability of a correct decision converges to one.[4]

We think that Martinelli (2006) contributes substantially towards the understanding of the efficiency properties of group decision making when there is no fixed cost associated with information acquisition. However, we believe that the existence of a fixed cost is an essential feature of many environments. Indeed, one has to pay the price of a newspaper, even if it will be thrown away later. The management of a company has to pay for a consultant, even if the consultant's work will be completely ignored. Similarly, a juror has to sit through a trial, even if he decides not to pay any attention. What these examples have in common is that a positive cost must be invested in the signals even if the precision is zero. These examples also suggest that exerting more effort might lead to more accurate information, though this is not modeled in our paper. Hence, we think that the result in our paper is an important complement to Martinelli (2006). If information acquisition has fixed as well as variable costs, then an asymptotic efficiency result, like the one in Martinelli (2006), does not hold. However, we conjecture that our result is still valid in some form. That is, the efficiency loss due to large committees is small.[5]

Numerous papers analyze the optimal decision rules in the presence of costly information. Persico (2004) discusses the relationship between the optimal decision rules and the accuracy of the signals. He shows that a voting rule that requires a strong consensus in order to upset the status quo is optimal only if the signals are sufficiently accurate. The intuition for the extreme case, where the decision rule is unanimity rule, is the following. Under unanimity rule, the probability of being pivotal is small. However, this probability increases as the signals become more accurate. Therefore, in order to provide a voter with an incentive to invest in information, the signals must be sufficiently accurate.

Li (2001), Gerardi and Yariv (2008), and Gershkov and Szentes (forthcoming) show that the optimal voting mechanism sometimes involves ex post inefficient decisions. That is, the optimal mechanism might specify inefficient decisions for certain signal profiles. We believe that in many situations such a commitment device is not available, which is why we simply restrict attention to ex post efficient decision rules. We believe that this is the appropriate assumption in the context of a deliberating committee in which there is no conflict of interest among individuals.

Section 2 describes the model. The main theorems are stated and proved in Section 3. Section 4 concludes. Some of the proofs are relegated to the Appendix.

2. The model

There is a population consisting of N (> 1) individuals. The state of the world, ω, can take one of two values, 1 and -1, with $\Pr[\omega = 1] = \pi \in (0, 1)$. The society must make a

[4]In his accompanying paper, Martinelli (2007) analyzes a model in which information has a fixed cost, voters are heterogeneous in their costs, and abstention is not allowed. He shows that if, on the one hand, the support of the cost distribution is not bounded away from zero, asymptotic efficiency can be achieved. If, on the other hand, the cost is bounded away from zero, and the number of voters is large, nobody acquires information in any equilibrium.

[5]The difficulty of solving these models is that it is particularly hard to characterize the set of all equilibria.

decision, d, which is either 1 or -1. There is no conflict of interest among individuals. Each individual has a benefit of $u(d,\omega)$ if decision d is made when the state of the world is ω. In particular,

$$u(d,\omega)=\begin{cases}0 & \text{if } d=\omega\\ -q & \text{if } d=-1 \text{ and } \omega=1\\ -(1-q) & \text{if } d=1 \text{ and } \omega=-1,\end{cases}$$

where q $(\in(0,1))$, indicates the severity of a type-I error.[6] Each individual can purchase a signal at a cost c (>0) at most once. Signals are iid across individuals, conditional on the realization of the state of the world. The ex post payoff of an individual who invests in information is $u-c$. Each individual maximizes her expected payoff.

There are two stages of the decision-making process. At Stage 1, k $(\leq N)$ members of the society are designated at random to serve on a committee. At Stage 2, the committee members decide simultaneously and independently whether or not to invest in information. Then, the efficient decision is made given the signals collected by the members.

We do not model explicitly how committee members deliberate at Stage 2. Since there is no conflict of interest among the members, it is easy to design a communication protocol that efficiently aggregates information. Alternatively, one can assume that the acquired information is *hard*. Hence, no communication is necessary for making the ex post efficient decision. We focus merely on the committee members' incentives to acquire information.

Next, we turn to the definition of social welfare. First, let μ denote the ex post efficient decision rule. That is, μ is a mapping from sets of signals into possible decisions. If the signal profile is $(s_1,\ldots,s_n)$, where n is the number of signals acquired, then

$$\mu(s_1,\ldots,s_n)=1 \Leftrightarrow \mathbb{E}_\omega[u(1,\omega)\mid s_1,\ldots,s_n]\geq \mathbb{E}_\omega[u(-1,\omega)\mid s_1,\ldots,s_n].$$

Social welfare is measured by the expected sum of the payoffs of the individuals,

$$\mathbb{E}_{s_1,\ldots,s_n,\omega}[Nu(\mu(s_1,\ldots,s_n),\omega)-cn],$$

where the expectation also takes into account possible randomization by the individuals. That is, n can be a random variable.

If the committee is large, then a member might prefer to save the cost of information acquisition and choose to rely on the opinions of others. By contrast, if k is too small, there is too little information to aggregate, and thus the final decision is likely to be inefficient. The questions are: What value of k maximizes ex-ante social welfare? How big is the welfare loss in sub-optimal committees? Next, we define the optimal committee size formally.

[6] In the jury context, where $\omega=1$ corresponds to the innocence of the suspect, q indicates the severity of the error of convicting an innocent.

Definition. The *optimal size of the committee* is the smallest committee size k^* ($\in \mathbb{N}$) such that there is an equilibrium in the committee with k^* members that maximizes social welfare among all equilibria in any committee.

In our model, the optimal committee size is k^* if (i) there is an equilibrium in the committee with k^* members that maximizes social welfare among all equilibria in any committee, and (ii) each member acquires information with positive probability in this equilibrium. This is because a member who does not invest in information can be eliminated from the committee without changing the incentives of the other members.

We compute social welfare in all equilibria in suboptimal committees and compare the welfare loss in these committees. Since the signals are iid conditional on the state of the world, the expected benefit of an individual from the ex post efficient decision is a function of the number of signals acquired. We define this function as follows:

$$\eta(n) = \mathbb{E}_{s_1,\ldots,s_n,\omega}[u(\mu(s_1,\ldots,s_n),\omega)].$$

We assume that the signals are informative about the state of the world, but only imperfectly so. That is, as the number of signals goes to infinity, the probability of making the correct decision is strictly increasing and converges to one. Formally, the function η is strictly increasing and $\lim_{n\to\infty}\eta(n) = 0$. An individual's marginal benefit from collecting an additional signal, when n signals are already obtained, is

$$g(n) = \eta(n+1) - \eta(n).$$

Note that $\lim_{n\to\infty} g(n) = 0$. For our main theorem to hold, we need the following assumption.

Assumption 1. The function g is log-convex.[7] That is, $g(n+1)/g(n)$ is increasing in n ($\in \mathbb{N}$).

Whether or not this assumption is satisfied depends only on the primitives of the model—that is, on the distribution of the signals and on the parameters q and π. An immediate consequence of the assumption is the following.

Remark 1. The function g is decreasing.

Proof. Suppose to the contrary that there exists an integer $n_0 \in \mathbb{N}$ such that $g(n_0+1) > g(n_0)$. Since $g(n+1)/g(n)$ is increasing in n, it follows that $g(n+1) > g(n)$ whenever $n \geq n_0$. Hence $g(n) > g(n_0) > 0$ whenever $n > n_0$. This implies that $\lim_{n\to\infty} g(n) \neq 0$, which is a contradiction. $\square$

Next, we explain that Assumption 1 essentially means that the marginal value of a signal decreases *rapidly*. Notice that the function g being decreasing means that the

[7] The standard definition of convex functions requires the functions to have convex domains. This definition, however, can be naturally extended to functions with non-convex domains by requiring the convexity inequality to hold only on the domains (see Peters and Wakker 1987).

marginal social value of an additional signal is decreasing. We think that this assumption is satisfied in most economic and political applications. How much more does Assumption 1 require? Since g is decreasing and $\lim_{n\to\infty} g(n) = 0$, there always exists an increasing sequence $\{n_m\}_{m=1}^{\infty} \subset \mathbb{N}$ such that $g(n_m) - g(n_m + 1)$ is decreasing in m. Hence, it is still natural to restrict attention to information structures where the second difference in the social value of a signal, $g(n) - g(n+1)$, is decreasing. Recall that Assumption 1 is equivalent to $(g(n) - g(n+1))/g(n)$ being decreasing. That is, Assumption 1 requires that the second difference in the value of a signal not only decreases, but does so at an increasing rate.

In general, it is hard to check whether this assumption holds because it is often difficult (or impossible) to express $g(n)$ analytically. The next section provides examples where the assumption is satisfied.

