# Deliberation, Preference Uncertainty and Voting Rules*

David Austen-Smith

MEDS, Kellogg Graduate School of Management

Northwestern University

Evanston, IL 60208

Timothy J. Feddersen

MEDS, Kellogg Graduate School of Management

Northwestern University

Evanston, IL 60208

July 2005, Revised November 2005

**Abstract**

A deliberative committee is a group of at least two individuals who first debate about what alternative to choose prior to these same individuals voting to determine the choice. We argue, first, that uncertainty about individuals' private preferences is necessary for full information sharing and, second, demonstrate in a very general setting that the condition under which unanimity can support full information revelation in debate amounts to it being common knowledge that all committee members invariably share identical preferences over the alternatives. It follows that if ever there exists an equilibrium with fully revealing debate under unanimity rule, there exists an equilibrium with fully revealing debate under any voting rule. Moreover, the converse is not true of majority rule if there is uncertainty about individuals' preferences.

Committees typically talk before making a decision. When does talking help committees make better decisions? Models of committee decision making show that when there is an underlying consensus, committee members comfortably share any relevant information and talking unambiguously improves committee performance independently of the committee's decision making rule (e.g. Coughlan, 2000). Such consensus, however, is unlikely to be the norm. And when there is no consensus, committee members may have an incentive not to disclose their information fully or accurately. Moreover, their incentives to share information may depend on details of the committee decision rule (e.g. Austen-Smith and Feddersen, 2005). Important questions have also been raised about what committees ought to talk about (see, for instance, Essay VI in Rawls, 1993); in particular, to what extent should talk focus on issues of common, rather than idiosyncratic, concern?

Answers to the preceding questions (among others) are central to the theory of deliberative democracy and to our understanding of exactly how talk affects political decision making. And while some progress has certainly been made on understanding the issues from a variety of analytical and substantive perspectives, there remains a great deal to be learned. In this paper, we address the questions in the context of the informational role of talk, paying attention to how individuals' private and public interests affect the sharing of information through talk prior to committee choices under various decision making rules.

We consider a committee that has to decide between two alternatives by voting. Members of the committee have private information about the relative merits of the alternatives and can talk, that is, deliberate, prior to the vote. An underlying consensus exists in the committee if, given the public revelation of all private information, committee members always agree on which alternative should be chosen. In this case, deliberation can result in the full sharing of information and a unanimous decision by the committee whatever voting rule is in place. However, if there is no underlying consensus, committee members might sometimes disagree on the best alternative despite all of the relevant information being shared. In this case, we say that committee members have different *biases*. And when committee members have different biases and these biases are common knowledge, full information sharing during deliberation is problematic (Coughlan, 2000; Meirowitz, 2005). In this paper, we demonstrate the impact of uncertainty regarding individuals' biases on the performance of deliberation under different voting rules. We show that such uncertainty allows full information sharing in deliberation under non-unanimous rules but not under unanimity rule.

The framework within which we address the issue is the basic Condorcet Jury Theorem model with incomplete information (see, for example, Austen-Smith and Banks, 1996; Feddersen and Pesendorfer, 1998; McLennan, 1998). Suppose, for example, there is a three person jury which is to decide the fate of a defendant who is either guilty or innocent with equal probability. The jurors agree that the best committee decision is to convict the defendant if guilty and to acquit him otherwise, and each of them privately receives an inconclusive, but informative, piece of evidence (a *signal*) concerning the truth. Where they disagree, however, is in exactly how much evidence suffices for conviction. Whereas two jurors are cautious (that is, have a *high bias* in favor of acquittal), willing to convict only if all three pieces of the available evidence suggest guilt, the third is less concerned about making a mistake (has a *low bias* in favor of acquittal) and is willing to convict if there is at least one such piece of evidence among the three.

When jurors' biases are known then, if they vote without talking the committee will, depending on the voting rule, sometimes make mistakes (Austen-Smith and Banks, 1996; Feddersen and Pesendorfer, 1998). The committee makes a mistake when it chooses an outcome that would not have been chosen under the given voting rule had all private information about the defendant's guilt or innocence been revealed prior to the vote. This is the case in our simple example. Under either majority or unanimity rule, the defendant would be convicted when all three private signals are revealed only if all of the signals were guilty. To see that errors can be made when individuals' signals are not commonly known, assume the committee votes under the unanimity rule and suppose further that each juror votes to convict if and only if he or she observes evidence of guilt. Then the outcome is that the defendant is convicted if and only if all three jurors have observed evidence of guilt, exactly the outcome conditional upon all private information becoming public. The problem here, however, is that the low bias juror knows that the only circumstance in which his vote can change the outcome is when the high bias jurors have voted to convict and, therefore, must have observed evidence of guilt. But then the low bias person prefers to vote to convict independent of his private signal and thus has no incentive to vote as supposed. It follows that under unanimity rule the defendent will sometimes be convicted when the low biased juror has observed an innocent signal and the two high biased jurors would prefer to acquit.

