Information Aggregation, Costly Voting and Common Values

Sayantan Ghosal and Ben Lockwood[*]
University of Warwick

January, 2002

Abstract

In a model of majority voting with common values and costly but voluntary participation, we show that too little information is aggregated at a voting equilibrium: in the vicinity of a voting equlibrium, it is *always* Pareto-improving for more agents, on the average, to vote. In addition, we show that multiple Pareto ranked voting equilibria may exist and moreover, majority voting with compulsory participation can Pareto dominate majority voting with voluntary participation.
Keywords: Voting, information, pivot, externality, inefficiency.
JEL Classification Numbers: D72, D82.

[*]We would like to thank Massimo Morelli for his comments. Address for correspondance: Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom. E-mails: S.Ghosal@warwick.ac.uk, and B.Lockwood@warwick.ac.uk.

1. Introduction

Many decisions are made by majority voting. In most cases, participation in the voting process is both voluntary and costly. The question then arises whether the level of participation is efficient i.e. is there too much or too little voting?

A recent paper, Borgers (2001) addresses this issue in a model with costly voting and private values. He identifies a negative “pivot” externality from voting: the decision of one voter to vote lowers the probability that any other voter is pivotal, and thus reduces the benefit to voting of all other agents. A striking result of the paper is that the negative “pivot” externality implies that compulsory voting is *never* desirable: all voters are strictly better off at the (unique) voluntary voting equilibrium. An implication of this global result is a local one: in the vicinity of an equilibrium it is *always* Pareto-improving for fewer agents, on the average, to vote.

In this paper, we re-examine the nature of inefficiency of majority voting in a model with common values, costly participation and symmetrically informed voters. In addition to the negative “pivot” externality identified by Borgers (2001), we identify a positive informational externality: an individual voter, by basing his voting decision on his informative signal, improves the quality of the collective decision for all voters. It follows that in our setup, a new issue arises: does majortity voting aggregagte too much or too little information? If at the voting equilibrium, there is too little voting, not enough information is aggrgated. On the other hand, if at a voting equilibrium, there is too much voting, too much information will be aggregated. On the face of it, it is not at all obvious, in general, which of these two externalities dominates. Consequently, it is not clear, in general, whether there is too much or too little information is aggregated by the voting process.

With costly voting, common values and symmetrically informed voters, we show, completely generally, that the second, positive informational externality, always dominates the first, negative “pivot” externality. We show that this reverses the nature of the inefficiency of voting equilibrium: in the vicinity of a voting equilibrium it is *always* Pareto-improving for more agents, on the average, to vote[1]. It follows that with costly participation, majority voting fails to aggregate information efficiently because not enough voters, on the average, participate in the voting process. One policy implication of our result is that by providing voting subsidies and recovering the cost through lump-sum taxes, all voters can

[1] Although the assumption that voters are symmetrically (partially) informed simplifies the statement of our results, allowing for asymmetrically informed non-partisan voters wouldn’t alter the nature of the inefficiency of voting equilibria we identify. (CHECK THIS)

be made better-off.

We show that an additional difference between the private and common value cases is that with common values, that there can be multiple voting equilibria. When three or more voting equilibria exist, we show that there must be at least two equilibria such that the equilibrium with more voters, on the average, Pareto dominates the equilibrium with fewer voters, on the average. We also show that there are conditions under which compulsory voting Pareto dominates voluntary majority voting. By contrast, in Borgers (2001), there is a unique voting equilibrium with voluntary majority voting; furthermore, voluntary majority voting always Pareto dominates compulsory majority voting.

This paper is related to a number of different literatures on voting. In a model of voting with common values with partisan voters and uninformed agents, Feddersen and Pessendorfer (1996)[2] show that voting does not aggregate information efficiently in a finite electorate. The presence of partisan voters and uninformed voters add noise to the voting process. In their set-up, voting fails to aggregate information efficiently as too many voters participate. In the absence of partisan voters, uninformed voters would abstain, implying that majority voting would aggregate information efficiently: as participation is costless, the remaining, symmetrically informed, non-partisan voters would all vote[3]. The above conclusion, leaves open the possibility that with common values even with symmetrically informed, non-partisan voters, majority voting may fail to aggregate information efficiently if participation is costly. However, the failure to aggregate information efficiently can occur in two mutually contradictory ways: efficient aggregation of information could fail either because there is too much voting or because there is too little voting. In contrast to Feddersen and Pessendorfer (1996), in our model, all agents are symmetric, receive (partially) informative signals, and have identical priors[4]. Our result that majority voting fails to aggregate information efficiently because, on the average, there is too little voting complements and extends Feddersen and Pessendorfer (1996).

A related strand of work focuses on information aggregation in juries (see, for instance, Austen-Smith and Banks (1996), Feddersen and Pessendorfer (1998)). The positive information externality we identify is similar to the information aggregation in a jury, except

[2]Other papers on information aggregation include Feddersen and Pessendorfer (1997) and Dekel and Piccione (2000).

[3]Remark that voting truthfully is a strictly dominant strategy for informed, non-partisan voters in Feddersen and Pessendorfer (1996).

[4]The assumption of identical priors rules out the possibility of partisan voters in our model. The assumption that all voters receive (partially) informative signals ensures that there are no uninformed voters in our model.

that the voting rule is majority, rather than unanimity. However, our characterization of the nature of the inefficiency of voting equilibria is new. Osborne and Rosenthal and Turner (2000) also study a model of costly participation. However, the focus of our paper and the formal model differs from their paper. They do not explicitly model voting and agents have complete information. Moreover, they do not consider the efficiency of participation equilibria.