2.1 *Examples for the log-convexity assumption*

First, suppose that the signals are normally distributed around the true state of the world. The log-convexity assumption is satisfied for the model where $\pi + q = 1$. That is, the society would be indifferent between the two possible decisions if information acquisition were impossible. The assumption is also satisfied even if $\pi + q \neq 1$, if the signals are sufficiently precise. Formally, we have the following result, which, like the other results in this section, is proved in the Appendix.

Proposition 1. *Suppose that* $s_i \sim N(\omega, \sigma)$.

(i) If $q + \pi = 1$ *then Assumption 1 is satisfied.*

(ii) For all q, π, *there exists* $\varepsilon_{q,\pi} > 0$ *such that Assumption 1 is satisfied if* $\varepsilon_{q,\pi} > \sigma$.

In our next example the signal is ternary—that is, its possible values are $\{-1, 0, 1\}$. In addition,

$$\Pr(s_i = \omega \mid \omega) = pr,\ \Pr(s_i = 0 \mid \omega) = 1 - r, \text{ and } \Pr(s_i = -\omega \mid \omega) = (1-p)r.$$

Notice that r $(\in (0,1))$ is the probability that the realization of the signal is informative, and p is the precision of the signal conditional on being informative.

Proposition 2. *Suppose that the signal is ternary. Then for any* $r (\in (0,1))$ *there exists a threshold* $\overline{p}(r) \in (0,1)$ *such that if* $p > \overline{p}(r)$, *Assumption 1 is satisfied.*

Next, we provide an example where the log-convexity assumption is not satisfied. Suppose that the signal is binary—that is, $s_i \in \{-1, 1\}$ and

$$\Pr(s_i = \omega \mid \omega) = p,\ \Pr(s_i = -\omega \mid \omega) = 1 - p.$$

Proposition 3. *If the signal is binary then Assumption 1 is not satisfied.*

Since binary signals are commonly assumed in the literature on committee design, we further explain the negative result in Proposition 3. For simplicity, assume that the prior is symmetric. In this case, the ex post optimal decision after receiving a set of signals is 1 if and only if there are (weakly) more signals 1 in the set than signals -1. The marginal social value of a signal can be positive only if the ex post optimal decision is different from the one without this signal with positive probability. This can happen only if there are as many signals 1 as signals -1 prior to receiving the additional signal, and in particular, if the number of signals is even. This implies that the marginal social values of the second, fourth, sixth, etc. signals are all zero. Hence, the function g is not even decreasing.

Notice that in this example, even if the cost of information is zero there always exists an equilibrium in which only one individual obtains a signal. Indeed, if an individual knows that there is exactly one signal collected by the others, he has no incentive to invest in information because he cannot have any effect on the optimality of the final decision. This argument is due to the fact that each signal realization has the same strength. Were an individual to know that his signal would be potentially very informative, he would invest if the cost is low enough, because, at least with small probability, his information induces a more informative decision.

Most of the literature assumes that committee members aggregate information by a binary voting procedure where members vote either yes or no (or abstain). Assuming binary signals in this case is perhaps less problematic because even if a member has a strong signal he cannot communicate it via his vote. We believe, however, that allowing variation in the strength of the signals is an important and realistic feature of the world. Allowing, for example, that a juror leans slightly towards a guilty verdict and at the same time another juror is sure that the defendant is innocent might have important consequences. We believe also that even if there is binary voting at the end of the deliberation, committee members are able to communicate their information to the others, at least partially. Focusing on the deliberation might be more important than analyzing the voting game.

3. Results

This section is devoted to our main theorems. We first characterize the set of equilibria for all k ($\in \mathbb{N}$). The next subsection shows that if k is small, the equilibrium is unique, and each member incurs the cost of information (Proposition 4). Section 3.2 describes the set of mixed-strategy equilibria for sufficiently large k (Proposition 5). Finally, Section 3.3 states and proves the main results (Theorems 1 and 2).

3.1 *Pure-strategy equilibrium*

Suppose that the size of the committee is k. If the first $k-1$ members acquire information, the expected *gain* from collecting information for the kth member is $g(k-1)$. She is willing to invest if this gain exceeds the cost of the signal; that is, if

$$c < g(k-1). \tag{1}$$

This inequality is the incentive compatibility constraint that guarantees that a committee member is willing to invest in information if the size of the committee is k.[8]

Proposition 4. *Let k denote the size of the committee. There exists $k^P \in \mathbb{N}$ such that there exists a unique equilibrium in which each member invests in a signal with probability one if and only if $k \le \min\{k^P, N\}$. Furthermore, the social welfare generated by these equilibria is monotonically increasing in k ($\le \min\{k^P, N\}$).*

Proof. Recall from Remark 1 that g is decreasing and $\lim_{k\to\infty} g(k) = 0$. Therefore, for any positive cost[9] $c < g(0)$, there exists a unique $k^P \in \mathbb{N}$ such that

$$g(k^P) < c < g(k^P - 1). \tag{2}$$

First, we show that if $k \le k^P$ then there is a unique equilibrium in which each committee member invests in information. Suppose that in an equilibrium, the first $k-1$ members randomize according to the profile $(r_1, \dots, r_{k-1})$, where $r_i \in [0,1]$ denotes the probability that the ith member invests. Let I denote the number of signals collected by the first $k-1$ members. Since the members randomize, I is a random variable. Notice that $I \le k-1$ and

$$\mathbb{E}_{r_1,\dots,r_{k-1}}[g(I)] \ge g(k-1)$$

because g is decreasing. Notice also that from $k \le k^P$ and (2), it follows that

$$g(k-1) > c.$$

Combining the previous two inequalities, we get

$$\mathbb{E}_{r_1,\dots,r_{k-1}}[g(I)] > c.$$

This inequality implies that no matter what the strategies of the first $k-1$ members are, the kth member strictly prefers to invest in information. From this observation, the existence and uniqueness of the pure-strategy equilibrium follow. It remains to show that if $k > k^P$, such a pure-strategy equilibrium does not exist. If $k > k^P$, then $g(k-1) < c$. Therefore, the incentive compatibility constraint, (1), is violated, and there is no equilibrium where each member incurs the cost of the signal.

Finally, we must show that the social welfare generated by these pure-strategy equilibria is increasing in k ($\le \min\{k^P, N\}$). Notice that since $N > 1$,

$$c < g(k-1) = \eta(k) - \eta(k-1) < N(\eta(k) - \eta(k-1)).$$

[8]In what follows, we ignore the case where there exists $k \in \mathbb{N}$ such that $c = g(k)$. This equality does not hold generically, and would have no effect on our results.

[9]If $c > \eta(1) - \eta(0)$, then nobody has an incentive to collect information, hence $k^P = 0$.

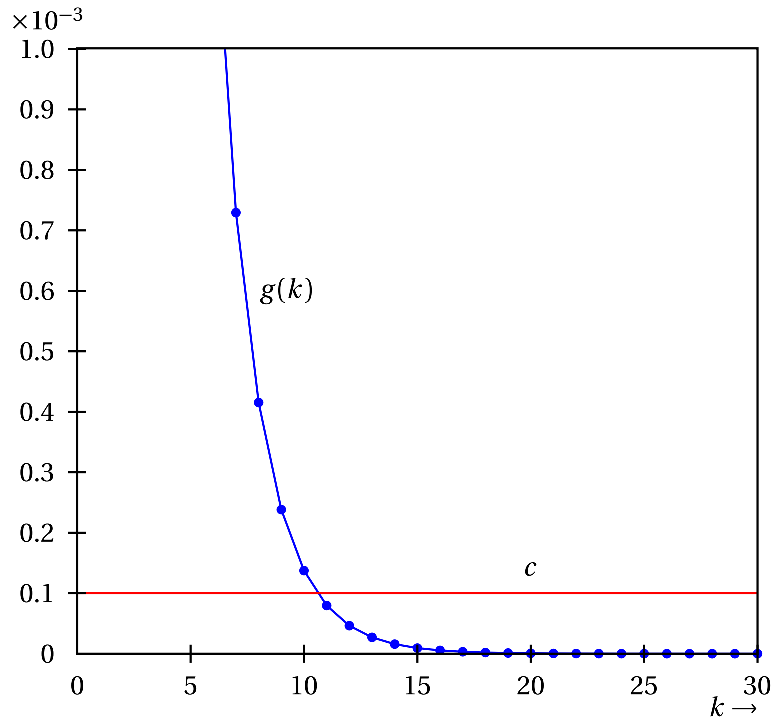

FIGURE 1. Expected gain $g(k)$ and the cost c.