A similar argument works for majority rule. In this case, all three jurors have an incentive to vote independently of their signals: if any individual is pivotal when others are assumed to be voting with their signals, there must be an innocent signal, giving the two high bias jurors an incentive to vote for acquital irrespective of their signals; and there must be a guilty signal, giving the low bias juror an incentive to vote to convict irrespective of his signal. Thus the committee can again make errors.

As a consequence of these observations, a central question in the voting literature concerns the circumstances under which communication prior to voting – deliberation – may improve committee decision-making. An important contribution here is Coughlan (2000), who considers the possibility of talking prior to voting in a variant of the Feddersen and Pesendorfer (1998) model in which committee members may or may not have identical preferences. Coughlan models deliberation as a simultaneous cheap talk stage prior to voting, in which each member of the committee privately votes in a non-binding straw poll.[1] The aggregate results of the straw poll are revealed and then the final vote is taken. He shows that if all committee members have identical (ordinal) preferences then deliberation solves the information aggregation problem for all voting rules. However, if committee members' preferences (their biases) are known but not identical (as in our example above) then full information sharing cannot occur for any voting rule.

To appreciate why Coughlan's result holds, first note that the influence of any one committee member's speech during deliberation depends in large part on what others believe prior to hearing that speech. That is, given a voting rule and given certainty regarding the preferences of others, a juror can identify the information possessed by others under which the revelation of his or her own information is pivotal. In our example above with unanimity rule, all jurors know that if the rule requires all jurors to vote for conviction, then the only circumstance in which any one juror's information is relevant is when the two high types have both observed the guilty signal. In this case the low bias juror always prefers conviction and so does better by misreporting his signal and voting to convict in the (cheap talk stage) straw poll. This logic applies more generally and, as a result, when preferences are diverse and commonly known, there can be no guarantee for any voting rule that all individuals tell the truth in deliberation (that is, vote with their signals in Coughlan's straw poll).

In this paper we provide a simple formal example and a general theorem. The example suffices to show that Coughlan's result does not extend beyond unanimity rule to the case in which there is uncertainty about committee members' underlying preferences as well as uncertainty about their private information regarding the guilt or innocence of the defendant; the theorem proves for a wide class of environments that full information sharing under unanimity rule is impossible, even in the presence of preference uncertainty. In general, therefore, uncertainty over preferences can permit truth-telling under majority rule but not under the unanimity rule.[2]

The key intuition is that, in the case of majority rule with two possible sorts of bias (as in the example above), preference uncertainty means that no individual knows whether or not those with his or her bias constitutes a majority (an intuition also explored by Meirowitz, 2005). As before, suppose all others report their signals truthfully in deliberation (Coughlan's straw poll stage). Then there is a chance that lying by, say, a low bias juror convinces a majority of high bias jurors to convict when they might otherwise acquit; but lying also runs the risk of convincing a majority of low bias jurors to convict when no low bias person would prefer this outcome under full information. Jurors must balance out the likelihood of these two events. When information is sufficiently good, jurors put more weight on the event that other jurors have observed information similar to their own and therefore (if deliberation matters) it is more likely that the other jurors share her bias, in which case telling the truth is in the individual's best interest. But with unanimity rule, each agent has a veto in favor of the status quo alternative (acquittal in the case of jury trials). There is therefore no downside risk to lying: the only pivotal event at the communication stage is when all others have information such that, by speaking in favor of conviction, a juror convinces the others to vote to convict when at least one of them would not otherwise do so; so either the individual, on the basis of what he or she learns from others' straw votes, wishes to convict and lying turns out to be in his or her interest or, given what is learned from others, the individual wishes to acquit and can insure this simply by voting to acquit whatever her or she might have said in debate.