In the next section, set out the model and prove some preliminary results. Section 3 analyses the inefficiency of voting equilibria. The last section concludes. Some technical material is gathered in the appendix.

2. The Model and Preliminary Results

2.1. The Set Up

The model of costly voting we study modifies Borgers (2001) to allow for common values. There is a set $N = \{1, ..n\}$ of agents, who can collectively choose between two alternatives, A and B. Voters have identical payoffs over alternatives, but their payoffs are state-dependent. Specifically, there are two states of nature s_A, s_B. The the payoff for all voters is a map $v : \{A, B\} \times \{s_A, s_B\} \rightarrow \Re$ such that $v(A, s_A) = v(B, s_B) = 1$ and 0 otherwise.

Agents have identical priors over the two states: all believe that each state is equally likely. However, prior to the decision to vote, voters receive private signals about the state of nature. Specifically, each $i \in N$ privately observes signal $\sigma^i \in \{\sigma_A, \sigma_B\}$ and σ^i is uncorrelated with σ^j for all $i, j \in N$. We assume that signals are informative i.e. the probability of signal σ_k, conditional on state s_k is $\rho > 0.5$, $k = A, B$.

We also assume that voting is costly. Specifically, each voter $i \in N$ incurs a privately observed cost of voting, c^i, if he chooses to vote. If he does not vote, he does not pay this cost. We assume that the c^i are independently and identically distributed across individuals: c^i is distributed on support $[\underline{c}, \bar{c}] \subset \Re_{++}$ with the probability distribution $F(c)$. Moreover, we also assume that c^i is independently distributed from σ^i for each individual i. The sequence of events is as follows.

Step 0. The state of the world is realized.

Step 1. Nature generates a pair (σ^i, c^i) for each $i \in N$ which is transmitted privately to each i.

Step 2. All $i \in N$ simultaneously choose to vote for A,vote for B,or abstain from voting.

Step 3. The alternative with the most votes is selected. If both A, B get equal numbers of votes, each is selected with probability 0.5.

Note that Step 3 embodies the assumption that there is no distinguished status quo. In particular, if no-one votes, each alternative is selected with equal probability.

As in Borgers (2001), voters are modelled symmetrically: all players have identical priors, receive partially informative signals. This rules out both partisan voters and uninformed voters as in Feddersen and Pessendorfer (1996) model with common values.

2.2. Voting Equilibrium

In the above environment, the n voters play a game of incomplete information. In this game, voters' strategies are maps from their private information to a decision at step 2. Formally, a strategy for i is of the form $\gamma^i : \{s_A, s_B\} \times [\underline{c}, \bar{c}] \to \{A, B, \text{abstain}\}$. We assume that voters behave *symmetrically* i.e. $\gamma^i = \gamma$, all $i \in N$, and we look for a symmetric Bayesian equilibrium[5]. We only look for equilibria with sincere voting i.e. each voter j votes for A (resp. B) iff $\sigma^j = \sigma_A$ (resp. σ_B), as any strategy where a voter votes insincerely is weakly dominated.

To characterize the equilibrium γ, we begin by calculating the benefit of voting relative to abstaining for some $i \in N$, conditional on his information (σ^i, c^i). Let p^j denote the ex-ante probability, before learning (σ^j, c^j), that any individual $j \neq i$ votes. By the assumed symmetry of strategies, $p^i = p$, all $i \in N$. So, the probability that exactly l other voters have chosen to vote is given by

$$v(l : p) = \binom{n-1}{l} p^l (1-p)^{n-1-l}. \tag{2.1}$$

We begin with the case when l is even and $l > 0$. Then, i is pivotal only when there is a tie i.e. exactly $\frac{l}{2}$ voters vote for A while the other $\frac{l}{2}$ voters vote for B. To calculate the probability of this event, we now have to calculate how i believes that others will vote, given his private information, (σ^i, c^i). Obviously, c^i is uninformative. If the true state of the world is s_A, then i believes that any $j \neq i$ will receive a signal σ^j with probability ρ and so will vote for alternative A with probability ρ. So, in this event, i assesses the probability that he is pivotal at

$$\pi(l : \rho) = \binom{l}{\frac{l}{2}} \rho^{\frac{l}{2}} (1-\rho)^{\frac{l}{2}} \tag{2.2}$$

[5] We are, therefore, ignoring assymmetric Bayesian equilibria.

If on the other hand, the true state of the world is s_B, i believes that $j \neq i$ will vote for alternative B with probability ρ. In this event, i assesses the probability that he is pivotal at (2.2) again. So, whatever the state of the world, and therefore whatever his private signal σ^i, i believes that the probability that he is pivotal, given l other voters, is simply (2.2).

In this event, what is his gain to voting? Recall that conditional on (say) $\sigma^i = \sigma_A$ he believes the probability that the state of the world is s_A is ρ. If he does not vote, both alternatives will be selected with probability 0.5, and this yields him a payoff of 0.5. If he does vote, how will he vote? As he is pivotal, he can deduce that the l agents who vote have received $l/2$ signals in favor of A, and $l/2$ in favor of B. In this case, the voting behavior of other players conveys no additional information to voter i about the state of the world, and so he votes according to his signal i.e. sincerely. So, alternative A will be selected, which gives him expected payoff ρ. So, his gain to voting (ignoring voting costs) is $\rho - 0.5$.