After adding $N\eta(k-1)-ck$, we get

$$N\eta(k-1)-c(k-1)<N\eta(k)-ck.$$

The left-hand side is the social welfare generated by the equilibrium in a committee of size $k-1$, while the right-hand side is the social welfare induced by a committee of size k. □

Figure 1 is the graph of $g(k)$ and c when $s_i \sim N(\omega,1)$, $\pi=.3$, $q=.7$, and $c=10^{-4}$. The expected gain is decreasing and log-convex. In this example, $k^P=11$.

The amount of information purchased in any equilibrium is inefficiently small. This is because when a committee member decides whether or not to invest, she considers her private benefit rather than the society's benefit. Since information is a public good, its social benefit is bigger than its individual benefit. Hence, the total number of signals acquired in an equilibrium is smaller than the socially optimal number. This is why the social welfare is monotonically increasing in the committee size k as long as $k\le k^P$.

Mukhopadhaya (2003) has proved a result that corresponds to the statement of Proposition 4 for the case where the signals are binary. He has also shown by numerical examples that mixed-strategy equilibria can yield lower expected welfare in large committees than in small committees. Our analysis goes further by analytically comparing the expected welfare of all mixed-strategy equilibria.

3.2 *Mixed-strategy equilibrium*

Suppose now that the size of the committee is larger than k^P. We consider strategy profiles in which the committee members can randomize when making a decision about incurring the cost of information acquisition. The following proposition characterizes the set of mixed-strategy equilibria (including asymmetric ones).

We show that each equilibrium is characterized by a pair of integers (a,b). In the committee, a members invest in a signal with probability one and b members acquire information with a positive probability that is less than one. The rest of the members, $k-(a+b)$ in number, do not incur the cost. We call such an equilibrium a *type*-(a,b) *equilibrium*.

Proposition 5. *Let the committee size be* k $(> k^P)$. *There exists an equilibrium where* a *members invest for sure,* b *members invest with probability* $r \in (0,1)$, *and* $k-(a+b)$ *members do not invest, if and only if*

$$a \le k^P \le a+b \le k, \tag{3}$$

where the first two inequalities are strict whenever $b > 0$.

Proof. First, we explain that if, in an equilibrium in which one member invests with probability $r_1 \in (0,1)$ and another invests with probability $r_2 \in (0,1)$, then $r_1 = r_2$. Since the marginal benefit from an additional signal is decreasing, our games exhibit strategic substitution. That is, the more information the others acquire, the less incentive a member has to invest. Hence, if $r_1 < r_2$, then the individual who invests with probability r_1 faces more information in expectation and has less incentive to invest than the individual who invests with probability r_2. On the other hand, since $r_1, r_2 \in (0,1)$, both individuals must be exactly indifferent between investing and not investing, a contradiction. Now, we formalize this argument. Let $r_i \in [0,1]$ $(i = 1,\dots,k)$ be the probability that the ith member collects information in an equilibrium. Suppose that r_1, $r_2 \in (0,1)$ and $r_1 > r_2$. Let I_{-1} and I_{-2} denote the number of signals collected by members $2,3,\dots,k$ and by members $1,3,\dots,k$, respectively. Notice that since $r_1 > r_2$ and g is decreasing,

$$\mathbb{E}_{r_2,r_3,\dots,r_k}[g(I_{-1})] > \mathbb{E}_{r_1,r_3,\dots,r_k}[g(I_{-2})]. \tag{4}$$

On the other hand, a member who strictly randomizes must be indifferent between investing and not investing. Hence, for $j = 1,2$,

$$\mathbb{E}_{r_j,r_3,\dots,r_k}[g(I_{-j})] = c.$$

This equality implies that (4) should hold with equality, which is a contradiction. Therefore, each equilibrium can be characterized by a pair (a,b) where a members collect information for sure and b members randomize but collect information with the same probability.

It remains to show that there exists a type-(a,b) equilibrium if and only if (a,b) satisfies (3). First, notice that whenever $k > k^P$, in all pure-strategy equilibria k^P members invest with probability one and the remaining members never invest. In addition, the pair $(k^P, 0)$ satisfies (3). Therefore, we have to show only that there exists an equilibrium of type-(a,b) where $b > 0$ if and only if (a,b) satisfies

$$a < k^P < a + b \leq k. \tag{5}$$

Suppose that a members invest in information for sure and $b-1$ invest with probability r. Let $G(r;a,b)$ denote the expected gain from acquiring information for the $(a+b)$th member. That is,

$$G(r;a,b) = \sum_{i=0}^{b-1} \binom{b-1}{i} r^i (1-r)^{b-1-i} g(a+i).$$

We claim that there exists a type-(a,b) equilibrium if and only if there exists $r \in (0,1)$ such that $G(r;a,b) = c$. Suppose first that such an r exists. We first argue that there exists a type-(a,b) equilibrium in which b members invest with probability r. This means that the b members, who are randomizing, are indifferent between investing and not investing. The a members who invest for sure strictly prefer to invest because the marginal gain from an additional signal exceeds $G(r;a,b)$. Similarly, those members who do not invest, who number $k-(a+b)$, are strictly better off not investing because their marginal gains are strictly smaller than $G(r;a,b)$. Next, we argue that if $G(r;a,b) = c$ does not have a solution in $(0,1)$, then there exists no type-(a,b) equilibrium. This immediately follows from the observation that if b members are strictly randomizing, they must be indifferent between investing and not investing, and hence $G(r;a,b) = c$. Therefore, it is sufficient to show that $G(r;a,b) = c$ has a solution in $(0,1)$ if and only if (5) holds.

Notice that $G(r;a,b)$ is strictly decreasing in r because g is strictly decreasing. Also observe that $G(0;a,b) = g(a)$ and $G(1;a,b) = g(a+b-1)$. By the Intermediate Value Theorem, $G(r;a,b) = c$ has a solution in $(0,1)$ if and only if $G(1;a,b) < c < G(0;a,b)$, which is equivalent to

$$g(a+b-1) < c < g(a). \tag{6}$$

Recall that k^P satisfies

$$g(k^P) < c < g(k^P - 1).$$

Since g is decreasing, (6) holds if and only if $a < k^P$ and $a + b > k^P$. That is, the two strict inequalities in (5) are satisfied. The last inequality in (5) must hold because $a+b$ cannot exceed the size of the committee, k. □

Figure 2 graphically represents the set of pairs (a,b) that satisfy (3).

According to the previous proposition, there are several equilibria in which more than k^P members acquire information with positive probability. A natural question to ask is: can these mixed-strategy equilibria be compared from the point of view of social welfare? The next proposition partially answers this question. We show that if one fixes the number of members who acquire information for sure, then as the number of randomizing members increases, the social welfare generated by the equilibrium decreases. This proposition plays an important role in determining the optimal size of the committee.

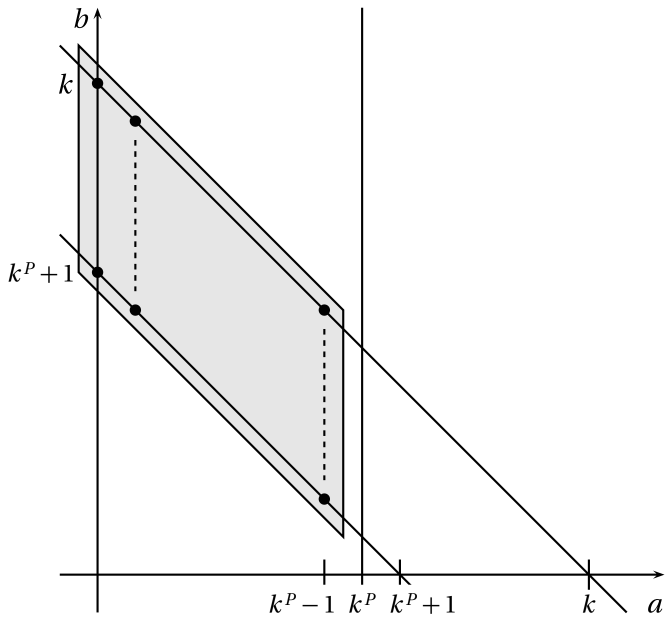

Figure 2. The set of mixed-strategy equilibria.

Proposition 6. *Suppose that $k \in \mathbb{N}$ is such that there are both type-(a,b) and type-$(a,b+1)$ equilibria. Then the type-(a,b) equilibrium generates strictly higher social welfare than the type-$(a,b+1)$ equilibrium.*

In order to prove this proposition we need the following result, which is proved in the Appendix.

Lemma 1. *(i) $G(r;a,b) > G(r;a,b+1)$ for all $r \in (0,1]$.*

(ii) $r_{a,b} > r_{a,b+1}$, where $r_{a,b}$ and $r_{a,b+1}$ are the solutions for r of $G(r;a,b)=c$ and $G(r;a,b+1)=c$ respectively.