Thus we argue that the unanimity rule is uniquely bad with respect to providing committee members with incentives to share relevant information prior to voting. In so doing, we illustrate an important and subtle point about the relationship between deliberation, preference uncertainty and voting rules.[3]

There is an extensive literature on deliberative democracy (see, for example, Cohen, 1989; Fishkin, 1993; and the essays in Bohman and Rehg, 1997). And among the issues addressed therein is a concern with what constitute legitimate, or admissible, deliberative reasons and arguments (e.g. Gutman and Thompson, 1995; Gauss, 1997). There is broad agreement here that a central (perhaps the central) characteristic of a legitimate reason for a collective decision is that it is a reason grounded on some concept of the "common good"; in particular, self-regarding reasons are deemed illegitimate in public debate over social decisions. Insofar as deliberation is intended, among other things, to encourage the pooling of private information salient to making a decision that reflects some notion of the common good, the observation in this paper, that a necessary condition for full information sharing in debate under any voting rule is that private preferences (biases) are not common knowledge, provides positive support for proscribing arguments that reveal such bias. In other words, any requirement that personal biases should be made clear before contributing to public deliberation, is inconsistent with a desire to promote the full revelation of information regarding the relative merits of the feasible collective decisions.

In contrast to its concern with the legitimacy of public arguments, the deliberative democracy literature has paid scant attention to how voting rules might affect the incentives for productive deliberation. And given the focus on deliberation as producing a collective consensus of some sort in favor of a particular decision, this is to be expected: with a consensus on the decision, the choice of voting rule is immaterial since the committee vote will be unanimous for that decision. Indeed, some support for this view is given by Coughlan's result that, with certainty about preferences, full revelation of information in debate is possible under any voting if and only if individuals' ordinal preferences over alternatives are the same. That is, consensus occurs as a result of deliberation if and only if all individuals always share a common bias about what constitutes a good decision prior to the debate. But, as we show in this paper, Coughlan's result does not extend to all voting rules once preferences are themselves subject to some uncertainty; in this case, the choice of voting rule does matter. Furthermore, the claim that, if there is a deliberative consensus, then the voting rule is immaterial, does not imply the converse is true, *viz.* that whatever voting rule the committee uses, a deliberative consensus is possible.

To our knowledge, the issue of how voting rules affect pre-vote deliberation has been addressed only within the formal, game-theoretic, literature on strategic deliberation in multi-person committees.[4] Austen-Smith (1990a,b) considers the role of debate in a spatial model of endogenous agenda-setting under majority rule; Calvert and Johnson (1998) look at the coordinating role of debate in a complete information model of committee decision-making; Meirowitz (2004) also worries about debate and coordination but under incomplete information; Meirowitz (2005) demonstrates (among other things) the importance of preference uncertainty for truthful deliberation in his model, anticipating part of the argument here; and Hafer and Landa (2003) and Aragones, Gilboa, Postlewaite and Schmeidler (2001) develop non-Bayesian models of argument. None of these papers consider unanimity rule. Apart from Coughlan (2000), discussed above, the most closely related contributions to the current paper are Doraszelski, Gerardi and Squintani (2003) and Austen-Smith and Feddersen (2005). Assuming cheap-talk debate and preference uncertainty, Doraszelski, Gerardi and Squintani prove an impossibility result for fully revealing equilibria in a two-person model with deliberation over a fixed binary agenda with unanimity rule, and Austen-Smith and Feddersen establish a similar claim for a particular three-person committee: both results are corollaries of the general theorem proved in this paper. Finally, in a recent and insightful contribution, Gerardi and Yariv (2004) adopt a quite different approach, either to the papers cited here or to this paper. In particular, unlike the focus on full information revelation in what follows, Gerardi and Yariv do not consider any qualitative properties of deliberation *per se*. Framing the issue as one of mechanism design under incomplete information, they instead study the abstract relationships between sets of sequential equilibrium outcomes achievable through unmediated cheap-talk communication in voting games.[5]

**An Example**

We develop the example introduced above more formally, both to illustrate the main ideas of the paper and to justify the claim that there exist environments wherein preference uncertainty can support full information sharing in debate under majority rule. Subsequently, we show our principle result that, even under preference uncertainty, unanimity rule is incapable of supporting full information revelation in debate.