What happens if l is odd? In that case, the only situation where $i's$ vote matters is when $\frac{l+1}{2}$ voters have voted for one alternative, and $\frac{l-1}{2}$ for the other. Suppose w.l.o.g. that $\sigma^i = \sigma_A$. Then, it is possible to compute[6] that $\Pr(\sigma^j = \sigma_A | \sigma^i = \sigma_A) = \delta = \rho^2 + (1-\rho)^2$. That is, conditional on his signal, i believes that all l voters will vote for A with probability δ, so he calculates the probability that he is pivotal as:

$$\begin{aligned} \pi(l : \delta) &= \binom{l}{\frac{l+1}{2}} \delta^{\frac{l+1}{2}}(1-\delta)^{\frac{l-1}{2}} + \binom{l}{\frac{l-1}{2}} \delta^{\frac{l-1}{2}}(1-\delta)^{\frac{l+1}{2}} \\ &= \binom{l}{\frac{l+1}{2}} \delta^{\frac{l-1}{2}}(1-\delta)^{\frac{l-1}{2}} \end{aligned} \tag{2.3}$$

In this case, the voting behavior of other players *does* convey information to i. In particular, in the first case, player $i's$ updated information set has $\frac{l+1}{2}+1$ signals favoring A and $\frac{l-1}{2}$ signals favoring B. Therefore, voter i prefers A, but knows that he does not need to vote for alternative A to be selected; there is *already* a majority for alternative A. In

[6] The proof is as follows. By definition,

$$\begin{aligned} \delta &= \Pr(\sigma^j = \sigma_A | \sigma^i = \sigma_A) \\ &= \Pr(\sigma^j = \sigma_A | s_A)\Pr(s_A | \sigma^i = \sigma_A) + \Pr(\sigma^j = \sigma_A | s_B)\Pr(s_B | \sigma^i = \sigma_A) \\ &= \rho \Pr(s_A | \sigma^i = \sigma_A) + (1-\rho)\Pr(s_B | \sigma^i = \sigma_A) \\ &= \rho^2 + (1-\rho)^2 \end{aligned}$$

the second case, player $i's$ updated information set has $\frac{l-1}{2}+1$ signals favoring A and $\frac{l+1}{2}$ signals favoring B. Therefore, voter i is indifferent between A and B. So, we conclude that voter i has *zero gain from voting when l is odd, even when he is pivotal.*

The final case is where $l = 0$. In this case, i is pivotal with probability 1, so we set $\pi(0:\rho) = 1$. In this event, the gain to voting is the same as in the even case with $l > 0$. So, the preceding discussion implies that the expected gross gain to voting, conditional on some $\sigma^i \in \{\sigma_A, \sigma_B\}$ is:

$$B(p \mid \sigma^i) = (\rho - \frac{1}{2}) \sum_{l=0}^{n-1} v(l:p)\beta(l:\rho) \tag{2.4}$$

where

$$\beta(l) = \begin{cases} \pi(l:\rho) & l \text{ even} \\ 0 & l \text{ odd} \end{cases} \tag{2.5}$$

Note that $B(p|\sigma^i)$ is independent of the value of the signal σ^i, so from now on, we write $B(p|\sigma^i) \equiv B(p)$ without ambiguity.

So, it is now clear that if all other voters play a voting strategy γ with voting probability p, then $i's$ (strict) best response is to vote if $c_i < B(p)$ and not if $c_i > B(p)$. Following Borgers, we call this a *threshold strategy,* and we denote the threshold generally by $\hat{c}$. Generally, c^* is an *equilibrium* threshold strategy if $c \leq B(F(c^*))$, all $c \leq c^*$, and $c \geq B(F(c^*))$, all $c \geq c^*$. We can now show that there is at least one symmetric Bayesian equilibrium in threshold strategies.

Proposition 1. *There is at least one symmetric Bayesian equilibrium where voter i votes iff $c^i \leq c^*$. If c^* solves $B(F(c)) = c$, then c^* is an equilibrium threshold. If $B(1) > \overline{c}$, then $c^* = \overline{c}$ is an equilibrium threshold. If $\rho - \frac{1}{2} < \underline{c}$, then $c^* = \underline{c}$ is the unique equilibrium threshold. Finally,*

$$B(1) = \begin{cases} (\rho - \frac{1}{2}) \left(\begin{array}{c} n-1 \\ (n-1)/2 \end{array} \right) \rho^{\frac{n-1}{2}}(1-\rho)^{\frac{n-1}{2}} & \textit{if } n-1 \textit{ even} \\ 0 & \textit{if } n-1 \textit{ odd} \end{cases}$$

Proof. Existence of some equilibrium follows from the continuity of $B(F(.))$ on $[\underline{c}, \overline{c}]$. The remaining parts follow directly from the definition of equilibrium, except the last part. This follows from the fact that $\sum_{l \in E} v(l:p)\pi(l:\rho) < 1$, so $B(p) < B(0)$, all $p > 0$, so if $B(0) < \underline{c}$, neither of the other types of equilibria are possible. Finally, the formula for $B(1)$ follows from (1.1)-(1.5). □

This result leaves open the possibility that multiple equilibria exist, and the following example confirms this.