Proof of Proposition 6. Suppose that a members collect information with probability one and b members do so with probability r. Let $f(r;a,b)$ denote the benefit of an individual; that is,

$$f(r;a,b)=\sum_{i=0}^{b}\binom{b}{i}r^i(1-r)^{b-i}\eta(a+i).$$

Clearly

$$\frac{\partial f(r;a,b)}{\partial r}=\sum_{i=1}^{b}\binom{b}{i}ir^{i-1}(1-r)^{b-i}\eta(a+i)-\sum_{i=0}^{b-1}\binom{b}{i}r^i(b-i)(1-r)^{b-i-1}\eta(a+i).$$

Notice that

$$\binom{b}{i}i=b\binom{b-1}{i-1} \quad \text{and} \quad \binom{b}{i}(b-i)=b\binom{b-1}{i}.$$

Therefore the right-hand side of the previous equality can be rewritten as

$$\sum_{i=1}^{b} b\binom{b-1}{i-1} r^{i-1}(1-r)^{b-i}\eta(a+i) - \sum_{i=0}^{b-1} b\binom{b-1}{i} r^{i}(1-r)^{b-i-1}\eta(a+i).$$

After changing the notation in the first summation, this can be further rewritten as

$$\begin{aligned}\sum_{i=0}^{b-1} b\binom{b-1}{i} r^{i}(1-r)^{b-i-1}\eta(a+i+1) - \sum_{i=0}^{b-1} b\binom{b-1}{i} r^{i}(1-r)^{b-i-1}\eta(a+i)\\ = b\sum_{i=0}^{b-1}\binom{b-1}{i} r^{i}(1-r)^{b-i-1}\left[\eta(a+i+1)-\eta(a+i)\right].\end{aligned}$$

This last expression is just $bG(r;a,b)$, and hence we have

$$\frac{\partial f(r;a,b)}{\partial r} = bG(r;a,b).$$

Next, we show that

$$f(r_{a,b};a,b) - f(r_{a,b+1};a,b+1) > b(r_{a,b} - r_{a,b+1})c. \tag{7}$$

Since $f(0;a,b) = f(0;a,b+1) = \eta(a)$,

$$\begin{aligned}&f(r_{a,b};a,b) - f(r_{a,b+1};a,b+1)\\ &\quad= \left[f(r_{a,b};a,b) - f(0;a,b)\right] - \left[f(r_{a,b+1};a,b+1) - f(0;a,b+1)\right]\\ &\quad= b\int_0^{r_{a,b}} G(r;a,b)\,dr - b\int_0^{r_{a,b+1}} G(r;a,b+1)\,dr.\end{aligned}$$

By part (i) of Lemma 1, this last difference is larger than

$$b\int_0^{r_{a,b}} G(r;a,b)\,dr - b\int_0^{r_{a,b+1}} G(r;a,b)\,dr = b\int_{r_{a,b+1}}^{r_{a,b}} G(r;a,b)\,dr.$$

By part (ii) of the lemma, we know that $r_{a,b+1} < r_{a,b}$. In addition, since G is decreasing in r, this last expression is larger than

$$b(r_{a,b} - r_{a,b+1})G(r_{a,b};a,b).$$

Recall that $r_{a,b}$ is defined such that $G(r_{a,b};a,b) = c$ and hence we can conclude (7).

Let $S(a,b)$ denote the social welfare in the type-(a,b) equilibrium; that is,

$$S(a,b) = Nf(r_{a,b};a,b) - c(a + br_{a,b}).$$

Then

$$\begin{aligned}S(a,b) - S(a,b+1) &= Nf(r_{a,b};a,b) - c(a+br_{a,b}) - [Nf(r_{a,b+1};a,b+1) - c(a+br_{a,b+1})]\\ &> Nb(r_{a,b} - r_{a,b+1})c - cb(r_{a,b} - r_{a,b+1})\\ &= (N-1)cb(r_{a,b} - r_{a,b+1}) > 0,\end{aligned}$$

where the first inequality follows from (7), and the last one follows from part (ii) of Lemma 1. □

3.3 *The main theorems*

First, we show that the optimal committee size is either k^P or k^P+1. Second, we prove that if $k>k^*$, then even the worst possible equilibrium yields higher social welfare than the unique equilibrium in the committee of size k^*-2.

Theorem 1. *The optimal committee size, k^*, is either k^P or k^P+1.*

We emphasize that for a certain set of parameter values, the optimal size is $k^*=k^P$, and for another set, $k^*=k^P+1$.

Proof. Suppose that k^* is the optimal size of the committee and the equilibrium that maximizes social welfare is of type-(a,b). By the definition of optimal size, $a+b=k^*$. If $b=0$, then all of the committee members invest in information in this equilibrium. From Proposition 4, $k^*\le k^P$ follows. In addition, Proposition 4 states that the social welfare is increasing in k as long as $k\le k^P$. Therefore, $k^*=k^P$ follows. Suppose now that $b>0$. If there exists an equilibrium of type-$(a,b-1)$, then, by Proposition 6, k^* is not the optimal committee size. Hence if the size of the committee is k^*, there does not exist an equilibrium of type-$(a,b-1)$. By Proposition 5, this implies that the pair $(a,b-1)$ violates the inequality chain (3) with $k=k^*$. Since the first and last inequalities in (3) hold because there is a type-(a,b) equilibrium, it must be the case that the second inequality is violated. That is, $k^P\ge a+b-1=k^*-1$. This implies that $k^*\le k^P+1$. Again, from Proposition 4, it follows that $k^*=k^P$ or k^P+1. □

Next, we turn our attention to the potential welfare loss due to oversized committees.

Theorem 2. *In any committee of size k $(>k^*)$, all equilibria induce higher social welfare than the unique equilibrium in the committee of size k^*-2.*

The following lemma plays an important role in the proof. We point out that the proof of this lemma is the only place where we use the log-convexity assumption (Assumption 1). For our previous results, we need the function g only to be decreasing, which is a consequence of Assumption 1.

Lemma 2. *For all $k\ge 1$ and $i\in\mathbb{N}$,*

$$g(k-1)\{g(i)-g(k)\}\ge\{g(k)-g(k-1)\}\{\eta(i)-\eta(k)\}, \tag{8}$$

with equality if and only if $i=k$ or $k-1$.

Proof of Theorem 2. Recall that $S(a,b)$ denotes the expected social welfare generated by an equilibrium of type-(a,b). Using this notation, we have to prove that $S(k^*-2,0)<S(a,b)$. From Theorem 1, we know that $k^*=k^P$ or k^P+1. By Proposition 4, $S(k^P-2,0)<S(k^P-1,0)$. Therefore, in order to establish $S(k^*-2,0)<S(a,b)$, it is enough to show that

$$S(k^P-1,0)<S(a,b) \tag{9}$$

for all pairs of (a,b) that satisfy (3).

Notice that if $a+i$ members invest in information, which happens with probability $\binom{b}{i} r_{a,b}^i (1-r_{a,b})^{b-i}$ in a type-(a,b) equilibrium, the social welfare is $N\eta(a+i)-c(a+i)$. Therefore,

$$
\begin{aligned}
S(a,b) &= \sum_{i=0}^{b} \binom{b}{i} {r_{a,b}}^i (1-r_{a,b})^{b-i} \Big[N\eta(a+i)-c(a+i)\Big] \\
&= \left\{ \sum_{i=0}^{b} \binom{b}{i} {r_{a,b}}^i (1-r_{a,b})^{b-i} \Big[N\eta(a+i)-ci\Big] \right\} - ca \\
&= N \left\{ \sum_{i=0}^{b} \binom{b}{i} {r_{a,b}}^i (1-r_{a,b})^{b-i} \eta(a+i) \right\} - c(a+br_{a,b}).
\end{aligned}
$$

In the last equation, we use the identity $\sum_{i=0}^{b} \binom{b}{i} {r_{a,b}}^i (1-r_{a,b})^{b-i} i = br_{a,b}$. Therefore, (9) can be rewritten as

$$
N\eta(k^P-1)-c(k^P-1) < N \left\{ \sum_{i=0}^{b} \binom{b}{i} r_{a,b}^i (1-r_{a,b})^{b-i} \eta(a+i) \right\} - c(a+br_{a,b}).
$$

Since $a \leq k^P-1$ by (3) and $b \leq N$, the right-hand side of this inequality is larger than

$$
N \left\{ \sum_{i=0}^{b} \binom{b}{i} r_{a,b}^i (1-r_{a,b})^{b-i} \eta(a+i) \right\} - c(k^P-1+Nr_{a,b}).
$$