In the example there is a three person jury, or committee, $N = \{1, 2, 3\}$, that has to choose under a given voting rule whether to acquit $(A)$ or to convict $(C)$ a defendant. In the case that the rule is unanimity, assume the *status quo* alternative is to acquit the defendant, $A$. The defendant is either guilty $(G)$ or innocent $(I)$. We assume that $G$ and $I$ occur with equal probability. Each juror $j$ gets a signal $s_j = g$ or $s_j = i$ that is correlated with the true state; specifically

$$\Pr(g|G) = \Pr(i|I) = p$$

with $p \in (.5, 1)$. Let $\mathbf{s} = (s_1, s_2, s_3)$ be the realized profile of signals.

Jurors have preferences that differ only in their bias, $b$, in favor of acquittal. Suppose jurors are either high bias types $(h)$ or low bias types $(l)$. High bias people prefer to convict if and only if all three committee members have observed the guilty signal $(\mathbf{s} = (g, g, g))$, whereas low bias people prefer to convict if and only if at least one committee member observed the guilty signal $(\mathbf{s} \neq (i, i, i))$. Let $u(A, b, \mathbf{s})$ be the payoff from acquittal $A$ given the individual has bias $b$ and the the profile of signals is $\mathbf{s}$, and so forth. To make the point regarding preference uncertainty most simply, assume that,

$$u(A, h, \mathbf{s}) = -1 - u(C, h, \mathbf{s}) = \begin{cases} 0 & if \quad \mathbf{s} \neq (g, g, g) \\ -1 & if \quad \mathbf{s} = (g, g, g) \end{cases}$$

and

$$u(A,l,\mathbf{s}) = -1 - u(C,l,\mathbf{s}) = \begin{cases} 0 & if \quad \mathbf{s} = (i,i,i) \\ -1 & if \quad \mathbf{s} \neq (i,i,i) \end{cases}.$$

Finally, suppose that each voter knows his or her own bias (high or low) but is unsure of the other voters' thresholds of reasonable doubt. Formally, let $q \in (0,1)$ be the *ex ante* probability that a juror has a high bias.

As discussed in the previous section, if there is no communication prior to the jury voting over the agenda, there is an incentive for at least one sort of juror to ignore any of his or her evidence (signal) regarding the guilt or innocence of the defendant when casting a vote. And this can lead to the jury making an error relative to the choice it would make were all of the evidence shared among its members.

Suppose there is an opportunity to share any private information regarding the guilt or innocence of the defendant through a non-binding straw poll prior to voting. Assume that all jurors are expected to reveal their private signal truthfully in the straw poll. If everyone believes that the straw poll contains honest information then behavior at the voting stage becomes simple: those with a high bias vote to convict if everyone voted guilty in the straw poll and vote to acquit otherwise; those with a low bias vote to convict if everyone voted innocent in the straw poll and vote guilty otherwise. Given honest reporting in the straw poll, this voting behavior is the same under either unanimity rule or majority rule. Thus the committee makes no mistakes under either rule relative to the information available to jurors as a group. The problem, therefore, is to determine whether or not honest reporting constitutes equilibrium behavior at the straw poll stage.

Consider unanimity rule. When does it matter what a person reports, i.e., when does voting guilty rather than innocent in the straw poll change the outcome of the subsequent vote? Consider the case of a low bias juror who has observed an innocent signal. It must be that by reporting innocent the defendent is acquitted while by reporting guilty the defendant is convicted. Given that the others have reported their information truthfully in the straw poll, the low bias juror's report in the straw poll does not affect how she votes. Furthermore, under unanimity rule, if her report in the straw poll changes the outcome then it must be because she is voting to convict, for otherwise her vote in the straw poll would not matter. Now, given that she is voting to convict, it follows that all other low bias jurors would vote to convict if she reports honestly and would certainly vote to convict if she reported guilty. So, the only time her report in the straw poll matters is when there is at least one high bias person among the other two and both the other reports are guilty. But in that event the low bias juror is better off reporting guilty in the straw poll. Thus, uncertainty about preferences does not change the incentive to lie in deliberations for those with a low bias under unanimity rule.

Now consider majority rule and suppose as before that everyone honestly reports their information in the pre-vote straw poll. Suppose that a particular individual has a low bias. If this individual has observed a guilty signal then surely telling the truth in the straw poll is the best decision; so suppose she has observed the innocent signal. Under majority rule there are two possible situations in which her report in the straw poll could change the outcome: in situation $S_1$, both of the other jurors have a high bias and have received guilty signals; and in situation $S_2$, both of the other jurors have a low bias and have observed innocent signals. So in situation $S_1$, the low bias individual with an innocent signal prefers conviction to acquittal. But if she tells the truth and reports an innocent signal in $S_1$, the defendant is acquitted. Conversely, in situation $S_2$, where she prefers acquittal to conviction, telling the truth is the only way to secure this outcome. To determine which is the dominant incentive we have to compute the net expected payoff from truth-telling.