Example 1 (**Multiple Equilibria**).
Assume $n = 3$, and that c is uniform on $[0, \overline{c}]$. In this case,

$$B(p) = (\rho - 0.5)[2p^2\rho(1 - \rho) + (1 - p)^2] \tag{2.6}$$

Note that $p^* = F(c^*) = c^*/\overline{c}$, so assuming an interior equilibrium, the equilibrium condition $B(F(c)) = c$ can be rewritten in terms of p as $B(p) = p\overline{c}$, or explicitly as

$$(\rho - 0.5)[2p^2\rho(1 - \rho) + (1 - p)^2] = p\overline{c} \tag{2.7}$$

This is a quadratic in p, with two roots:

$$p = \frac{(2+\alpha) \pm \sqrt{(2+\alpha)^2 - 8\rho(1-\rho) - 4}}{2[2\rho(1-\rho)+1]} \tag{2.8}$$

where $\alpha = \overline{c}/(\rho - 0.5)$. If we take $\rho = 0.75$, and $\overline{c} = 0.09$, then it is easy to check that the two roots are

$$p^* = \frac{1.3119}{1.375},\ p^{**} = \frac{1.0481}{1.375}$$

i.e. the game has two interior equilibria. Note also for these numbers that $B(1) = 0.09375 > \overline{c}$, so there is also a corner equilibrium where $p^{***} = 1$. $\blacksquare$

This is illustrated below. It is clear from the Figure that multiple equilibria are due to the non-monotonicity of the benefit function $B(p)$. This is in contrast to the case of private values, where $B_{PV}(p)$ is strictly decreasing in p, and hence there is a unique equilibrium (Borgers(2001), Proposition 2).

Figure 1 in here

2.3. Comparing Common and Private Values

The equilibria of our model can in fact be compared to the voting equilibrium with private values in Borgers(2001). The first step is to note that the gain from voting with private values is, in our notation:

$$B_{PV}(p) = \frac{1}{2}\sum_{l=0}^{n-1} v(l : p)\pi(l : 0.5) \tag{2.9}$$

Note three differences between (2.4) and (2.9). First, in the private values case, there is a benefit to voting even when the number of voters is odd. Second ρ is replaced by 0.5 in $\pi(l : .)$ as any voter cannot predict how any other will vote, given that he decides to vote at all. As $\rho(1 - \rho)$ is maximized at $\rho = 0.5$, we can assert that $\pi(l : 0.5) > \pi(l : \rho)$, all $\rho \neq 0.5$. Finally, the benefit from one's most preferred alternative relative to random selection rises from $\rho - 0.5$ to 0.5 as in the private values case, voters are *sure* which alternatives are best. It is clear that all these three differences raise the benefit to voting in the private values case, so that it is always true that

$$B_{PV}(p) > B(p),\ 0 \leq p \leq 1 \tag{2.10}$$

Moreover, as shown by Borgers, $B_{PV}(p)$ is decreasing in p, and so there is always a unique equilibrium in the private values case. Let c_{PV} be the unique equilibrium cost' threshold in the private values case, and let $c_{\max}$ be the *highest* equilibrium cutoff in the common values case (this is well-defined by Proposition 1). Then we have:

Proposition 2. $c_{\max} \leq c_{PV}$, and $c_{\max} < c_{PV}$ *if* $c_{PV} < \bar{c}$.

Proof. Case 1. $B_{PV}(1) > \bar{c}$. Then $c_{PV} = \bar{c}$. Also, $B_{PV}(1) > B(1)$ by (2.10), so by Proposition 1, with common values, $c_{\max} \leq c_{PV}$.

Case 2. $B_{PV}(\tilde{c}) = \tilde{c}$, some $\tilde{c} \in [\underline{c}, \bar{c})$. In this case, as $B(p)$ lies everywhere below $B_{PV}(p)$, and $p = F(c)$ is increasing in c, $B(F(c)) < B_{PV}(F(c))$, $c \in [\underline{c}, \bar{c}]$. Therefore, if c^* *solves* $B(F(c)) = c$ then $c^* < \tilde{c}$ and in particular $c_{\max} < c_{PV}$. □

So, in a well-defined sense, a switch from private to common values lowers the probability of voting, and thus the fraction of the electorate who vote, when n is large.

3. The Inefficiency of Voting Equilibria

The central and striking claim of Borgers' paper is that the negative externalities from voting decisions imply that compulsory voting is never desirable. An implication of this "global" result is a "local" one: starting at the Bayes-Nash equilibrium cutoff c^*, it is *always* Pareto-improving to lower the cutoff slightly, so that fewer agents vote on average. This could be implemented (for example) by taxing voting and returning the revenue as a lump-sum.

By contrast, in the common values case, there are two externalities at work. First, the negative externality identified by Borgers is still present. We call it the *pivot* externality, as the voting decision by any i lowers the probability that j is pivotal. There is now another externality at work, however: when $j \neq i$ votes, i benefits from this decision as

j bases his voting decision on his informative signal. So, we call this the *informational* externality. This effect is similar to the information aggregation in a jury, except that the voting rule is majority, rather than unanimity.