Hence it suffices to show that

$$
N\eta(k^P-1)-c(k^P-1) < N \left\{ \sum_{i=0}^{b} \binom{b}{i} r_{a,b}^i (1-r_{a,b})^{b-i} \eta(a+i) \right\} - c(k^P-1+Nr_{a,b}).
$$

After adding $c(k^P-1)$ to both sides and dividing through by N, we have

$$
\eta(k^P-1) < \left\{ \sum_{i=0}^{b} \binom{b}{i} r_{a,b}^i (1-r_{a,b})^{b-i} \eta(a+i) \right\} - cr_{a,b}. \tag{10}
$$

The left-hand side is the payoff of an individual if k^P-1 signals are acquired by others, while the right-hand side is the payoff of an individual who is randomizing in a type-(a,b) equilibrium with probability $r_{a,b}$. Since this individual is indifferent between randomizing and not collecting information, the right-hand side of (10) can be rewritten as

$$
\sum_{i=0}^{b-1} \binom{b-1}{i} r_{a,b}^i (1-r_{a,b})^{b-1-i} \eta(a+i).
$$

Hence (10) is equivalent to

$$
\eta(k^P-1) < \sum_{i=0}^{b-1} \binom{b-1}{i} r_{a,b}^i (1-r_{a,b})^{b-1-i} \eta(a+i). \tag{11}
$$

By Lemma 2,

$$
\begin{aligned}
g(k^P-1)&\left\{\sum_{i=0}^{b-1}\binom{b-1}{i} r_{a,b}^i(1-r_{a,b})^{b-1-i} g(a+i)-g(k^P)\right\} \\
&\geq \{g(k^P)-g(k^P-1)\}\left\{\sum_{i=0}^{b-1}\binom{b-1}{i} r_{a,b}^i(1-r_{a,b})^{b-1-i}\eta(a+i)-\eta(k^P)\right\}.
\end{aligned}
\tag{12}
$$

Notice that

$$
\sum_{i=0}^{b-1}\binom{b-1}{i} r_{a,b}^i(1-r_{a,b})^{b-1-i} g(a+i)=c<g(k^P-1), \tag{13}
$$

where the equality guarantees that a member who is randomizing is indifferent between investing and not investing, and the inequality holds by (2). Hence from (12) and (13),

$$
\begin{aligned}
g(k^P-1)&\{g(k^P-1)-g(k^P)\} \\
&> \{g(k^P)-g(k^P-1)\}\left\{\sum_{i=0}^{b-1}\binom{b-1}{i} r_{a,b}^i(1-r_{a,b})^{b-1-i}\eta(a+i)-\eta(k^P)\right\}.
\end{aligned}
$$

Since $g(k^P-1)-g(k^P)>0$, this inequality is equivalent to

$$
g(k^P-1)>\eta(k^P)-\sum_{i=0}^{b-1}\binom{b-1}{i} r_{a,b}^i(1-r_{a,b})^{b-1-i}\eta(a+i).
$$

Finally, since $\eta(k^P)-g(k^P-1)=\eta(k^P-1)$, this inequality is just (11). □

The two graphs of Figure 3 show the social welfare in the worst equilibrium as a function of the committee size for an example. In the example, the prior is symmetric, and the parameters are $N=100$, $s_i \sim N(\omega,1)$, $\pi=.3$, $p=.7$, and $c=10^{-4}$. In this case we have $k^P=11$, and $k^*=12$. The two graphs are the same except that the scalings of the vertical axes differ. One can see that the welfare loss due to oversized committees is quite small.

4. Conclusion

In this paper, we discuss the optimal committee size and the potential welfare losses associated with oversized committees. We focus on environments in which there is no conflict of interest among individuals, but information acquisition is costly. First, we confirm that the optimal committee size is bounded. In other words, the Condorcet Jury Theorem fails to hold: larger committees might induce smaller social welfare. However, we show also that the welfare loss due to oversized committees is surprisingly small. In an arbitrarily large committee, even the worst equilibrium generates welfare higher than does an equilibrium in a committee with two less members than the optimal committee. Our results suggest that carefully designing committees might be not as important as has been thought.

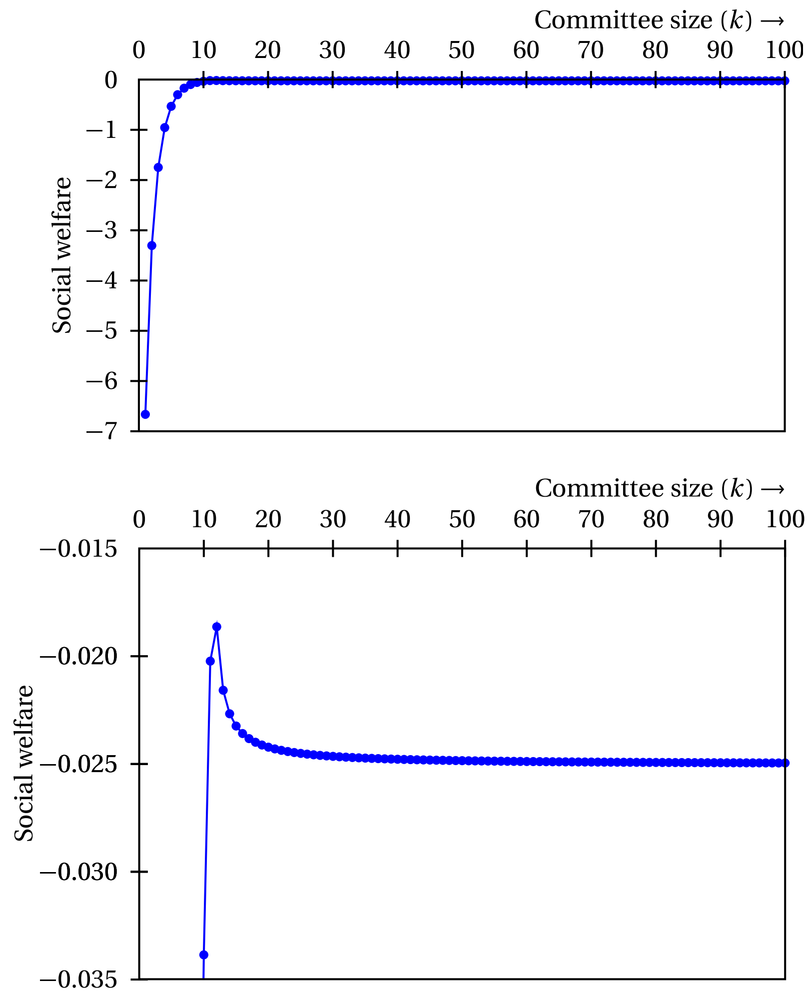

FIGURE 3. Social welfare as a function of the committee size k for $N = 100$, $\pi = 0.3$, $q = 0.7$, $\sigma = 1$, $c = 0.0001 \Rightarrow k^P = 11$, $k^* = 12$.

APPENDIX

LEMMA 3. *Suppose that η ($\in C^1(\mathbb{R}_+)$) is absolutely continuous and strictly increasing, and $\eta'(k+1)/\eta'(k)$ is strictly increasing for $k > \varepsilon$, where $\varepsilon \geq 0$. Let $g(k) = \eta(k+1) - \eta(k)$ for all $k \geq 0$. Then $g(k+1)/g(k) < g(k+2)/g(k+1)$ for $k \geq \varepsilon$.*

PROOF. Fix k ($\geq \varepsilon$). Notice that $\eta'(k+2)/\eta'(k+1) < \eta'(t+2)/\eta'(t+1)$ is equivalent to $\eta'(k+2)\eta'(t+1) < \eta'(k+1)\eta'(t+2)$. Therefore

$$\eta'(k+2)\int_k^{k+1} \eta'(t+1)\,dt < \eta'(k+1)\int_k^{k+1} \eta'(t+2)\,dt$$
$$\Leftrightarrow \eta'(k+2)[\eta(k+2) - \eta(k+1)] < \eta'(k+1)[\eta(k+3) - \eta(k+2)].$$

It follows that

$$\frac{\eta'(k+2)}{\eta'(k+1)} < \frac{\eta(k+3)-\eta(k+2)}{\eta(k+2)-\eta(k+1)} = \frac{g(k+2)}{g(k+1)}. \tag{14}$$

Similarly, for all $t \in (k,k+1)$, $\eta'(k+2)/\eta'(k+1) > \eta'(t+1)/\eta'(t)$ is equivalent to $\eta'(k+2)\eta'(t) > \eta'(k+1)\eta'(t+1)$. Therefore

$$\eta'(k+2)\int_k^{k+1} \eta'(t)\,dt > \eta'(k+1)\int_k^{k+1} \eta'(t+1)\,dt$$
$$\Leftrightarrow \eta'(k+2)[\eta(k+1)-\eta(k)] > \eta'(k+1)[\eta(k+2)-\eta(k+1)].$$