To state these two externalities formally, first define $W(p)$ to be the expected utility of i if he does not vote, but every $j \neq i$ votes sincerely with probability p. Then, if i does vote, his expected utility is $W(p) + B(p)$. Then, if all voters are following a threshold strategy with cutoff $\hat{c}$, the ex ante expected utility of any voter (i.e. prior to observation of (σ^i, c^i)) is:

$$\begin{aligned} U(\hat{c}) &= (1 - F(\hat{c}))W(F(\hat{c})) + F(\hat{c})\left(W(F(\hat{c})) + B(F(\hat{c}))\right) - \int_{\underline{c}}^{\hat{c}} cf(c)dc \qquad (3.1) \\ &= W(F(\hat{c})) + F(\hat{c})B(F(\hat{c})) - \int_{\underline{c}}^{\hat{c}} cf(c)dc \end{aligned}$$

So,

$$\frac{dU(\hat{c})}{d\hat{c}} = \frac{dW(p)}{dp}f(\hat{c}) + F(\hat{c})\frac{dB(p)}{dp}f(\hat{c}) + B(F(\hat{c})) - \hat{c})f(\hat{c})$$

where $p = F(c^*)$. Evaluated at an interior Bayes-Nash equilibrium i.e. $\hat{c} = c^*$, $\underline{c} < c^* < \overline{c}$, the last term vanishes. So, we have

$$\frac{dU(\hat{c})}{d\hat{c}}\Big|_{\hat{c}=c^*} = [\frac{dW(p)}{dp} + F(c^*)\frac{dB(p)}{dp}]f(c^*) \qquad (3.2)$$

The right-hand side of equation (3.2) therefore measures the sum of the two external effects referred to above, the pivot externality and the information externality. The pivot externality, by definition, only affects the gain to voting, and is captured in the derivative $B'(p)$. The information externality, by contrast, affects the welfare of any voter whether he votes or not, and so is measured by $W'(p)$ and part of the derivative $B'(p)$. So, we should expect $W'(p)$ to be positive and $B'(p)$ to be of indeterminate sign. The latter has already been suggested as a possibility by Example 1, where B is non-monotonic in p.

Now, note that $W(p)$ can be calculated as follows. Suppose first that the state of the world is s_A. Let $\alpha(l)$ be the probability that a strict majority observe signal $\sigma^j = \sigma_A$ when l agents vote. Then, A is chosen, which gives i a payoff of 1. Next, note that $\beta(l)$ in (2.5) above can be interpreted as the probability that exactly $l/2$ of $j \neq i$ observe signal $\sigma^j = \sigma_A$ when l agents vote. In this case, A, B are chosen with equal probability, which gives i a payoff of 0.5. So, in the event that s_A, $i's$ expected payoff is

$$W(p) = \sum_{l=0}^{n-1} v(l : p)[\alpha(l) + 0.5\beta(l)]$$

Now, in the event that the state is s_B, the payoff is the same i.e. $W(p)$. So, the ex ante expected utility of i must also be $W(p)$.

Now note that conditional on s_A, the probability that any one voter observes signal $\sigma^j = \sigma_A$ is ρ, so we have, for $l > 0$:

$$\alpha(l) = \begin{cases} \sum_{z=\frac{l}{2}+1}^{l} \binom{l}{z} \rho^z (1 - \rho)^{l-z}, & l \text{ even} \\ \sum_{z=\frac{l+1}{2}}^{l} \binom{l}{z} \rho^z (1 - \rho)^{l-z}, & l \text{ odd} \end{cases}$$

and when $l = 0$, $\alpha(0) = 0$. We are now in a position to state the following key result. .

Lemma 1. *For all* $1 \geq \rho > 0.5$, *and all* $\hat{c} \in (\underline{c}, \overline{c})$,

$$\frac{dW(p)}{dp} + F(\hat{c})\frac{dB(p)}{dp} > 0,\ p = F(\hat{c}) \qquad (3.3)$$

Proof. See Appendix. □

This result says that for all possible voting probabilities strictly between zero and one, the information externality dominates the pivot externality. It is worth noting that as part of the proof of Lemma 1, it is demonstrated that $W'(p) > 0$ i.e. an increase in the voting probability unambiguously increases the expected utility of non-voters. As remarked above, this is due to the information externality, which wholly determines $W'(p)$. We now have the main result of the paper:

Proposition 3. *For all* $1 \geq \rho > 0.5$, *at any interior symmetric Bayesian equilibrium* $c^* \in (\underline{c}, \overline{c})$, *the information externality dominates the pivot externality. Consequently, a small increase in the cutoff* $\hat{c}$ *from* c^* *is always ex ante Pareto-improving.*

Proof. It is immediate from (3.2) and (3.3) that $\frac{dU(\hat{c})}{d\hat{c}}|_{\hat{c}=c^*} > 0$. □

Proposition 3 contrasts with Borgers' sharply results. His global result with private values establishes that it is never optimal to force agents to vote i.e. to raise $\hat{c}$ to $\overline{c}$. However, the proof of this result also establishes the local result that a small *decrease* in the cutoff $\hat{c}$ from c^* is always ex ante Pareto-improving. In this sense, Proposition 3 shows how a move from private values to common values reverses the nature of the inefficiency of voting equilibria. Proposition 3 also complements (and extends) Feddersen and Pessendorfer (1996)'s result on inefficient aggregation of information through the voting process in the presence of partisan voters and uninformed voters. In this paper partisan voters are ruled out by assumption. With costly voting and symmetrically informed voters, Proposition 3 shows that in the vicinity of a Bayesian equilibrium, the positive informational externality always dominates the negative "pivot" externality and therefore, the voting process fails to aggregate information efficiently with a finite electorate as too few voters, on the average, vote at a Bayesian equilibrium.

Now consider two symmetric voting rules with cutoffs c^* and c^{**} such that $c^* < c^{**}$. Then, the difference between the expected payoffs at the two equilibria can be written as

$$U(c^{**}) - U(c^*) = \int_{c^*}^{c^{**}} \left(\frac{dW(p)}{dp} + F(c)\frac{dB(p)}{dp} \right) f(c)dc + \int_{c^*}^{c^{**}} (B(F(c)) - c) f(c)dc, p = F(\hat{c}) \tag{3.4}$$

By Lemma 1, we know that the first integral is positive. However, the sign of the second integral is ambiguous as $B(p)$ is, in general, non-monotonic. This makes it impossible to obtain global characterizations with common values. In particular, we cannot show that, in general, a Bayesian equilibrium with a higher cutoff value Pareto dominates a Bayesian equilibrium with a lower cutoff value. In general, it is also not possible to show that compulsory majority voting Pareto dominates Bayesian equilibrium outcomes with voluntary majority voting. However, the following results can be stated.

Proposition 4. *Suppose that there are multiple voting equilibria as described by cutoffs:* $c_1 < .. < c_k < ... < c_m$*, such that either (i)* $m \geq 2$*, and* $c_{PV} < \bar{c}$ *or (ii)* $m \geq 3$*. Then, for some* k*,* $1 \leq k \leq m - 1$*, the Bayesian equilibrium with the higher cutoff* c_{k+1} *Pareto dominates the Bayesian equilibrium with the lower cutoff* c_k*.*

Proof of Proposition 4. As $c_{PV} < \bar{c}$ remark that at $p = F(c_m)$, $\frac{dB(p)}{dp} < 0$. As $m \geq 2$, it follows that there is at least one Bayesian equilibrium with cutoff c_k, for some k, $1 \leq k \leq m - 1$ so that $\frac{dB(p)}{dp} > 0$, $p = F(c_k)$ for some $k < m$. As $\frac{dB(p)}{dp} > 0$, $p = F(c_k)$, for some $k < m$, $B(F(c)) > c$, $c \in (c_k, c_{k+1})$. Alternatively, suppose there exist at least three Bayesian equilibria. Then, there is at least one Bayesian equilibrium with cutoff c_k so that $\frac{dB(p)}{dp} \geq 0$, $p = F(c_k)$ for some $k < m$. As $\frac{dB(p)}{dp} \geq 0$, $p = F(c_k)$, for some $k < m$, $B(F(c)) \geq c$, $c \in (c_k, c_{k+1})$.So, in both cases, from (3.4), $U(c_{k+1}) > U(c_k)$ i.e. the Bayesian equilibrium with the cutoff c_{k+1} Pareto dominates the Bayesian equilibrium with cutoff c_k. □

The preceding proposition shows that with three or more Bayesian equilibria, it is always possible to find atl east two equilibria so that the voting equilibrium with the higher cutoff Pareto dominates the voting equilibrium with the lower cutoff. The next proposition provides a sufficient condition for compulsory voting to Pareto dominate some other Bayesian equilibrium with a lower cutoff.

Proposition 5. *Suppose that there are multiple voting equilibria as described by cutoffs:*

$c_1 < c_2 < .. < c_m$, and moreover, $B(1) \geq \overline{c}$. Then $c_m = \overline{c}$, and this equilibrium Pareto-dominates equilibrium c_{m-1}. Consequently, starting at c_{m-1}, imposing compulsory voting is Pareto-improving.

Proof of Proposition 5. First, given $B(1) \geq \overline{c}$, the fact that $c_m = \overline{c}$ follows directly from Proposition 1. Also, by definition of c_m, c_{m-1}, $B(F(c)) \geq c$, $c \in (c_{m-1}, c_m)$. So, from (3.4), $U(\overline{c}) = U(c_m) > U(c_{m-1})$ i.e. compulsory voting Pareto-dominates voluntary voting equilibrium c_m. □

The hypotheses of Proposition 5 can easily arise if n is odd, as demonstrated by Example 1 (although not if n is even, as then from Proposition 1, $B(1) = 0$). Nevertheless, it could be argued that the scenario described in Proposition 5 of limited interest, as compulsory voting is simply a method of coordinating voting strategies on a different equilibrium (the one with $\hat{c} = \overline{c}$), and thus does not involve true compulsion.

The following example shows that true compulsion may also lead to a Pareto-improvement. It has a unique equilibrium with $\hat{c} < \overline{c}$, and starting at this equilibrium, imposing compulsory voting leads to a strict Pareto-improvement. .

Example 2 (Compulsory Voting May be Desirable).
The Example is the same as Example 1 i.e. $n = 3$ and uniform distribution of costs. Ex ante payoffs in this example can be computed from formula (3.1), which in this case simplifies to

$$U(p) = W(p) + pB(p) - \frac{1}{\overline{c}} \int_0^{p\overline{c}} c dc = W(p) + pB(p) - \overline{c}p^2/2$$

for any voting probability p. We already have computed a formula for $B(p)$ in Example 1. Also, note that

$$W(p) = \sum_{l=0} v(l : p)[\alpha(l) + 0.5\beta(l)] = 0.5(1 - p)^2 + 2p(1 - p)\rho + p^2\rho$$

So, using (2.6), we conclude that

$$U(p) = 0.5(1 - p)^2 + 2p(1 - p)\rho + p^2\rho + p(\rho - 0.5)[2p^2\rho(1 - \rho) + (1 - p)^2] - \overline{c}p^2/2 \quad (3.5)$$

Now let $\rho = 0.75$, and ψ be the value of c for which the larger root of (2.7) is equal to 1. This will be the value for which $B(1) = \psi$, and $B(1) = (\rho - 0.5)2\rho(1 - \rho) = 0.09375$. Then from Figure 1, it is clear that for $\overline{c} > \psi$, there will be a unique equilibrium given by the smaller root to (2.7): the larger root is greater than 1 and so cannot be an equilibrium probability. So, take $\overline{c} = 0.0938$. Then $\alpha = \overline{c}/(\rho - 0.5) = 0.3752$. In this case, there is a unique interior equilibrium with voting probability given by the smaller root to (2.8) i.e.