It follows that

$$\frac{\eta'(k+2)}{\eta'(k+1)} > \frac{\eta(k+2)-\eta(k+1)}{\eta(k+1)-\eta(k)} = \frac{g(k+1)}{g(k)}. \tag{15}$$

From (14) and (15) it follows that

$$\frac{g(k+1)}{g(k)} < \frac{g(k+2)}{g(k+1)}$$

for all $k \geq \varepsilon$. □

Proof of Proposition 1. The sum of normally distributed signals is also normal: $\sum_{i=1}^k s_i \sim N(\omega k, \sigma\sqrt{k})$. The density function of $\sum_{i=1}^k s_i$ conditional on ω is

$$\frac{1}{\sigma\sqrt{k}}\phi\left(\frac{(\sum_{i=1}^k s_i)-\omega k}{\sigma\sqrt{k}}\right)$$

where $\phi(x) = (2\pi)^{-1/2}\exp(-x^2/2)$. The ex post efficient decision rule is given by

$$\mu(s_1,\ldots,s_k) = \begin{cases} 1 & \text{if } s_1+\cdots+s_k \geq \theta \\ -1 & \text{if } s_1+\cdots+s_k < \theta, \end{cases}$$

where $\theta = (\sigma^2/2)\log[(1-q)(1-\pi)/q\pi]$ is the cut-off value. Hence, for $k \in \mathbb{N}\setminus\{0\}$,

$$\begin{aligned} \eta(k) &= -q\pi \Pr\left[\mu(s_1,\ldots,s_k) = -1 \mid \omega = 1\right] - (1-q)(1-\pi)\Pr\left[\mu(s_1,\ldots,s_k) = 1 \mid \omega = -1\right] \\ &= -q\pi\Phi\left(\frac{\theta-k}{\sigma\sqrt{k}}\right) - (1-q)(1-\pi)\Phi\left(-\frac{\theta+k}{\sigma\sqrt{k}}\right) \end{aligned} \tag{16}$$

where Φ is the cdf of standard normal distribution. If $k=0$,

$$\eta(0) = \max\{-q\pi, -(1-q)(1-\pi)\}. \tag{17}$$

Notice that the right-hand side of (16) converges to that of (17) as k goes to zero.

Part(i). If $q+\pi = 1$, then $q\pi = (1-q)(1-\pi)$ and $\theta = 0$. Hence

$$\eta(k) = -2q\pi\Phi\left(\frac{-\sqrt{k}}{\sigma}\right) \quad \text{and} \quad \eta'(k) = \frac{q\pi}{\sigma}\frac{1}{\sqrt{k}}\phi\left(\frac{\sqrt{k}}{\sigma}\right) \text{ for } k>0.$$

Therefore

$$\frac{\eta'(k+1)}{\eta'(k)} = \sqrt{\frac{k}{k+1}}\exp\left(-\frac{1}{2\sigma^2}\right)$$

is increasing in k (>0). From Lemma 3, setting ε to be zero, it follows that $g(k+1)/g(k)$ is increasing in $k \in \mathbb{N}$.

Part(ii). By the definition of θ,

$$q\pi\phi\left(\frac{\theta-k}{\sigma\sqrt{k}}\right) = (1-q)(1-\pi)\phi\left(-\frac{\theta+k}{\sigma\sqrt{k}}\right).$$

Applying this to take the derivative of (16), we get

$$\begin{aligned}\eta'(k) &= -q\pi\phi\left(\frac{\theta-k}{\sigma\sqrt{k}}\right)\left\{\frac{\partial}{\partial k}\left(\frac{\theta-k}{\sigma\sqrt{k}}\right) + \frac{\partial}{\partial k}\left(-\frac{\theta+k}{\sigma\sqrt{k}}\right)\right\}\\ &= q\pi\phi\left(\frac{\theta-k}{\sigma\sqrt{k}}\right)\frac{1}{\sigma\sqrt{k}}.\end{aligned}$$

Next, we argue that for any ε (>0), $\eta'(k+1)/\eta'(k)$ is increasing for all $k>\varepsilon$ if σ is sufficiently small. For $k>0$,

$$\begin{aligned}\frac{\eta'(k+1)}{\eta'(k)} &= \sqrt{\frac{k}{k+1}}\frac{\phi\left(\frac{\theta-(k+1)}{\sigma\sqrt{k+1}}\right)}{\phi\left(\frac{\theta-k}{\sigma\sqrt{k}}\right)} = \sqrt{\frac{k}{k+1}}\frac{\exp\left(-\frac{1}{2\sigma^2}\left(\frac{\theta^2}{k+1}-2\theta+k+1\right)\right)}{\exp\left(-\frac{1}{2\sigma^2}\left(\frac{\theta^2}{k}-2\theta+k\right)\right)}\\ &= \sqrt{\frac{k}{k+1}}\exp\left(\frac{1}{2\sigma^2}\left(\frac{\theta^2}{k(k+1)}-1\right)\right)\\ &= \sqrt{\frac{k}{k+1}}\exp\left(\frac{L^2\sigma^2}{8k(k+1)}\right)\exp\left(\frac{-1}{2\sigma^2}\right)\end{aligned} \tag{18}$$

where $L=\log\{(1-q)(1-\pi)/(q\pi)\}$. Now suppose that $k>\varepsilon$. The last term in (18) has no influence on whether $\eta'(k+1)/\eta'(k)$ is increasing. The second term converges to 1 as σ goes to 0. Obviously, the first term is strictly increasing in k. Hence, $\eta'(k+1)/\eta'(k)$ is increasing in k ($>\varepsilon$) if σ is sufficiently small. By setting $\varepsilon\in(0,1)$ and using Lemma 3, we have shown that $g(k+1)/g(k) < g(k+2)/g(k+1)$ for all $k\ge 1$.

It remains to show that $g(1)/g(0) < g(2)/g(1)$. From the argument in the proof of Lemma 3, it follows that $\eta'(2)/\eta'(1) < \eta'(t+1)/\eta'(t)$ for all $t\in(1,2)$ implies $\eta'(2)/\eta'(1) < g(2)/g(1)$. Hence it is enough to show that $g(1)/g(0) < \eta'(2)/\eta'(1)$ for sufficiently small σ. Since $\lim_{\sigma\to 0} g(0) = -\eta(0) > 0$, it is enough to show that $\lim_{\sigma\to0}[g(1)/\{\eta'(2)/\eta'(1)\}] = 0$. In order to establish this equality, it is obviously enough to show that

$$\lim_{\sigma\to 0}\frac{\eta(k)}{\eta'(2)/\eta'(1)} = 0 \text{ for } k\in\{1,2\}. \tag{19}$$

Remember $L=\log\{(1-q)(1-\pi)/(q\pi)\}$. By (16), for $k>0$,

$$\eta(k) = -q\pi\left\{\Phi\left(\frac{L\sigma}{2\sqrt{k}}-\frac{\sqrt{k}}{\sigma}\right)+\exp(L)\Phi\left(-\frac{L\sigma}{2\sqrt{k}}-\frac{\sqrt{k}}{\sigma}\right)\right\},$$

which implies $\eta(k) \in O^E(\Phi(-\sqrt{k}/\sigma))$ as $\sigma \to 0$.[10] Using (18), $\eta'(2)/\eta'(1) \in O^E(\exp(-1/2\sigma^{-2})) = O^E(\phi(1/\sigma))$ as $\sigma \to 0$. By l'Hôpital's Rule, for $k \in \{1,2\}$,

$$\lim_{\sigma\to 0} \frac{\Phi\big(-\sqrt{k}/\sigma\big)}{\phi(1/\sigma)} = \lim_{\sigma\to 0} \frac{\phi\big(-\sqrt{k}/\sigma\big)\big(\sqrt{k}/\sigma^2\big)}{\phi(1/\sigma)(1/\sigma^3)} = 0,$$

which implies (19). □

Proof of Proposition 2. First, we claim that the ex post efficient decision rule $\mu : \{-1,0,1\}^k \to \{-1,1\}$ is the cut-off rule

$$\mu(s_1,\dots,s_k) = \begin{cases} 1 & \text{if } \sum_{i=1}^k s_i \ge \widehat{\theta}, \\ -1 & \text{if } \sum_{i=1}^k s_i < \widehat{\theta}, \end{cases} \tag{20}$$

where $\widehat{\theta} = \log[(1-q)(1-\pi)/q\pi]/\log[p/(1-p)]$. Suppose that the signal sequence $(s_1,\dots,s_k)$ is a permutation of

$$\{\underbrace{1,\dots,1}_{a},\underbrace{0,\dots,0}_{k-a-b},\underbrace{-1,\dots,-1}_{b}\}. \tag{21}$$

Then $\mu(s_1,\dots,s_k)=1$ if

$$\begin{aligned}\mathbb{E}_\omega[u(\omega,1)\mid s_1,\dots,s_k] &= -(1-q)(1-\pi)\frac{\Pr[s_1,\dots,s_k\mid\omega=-1]}{\Pr[s_1,\dots,s_k]} \\ &> \mathbb{E}_\omega[u(\omega,-1)\mid s_1,\dots,s_k] = -q\pi\frac{\Pr[s_1,\dots,s_k\mid\omega=1]}{\Pr[s_1,\dots,s_k]}.\end{aligned}$$

In addition,

$$\begin{aligned}\Pr[s_1,\dots,s_k\mid\omega=1] &= (pr)^a(1-r)^{k-a-b}(r(1-p))^b \\ \Pr[s_1,\dots,s_k\mid\omega=-1] &= (pr)^b(1-r)^{k-a-b}(r(1-p))^a.\end{aligned}$$

Hence, $\mu(s_1,\dots,s_k)=1$ if

$$-(1-q)(1-\pi)p^b(1-p)^a > -q\pi p^a(1-p)^b,$$

or equivalently,

$$a-b > \frac{\log\big(\frac{(1-q)(1-\pi)}{q\pi}\big)}{\log\big(\frac{p}{1-p}\big)} = \widehat{\theta}.$$

Since $\sum_{i=1}^k s_i = a-b$, (20) follows.

Now we consider the case where p converges to 1. Let ε denote $1-p$ and let $\Pr[a,b]$ denote the probability of a signal sequence that is a permutation of (21). Then

$$\begin{aligned}\Pr[a,b\mid\omega=1] &= C_k(a,b)(1-\varepsilon)^a\varepsilon^b \\ \Pr[a,b\mid\omega=-1] &= C_k(a,b)(1-\varepsilon)^b\varepsilon^a,\end{aligned}$$

[10] O^E is a version of Landau's O, which describes the exact order of the expression. Formally, $f(x) \in O^E(g(x))$ as $x \to a$ if and only if there exists $M > 0$ such that $\lim_{x\to a}|f(x)/g(x)| = M$.

where $C_k(a,b) = [k!/(a!b!(k-a-b)!)]r^{a+b}(1-r)^{k-a-b}$.[11] Notice that $C_k(a,b)$ is independent of ε and symmetric with respect to a and b. We have[12]

$$\begin{aligned}
&\Pr[a-b \le -1 \mid \omega = 1] = C_k(0,1)\varepsilon + O(\varepsilon^2)\\
&\Pr[a-b \ge 0 \mid \omega = -1] = C_k(0,0) + \{C_k(1,0) + C_k(1,1)\}\varepsilon + O(\varepsilon^2).
\end{aligned}$$

Observe that $|\widehat{\theta}| < 1$ if p is close enough to one. Without loss of generality, assume that $q + \pi \ge 1$. Then $-1 < \widehat{\theta} \le 0$. Hence

$$\begin{aligned}
\eta(k) &= -q\pi \Pr[a-b \le -1 \mid \omega = 1] - (1-q)(1-\pi)\Pr[a-b \ge 0 \mid \omega = -1]\\
&= -q\pi C_k(0,1)\varepsilon - (1-q)(1-\pi)[C_k(0,0) + \{C_k(1,0) + C_k(1,1)\}\varepsilon] + O(\varepsilon^2).
\end{aligned}$$

Then

$$g(k) = \eta(k+1) - \eta(k) = A(k) + B(k)\varepsilon + O(\varepsilon^2),$$

where

$$\begin{aligned}
A(k) &= -(1-q)(1-\pi)D_k(0,0)\\
B(k) &= -q\pi D_k(0,1) - (1-q)(1-\pi)[D_k(1,0) + D_k(1,1)]
\end{aligned}$$

and $D_k(a,b) = C_{k+1}(a,b) - C_k(a,b)$. Using this notation,

$$\begin{aligned}
\frac{g(k+1)}{g(k)} &= \frac{A(k+1) + B(k+1)\varepsilon + O(\varepsilon^2)}{A(k) + B(k)\varepsilon + O(\varepsilon^2)} = \frac{A(k+1)\big(1 + \frac{B(k+1)}{A(k+1)}\varepsilon + O(\varepsilon^2)\big)}{A(k)\big(1 + \frac{B(k)}{A(k)}\varepsilon + O(\varepsilon^2)\big)}\\
&= \frac{A(k+1)}{A(k)}\left(1 + \left(\frac{B(k+1)}{A(k+1)} - \frac{B(k)}{A(k)}\right)\varepsilon\right) + O(\varepsilon^2).
\end{aligned}$$

We want to show that $g(k+1)/g(k)$ is increasing in k if ε is sufficiently small. Since $A(k+1)/A(k) = 1-r$, it is sufficient to show that $B(k)/A(k)$ is convex in k. It is straightforward to see that

$$\frac{D_k(0,1)}{D_k(0,0)} = \frac{D_k(1,0)}{D_k(0,0)} = \frac{(k+1)r(1-r)^k - kr(1-r)^{k-1}}{(1-r)^{k+1} - (1-r)^k} = \frac{(k+1)r(1-r) - kr}{(1-r)^2 - (1-r)}$$

is a polynomial of k with degree 1, hence it has no influence on the convexity of $B(k)/A(k)$. On the other hand,

$$\frac{D_k(1,1)}{D_k(0,0)} = \frac{(k+1)kr^2(1-r)^{k-1} - k(k-1)r^2(1-r)^{k-2}}{(1-r)^{k+1} - (1-r)^k} = \frac{(k+1)kr^2(1-r) - k(k-1)r^2}{(1-r)^3 - (1-r)^2}$$

has a positive coefficient of k^2. Hence we conclude that $B(k)/A(k)$ is convex in k. □

[11] Define $C_k(a,b) = 0$ if $k < a+b$.

[12] $f(x) \in O(g(x))$ as $x \to 0$ if and only if there exists $\delta > 0, M > 0$ such that $|x| < \delta$ implies $\big|f(x)/g(x)\big| < M$.

Proof of Proposition 3. As in the proof of Proposition 2, the ex post efficient decision rule is defined by

$$\mu(s)=\begin{cases}1 & \text{if } \sum s_i \geq \theta\\ -1 & \text{otherwise,}\end{cases}$$

where $\theta = \log[(1-q)(1-\pi)/q\pi]/\log[p/(1-p)]$. By symmetry, we can assume $\theta \geq 0$ without loss of generality. First, suppose $\theta > 1$. Then $\eta(k) = -q\pi$ for $k < \theta$, and the marginal benefit from an additional signal is zero for $k < \theta - 1$. Therefore $g(k+1)/g(k)$ is not well-defined. Second, suppose $0 \leq \theta < 1$. We consider two different cases depending on whether k is even or odd.

Case 1: Suppose $k = 2m$, where $m \in \mathbb{N}$. Then the $(2m+1)$-st signal makes a difference if and only if m of the first $2m$ signals are positive and m are negative, and the $(2m+1)$-st signal is positive (denote this situation as piv_e). In such a case, the social decision changes from -1 to 1. Hence the gain is q if $\omega = 1$ and the loss is $(1-q)$ if $\omega = -1$. Therefore, the expected marginal benefit is

$$\begin{aligned}
g(2m) &= q\Pr[\omega=1, piv_e] - (1-q)\Pr[\omega=-1, piv_e]\\
&= q\left\{\pi\binom{2m}{m}p^m(1-p)^m p\right\} - (1-q)\left\{(1-\pi)\binom{2m}{m}p^m(1-p)^m(1-p)\right\}\\
&= \{pq\pi - (1-p)(1-q)(1-\pi)\}\binom{2m}{m}p^m(1-p)^m.
\end{aligned}$$

Case 2: Suppose $k = 2m+1$, where $m \in \mathbb{N}$. Then the $(2m+2)$-nd signal makes a difference if and only if the first $2m+1$ signals contain $m+1$ positive and m negative signals and the $(2m+2)$-nd signal is negative (denote this situation as piv_o). In such a case, the social decision changes from 1 to -1. Hence the loss is q if $\omega = 1$, and the gain is $1-q$ if $\omega = -1$. Therefore the expected marginal benefit is

$$\begin{aligned}
g(2m+1) &= -q\Pr[\omega=1, piv_o] + (1-q)\Pr[\omega=-1, piv_o]\\
&= -q\left\{\pi\binom{2m+1}{m}p^{m+1}(1-p)^{m+1}\right\}\\
&\qquad\qquad + (1-q)\left\{(1-\pi)\binom{2m+1}{m}p^{m+1}(1-p)^{m+1}\right\}\\
&= \{-q\pi + (1-q)(1-\pi)\}\binom{2m+1}{m}p^{m+1}(1-p)^{m+1}.
\end{aligned}$$