$$p^* = \frac{0.99947}{1.375} = 0.72689 \tag{3.6}$$

But then, from (3.5), after some simplification, we get:

$$U(p) = 0.5 + 0.75p \text{ ¡ } 0.7969p^2 + 0.34375p^3 \tag{3.7}$$

So, $U(1) = 0.79685 > 0.75613 = U(p^*)$ i.e. compulsory voting leads to a strict Pareto-improvement. Indeed, from (3.7), its can be shown that $U(p)$ is everywhere increasing in p 2 $[0, 1]$. k

Finally, we conclude this section by examining what happens when signals become uninformative. The following proposition shows that as signals become uninformative both the negative pivot externality and the positive information externality, and thus their overall effect, become negligible.

Proposition 6. *As signals become uninformative, i.e.* ρ ! 0.5, $\frac{dW(p)}{dp} + F(c^*)\frac{dB(p)}{dp}$ *tends to zero.*

Proof of Proposition 6. See Appendix. □

4. Conclusion

In this paper, we have shown that in a model of costly voting with common values, the nature of the inefficiency of voting equilibrium identified in Borgers (2001) is reversed: in the vicinity of a Bayesian equilibrium, it is *always* Pareto-improving for more agents, on the average, to vote. In addition, we have also shown that there Pareto ranked multiple Bayesian equilibria can exists and moreover, compulsory majority voting can Pareto dominate voluntary majority voting. The key behind all the results in this paper lies in the finding that there are two different externalities at work: the negative "pivot" externality identified by Borgers (2001) and the positive information externality. In the vicinity of a Bayesian equilibrium, the positive informational externality always outweighs the negative "pivot" externality implying that too few voters, on the average, participate in the voting process.

References

[1] Austen-Smith, D. and J.S. Banks (1996), "Information Aggregation, Rationality and the Condorcet jury theorem", *American Political Science Review, 90, pp.34-45.*

[2] Borgers, T. (2001), *"Costly Voting"*, mimeo, University College London.

[3] Dekel, E. and M. Piccione (2000), "Sequential Voting Procedures in Symmetric Binary Elections", *Journal of Political Economy, 108(1), pp.34-55.*

[4] Feddersen, T and W. Pessendorfer (1996), "The Swing Voter's Curse", *American Economic Review, 86(3), pp.3408-424.*

[5] Feddersen, T and W. Pessendorfer (1997), "Voting Behavior and Information Aggregation in Elections with Private Information", *Econometrica, 65(5), pp.1029-1058.*

[6] Feller, W. (1977), *"An introduction to probability theory and its applications, Volume I"*, Wiley Eastern Limited, New Delhi-Hyderabad.

[7] Lohmann, S. (1994), "Information Aggregation with Costly Political Action", *American Economic Review, 84(1), pp.518-530.*

[8] Osborne, M. and J. Rosenthal and M. Turner (2000), *"Meetings with Costly Participation"*, *American Economic Review, 90(4), pp.927-943.*

5. Appendix

Proof of Lemma 1. We, first, show that for $p' > p$,

$$W(p') - W(p) > 0.$$

Note that we can write for $p' > p$,

$$W(p') - W(p) = \sum_{l=0}^{n-1} [v(l : p') - v(l : p)][\alpha(l) + 0.5\beta(l)]$$

Now, by definition, of $\beta(l)$, plus the formulae for $\alpha(l)$ we have

$$W(p') - W(p) = \sum_{l \in T/E} [v(l : p') - v(l : p)] \left(\sum_{z=\frac{l+1}{2}}^{l} \binom{l}{z} \rho^z (1 - \rho)^{l-z} \right)$$
$$+ \sum_{l \in E} [v(l : p') - v(l : p)] \left(\sum_{z=\frac{l}{2}}^{l} \binom{l}{z} \rho^z (1 - \rho)^{l-z} - \frac{1}{2} \binom{l}{l/2} \rho^{l/2} (1 - \rho)^{l/2} \right)$$

were $T = \{0, 1, ..., n - 1\}$ and $E = \{x : x \textit{ is even and } x \leq n - 1\}$. This can be written more compactly as

$$W(p') - W(p) = \sum_{l \in T/E} [v(l : p') - v(l : p)](1 - G(\frac{l-1}{2}; \rho, l))$$
$$+ \sum_{l \in E} [v(l \;:\; p') - v(l : p)][(1 - G(\frac{l}{2} - 1; \rho, l)) - \frac{1}{2} g(\frac{l}{2}; \rho, l)] \tag{5.1}$$
$$= \sum_{l \in T/E} [v(l : p') - v(l : p)] G(\frac{l-1}{2}; 1 - \rho, l))$$
$$+ \sum_{l \in E} [v(l \;:\; p') - v(l : p)][G(\frac{l}{2} - 1; 1 - \rho, l)) - \frac{1}{2} g(\frac{l}{2}; \rho, l)] \tag{5.2}$$

where $G(x : \rho, l)$ is the cumulative distribution of the variable x distributed binomially with parameters ρ, l, and we have used the symmetry property of the Binomial that $1 - G(x : \rho, l) = G(x : 1 - \rho, l)$, and $g(x : \rho, l)$ is the distribution itself. As $v(. : p')$ first-order stochastically dominates $v(. : p)$, the RHS of (5.2) is positive as long as $G(\frac{l}{2} - 1; 1 - \rho, l), G(\frac{l-1}{2}; 1 - \rho, l)$ are increasing in l for $\rho > 0.5$ in the cases of l even and l odd respectively, and $g(\frac{l}{2}; \rho, l)$ is decreasing in l when l is even.