Recall that $0 \leq \theta < 1$, which is equivalent to $pq\pi - (1-p)(1-q)(1-\pi) > 0$ and $(1-q)(1-\pi) - q\pi \geq 0$. If $\theta > 0$,

$$\frac{g(2m+2)}{g(2m+1)} = \frac{pq\pi - (1-p)(1-q)(1-\pi)}{(1-q)(1-\pi) - q\pi},$$

which is a constant function of m, so that Assumption 1 does not hold. If $\theta = 0$, then $g(2m+1) = 0$ and $g(2m+2)/g(2m+1)$ is not well-defined. □

Proof of Lemma 1. *(i)*. Notice that

$$G(r;a,b)=\sum_{i=0}^{b-1}\binom{b-1}{i}r^i(1-r)^{b-i-1}g(a+i).$$

Since $r^i(1-r)^{b-i-1}=r^i(1-r)^{b-i}+r^{i+1}(1-r)^{b-i-1}$,

$$G(r;a,b)=\sum_{i=0}^{b-1}\binom{b-1}{i}\left[r^i(1-r)^{b-i}+r^{i+1}(1-r)^{b-i-1}\right]g(a+i).$$

Since g is decreasing,

$$\begin{aligned}G(r;a,b)&>\sum_{i=0}^{b-1}\binom{b-1}{i}r^i(1-r)^{b-i}g(a+i)+\sum_{i=0}^{b-1}\binom{b-1}{i}r^{i+1}(1-r)^{b-i-1}g(a+i+1)\\&=\sum_{i=0}^{b-1}\binom{b-1}{i}r^i(1-r)^{b-i}g(a+i)+\sum_{i=1}^{b}\binom{b-1}{i-1}r^i(1-r)^{b-i}g(a+i)\\&=\sum_{i=0}^{b}\left[\binom{b-1}{i}+\binom{b-1}{i-1}\right]r^i(1-r)^{b-i}g(a+i),\end{aligned}$$

where the first equality holds because we have just redefined the notation in the second summation, and the second equality holds because, by convention, $\binom{n}{-1}=\binom{n}{n+1}=0$ for all $n\in\mathbb{N}$. Finally, using $\binom{b-1}{i}+\binom{b-1}{i-1}=\binom{b}{i}$, we have

$$G(r;a,b)>\sum_{i=0}^{b}\binom{b}{i}r^i(1-r)^{b-i}g(a+i)=G(r;a,b+1).$$

(ii). By the definitions of $r_{a,b}$ and $r_{a,b+1}$, we have

$$c=G(r_{a,b};a,b)=G(r_{a,b+1};a,b+1),$$

and by part (i) of this lemma,

$$G(r_{a,b+1};a,b+1)<G(r_{a,b+1};a,b).$$

Therefore

$$G(r_{a,b};a,b)<G(r_{a,b+1};a,b).$$

Since $G(r;a,b)$ is strictly decreasing in r, $r_{a,b}>r_{a,b+1}$ follows. □

Proof of Lemma 2. The statement of the lemma is obvious if $i\in\{k-1,k\}$. It remains to show that (8) holds with strict inequality whenever $i\notin\{k-1,k\}$. First, notice that for any positive sequence $\{a_j\}_0^\infty$, if $a_{j+1}/a_j<a_{j+2}/a_{j+1}$ for all $j\in\mathbb{N}$, then

$$\frac{a_k}{a_{k-1}}>\frac{\sum_{j=i+1}^{k}a_j}{\sum_{j=i}^{k-1}a_j}\text{ for all }k>1\text{ and for all }i\in\{0,\dots,k-2\}.$$

Assumption 1 allows us to apply this result to the sequence $a_j = g(j)$, and hence, for all $k \geq 1$,

$$\frac{g(k)}{g(k-1)} > \frac{\sum_{j=i+1}^{k} g(j)}{\sum_{j=i}^{k-1} g(j)} = \frac{\eta(k+1)-\eta(i+1)}{\eta(k)-\eta(i)} \text{ for all } i \in \{0,\ldots,k-2\}.$$

Since $\eta(a) > \eta(b)$ if $a > b$, this implies that for all $i \in \{0,\ldots,k-2\}$,

$$g(k)[\eta(i)-\eta(k)] < g(k-1)[\eta(i+1)-\eta(k+1)]. \tag{22}$$

Similarly, for a positive sequence $\{a_j\}_0^\infty$, if $a_{j+1}/a_j < a_{j+2}/a_{j+1}$ for all $j \in \mathbb{N}$, then

$$\frac{a_k}{a_{k-1}} < \frac{\sum_{j=k+1}^{i} a_j}{\sum_{j=k}^{i-1} a_j} \text{ for all } k \geq 1 \text{ and for all } i \geq k+1.$$

Again, by Assumption 1, we can apply this result to the sequence $a_j = g(j)$ and get

$$\frac{g(k)}{g(k-1)} < \frac{\sum_{j=k+1}^{i} g(j)}{\sum_{j=k}^{i-1} g(j)} = \frac{\eta(i+1)-\eta(k+1)}{\eta(i)-\eta(k)} \text{ for all } i > k.$$

Multiplying through by $g(k-1)(\eta(i)-\eta(k))$, we get (22). That is, (22) holds whenever $i \notin \{k-1, k\}$. After subtracting $g(k-1)(\eta(i)-\eta(k))$ from both sides of (22), we get (8). □

References

Austen-Smith, David and Jeffrey S. Banks (1996), "Information aggregation, rationality, and the Condorcet jury theorem." *American Political Science Review*, 90, 34–45. [229]

Börgers, Tilman (2004), "Costly voting." *American Economic Review*, 94, 57–66. [229]

Condorcet, marquis de (Marie-Jean-Antoine-Nicolas de Caritat) (1785), *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale, Paris. [227]

Feddersen, Timothy and Wolfgang Pesendorfer (1997), "Voting behavior and information aggregation in elections with private information." *Econometrica*, 65, 1029–1058. [229]

Feddersen, Timothy and Wolfgang Pesendorfer (1998), "Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting." *American Political Science Review*, 92, 23–35. [229]

Gerardi, Dino and Leeat Yariv (2008), "Information acquisition in committees." *Games and Economic Behavior*, 62, 436–459. [230]

Gerling, Kerstin, Hans Peter Grüner, Alexandra Kiel, and Elisabeth Schulte (2005), "Information acquisition and decision making in committees: A survey." *European Journal of Political Economy*, 21, 563–597. [227]

Gershkov, Alex and Balázs. Szentes (forthcoming), "Optimal voting schemes with costly information acquisition." *Journal of Economic Theory*. [230]

Grofman, Bernard and Guillermo Owen (1986), "Condorcet models, avenues for future research." In *Information Pooling and Group Decision Making* (Bernard Grofman and Guillermo Owen, eds.), 93–102, JAI Press, Greenwich, Connecticut. [227]

Li, Hao (2001), "A theory of conservatism." *Journal of Political Economy*, 109, 617–636. [230]

Martinelli, César (2006), "Would rational voters acquire costly information?" *Journal of Economic Theory*, 129, 225–251. [229, 230]

Martinelli, César (2007), "Rational ignorance and voting behavior." *International Journal of Game Theory*, 35, 315–335. [230]

Miller, Nicholas (1986), "Information, electorates, and democracy: Some extensions and interpretations of the Condorcet jury theorem." In *Information Pooling and Group Decision Making* (Bernard Grofman and Guillermo Owen, eds.), 173–192, JAI Press, Greenwich, Connecticut. [227]

Mukhopadhaya, Kaushik (2003), "Jury size and the free rider problem." *Journal of Law, Economics, and Organization*, 19, 24–44. [229, 236]

Myerson, Roger B. (1998), "Extended Poisson games and the Condorcet jury theorem." *Games and Economic Behavior*, 25, 111–131. [229]

Persico, Nicola (2004), "Committee design with endogenous information." *Review of Economic Studies*, 71, 165–191. [230]

Peters, Hans J. M. and Peter P. Wakker (1987), "Convex functions on non-convex domains." *Economic Letters*, 22, 251–255. [232]

Submitted 2009-5-5. Final version accepted 2009-5-26. Available online 2009-5-26.