To prove the last statement, observe that

$$\frac{g(\frac{l}{2} + 1; \rho, l + 2)}{g(\frac{l}{2}; \rho, l)} = \frac{\rho(1 - \rho)(l + 2)}{4(l + 1)} \leq \frac{\rho(1 - \rho)}{3} < 1$$

as $l \geq 2$. To prove the statements about G, note from Feller (1968), page 173, equation 10.7, note that

$$G(x; 1 - \rho, l + 1) = G(x; 1 - \rho, l) - (1 - \rho) g(x; 1 - \rho, l) \tag{5.3}$$
$$G(x + 1; 1 - \rho, l + 1) = G(x; 1 - \rho, l) + \rho g(x + 1; 1 - \rho, l) \tag{5.4}$$

where $g(x; \rho, l)$ is the probability distribution function of the variable x distributed binomially with parameters ρ, l. First, consider the case where l is even. Using (5.3),(5.4), we have

$$G(\frac{l}{2}; 1 - \rho, l + 2) = G(\frac{l}{2}; 1 - \rho, l + 1) - (1 - \rho) g(\frac{l}{2}; 1 - \rho, l + 1) \tag{5.5}$$
$$= G(\frac{l}{2} - 1; 1 - \rho, l) + \rho g(\frac{l}{2}; 1 - \rho, l) - (1 - \rho) g(\frac{l}{2}; 1 - \rho, l + 1).$$

As $\rho > 0.5$, and $.g(\frac{l}{2}; 1 - \rho, l) > g(\frac{l}{2}; 1 - \rho, l + 1)$, we conclude from (5.5) that

$$G(\frac{l}{2} - 1; 1 - \rho, l) < G(\frac{l}{2}; 1 - \rho, l + 2). \tag{5.6}$$

as required. Now consider the case when l is odd. Again using (5.3),(5.4), it follows that

$$
\begin{aligned}
G(\frac{l-1}{2}+1; 1-\rho, l+2) &= G(\frac{l-1}{2}; 1-\rho, l+1) - (1-\rho)g(\frac{l-1}{2}; 1-\rho, l+1) \qquad (5.7) \\
&= G(\frac{l-1}{2}; 1-\rho, l) + \rho g(\frac{l-1}{2}; 1-\rho, l) - (1-\rho)g(\frac{l-1}{2}; 1-\rho, l+1).
\end{aligned}
$$

As $\rho > 0.5$, and $.g(\frac{l-1}{2}; 1-\rho, l) > g(\frac{l-1}{2}; 1-\rho, l+1)$, we conclude from (5.7) that,

$$
G(\frac{l-1}{2}; 1-\rho, l) < G(\frac{l-1}{2}; 1-\rho, l+2). \qquad (5.8)
$$

and therefore,

$$
W(p') - W(p) > 0
$$

Next, note that for $p' > p$,

$$
W(p') - W(p) + B(p') - B(p) = \sum_{l=0}^{n-1} [v(l:p') - v(l:p)][\alpha(l) + 0.5\beta(l) + (\rho - 0.5)\beta(l)] \qquad (5.9)
$$

By an argument identical for establishing (5.1) and (5.2), we obtain

$$
\begin{aligned}
& W(p') - W(p) + B(p') - B(p) \qquad (5.10) \\
= & \sum_{l \in T/E} [v(l:p') - v(l:p)]G(\frac{l-1}{2}; 1-\rho, l)) \\
+ & \sum_{l \in E} [v(l : p') - v(l:p)][G(\frac{l}{2} - 1; 1-\rho, l)) - (1-\rho)g(\frac{l}{2}; \rho, l)]. \qquad (5.11)
\end{aligned}
$$

As $v(.:p')$ first-order stochastically dominates $v(.:p)$, it immediately follows by an argument identical to the one establishing that $W(p') - W(p) > 0$, that the RHS of (5.11) is positive and therefore,

$$
W(p') - W(p) + B(p') - B(p) > 0.
$$

As both $W(p)$ and $B(p)$ are differentiable on $(\underline{c}, \overline{c})$ and $0 < p = F(\hat{c}) < 1$, it follows that $\frac{dW(p)}{dp} + F(\hat{c})\frac{dB(p)}{dp} > 0$, $p = F(\hat{c})$. □

Proof of Proposition 6 As $\rho \to 0.5$, at each $p \in [0, 1]$, $B(p) \to 0$ and therefore, at every Bayesian equilibrium, $c^* \to 0$. It follows that $F(c^*) \to 0$, and therefore, the pivot externality goes to zero. Also, when $\rho = 0.5$, $G(\frac{l-1}{2}; 1-\rho, l)) = 0.5$, l odd, $G(\frac{l}{2} - 1; 1-\rho, l)) = 0.5$, l even. So, the RHS of (5.11) becomes

$$
\rho \sum_{l \in T/E} [v(l:p') - v(l:p)]0.5 + \rho \sum_{l \in E} [v(l:p') - v(l:p)]0.5 = 0
$$

implying that

$$\frac{dW(p)}{dp} + F(c^*)\frac{dB(p)}{dp} = 0$$

at $\rho = 0.5$. As $F(c^*) = 0$, $\frac{dW(p)}{dp} = 0$. $\square$