Let $\Delta EU(S_j, l, i)$ denote the difference in payoff for a low bias person with an innocent signal from telling the truth and from lying in situation $S_j$. If she tells the truth in $S_1$, the defendant is acquitted and this low bias juror's payoff from this outcome is $-1$ (since at least one other juror has a guilty signal); and if she lies in $S_1$, the defendant is convicted and the low bias juror's payoff from this outcome is 0. Hence, $\Delta EU(S_1, l, i) = -1$. Similarly, calculate $\Delta EU(S_2, l, i) = 1$. Now, the probability of situation $S_1$ conditional on an individual having an innocent signal, is given by:

$$\begin{aligned}\Pr(S_1|i) &= \frac{q^2}{2}p^2(1-p) + \frac{q^2}{2}p(1-p)^2 \\ &= \frac{q^2}{2}p\,(1-p)\,.\end{aligned}$$

And situation $S_2$ occurs with conditional probability:

$$\Pr(S_2|i) = \frac{(1-q)^2}{2}\left[(1-p)^3 + p^3\right].$$

It follows that truth-telling is a best response for the low bias juror who has observed an innocent signal when

$$\begin{aligned}\sum_j \Pr(S_j|i)\Delta EU(S_j,l,i) &\geq 0 \Leftrightarrow \Pr(S_2|i) \geq \Pr(S_1|i)\\ &\Leftrightarrow (1-q)^2\left[(1-p)^3+p^3\right] \geq q^2 p(1-p)\\ &\Leftrightarrow \left(\frac{1-q}{q}\right)^2 \geq \frac{p(1-p)}{(1-p)^3+p^3}.\end{aligned}$$

Therefore, for any nondegenerate probability of another individual being a high type, $q \in (0,1)$, there exists a sufficiently high quality of information, $\hat{p}(q) \in (.5,1)$, such that honest signal revelation is the unique best response for all $p > \hat{p}(q)$ when the individual has a low bias and an innocent signal. (For $q < 1/2$, $\hat{p}(q) = 1/2$ and, for $q \in [.5,1)$, $\hat{p}(q)$ is strictly increasing in $q$.) And, by symmetry, the same is true for a juror with high bias who has observed a guilty signal. Thus there can exist equilibria with full information revelation under majority rule when there is preference uncertainty, but not otherwise.[6]

The example suffices to show that preference uncertainty and deliberation jointly can induce good committee decisions under majority rule. However, the fact that this felicitous result does not hold for unanimity rule in the example does not imply that it could not hold in other cases. We now show that in fact full information sharing in deliberation under unanimity rule is generally unavailable.

**Model and Result**

The committee $N = \{1, 2, ..., n\}$, $n \geq 2$, has to choose an alternative $z \in \{x, y\}$; let $x$ be the *status quo* policy. Each individual $i \in N$ has private information $(b_i, s_i) \in B \times S$, where $b_i$ is a preference parameter, or *bias*, and $s_i$ is a *signal* regarding the alternatives. Assume the sets $B$ and $S$ are finite and common across individuals $i \in N$. Write $B^n \equiv \mathbf{B}$ and $S^n \equiv \mathbf{S}$; a *situation* is any pair $(\mathbf{b}, \mathbf{s}) \in \mathbf{B} \times \mathbf{S}$, where $\mathbf{b} = (b_1, ..., b_n)$, $\mathbf{s} = (s_1, ..., s_n)$. And with a convenient abuse of language, call any profile of individual signals $\mathbf{s} \in \mathbf{S}$ a *state*, as such a profile exhausts all of the relevant information for the collective decision. Let $p(\mathbf{b}, \mathbf{s})$ be the probability that situation $(\mathbf{b}, \mathbf{s}) \in \mathbf{B} \times \mathbf{S}$ obtains.

For any committee member $i \in N$, $i$'s preferences over $\{x, y\}$ depend exclusively on $i$'s own bias $b_i \in B$ and on the state $\mathbf{s} \in \mathbf{S}$: given a bias $b$ and state $\mathbf{s}$, an individual's payoff from a committee decision $z \in \{x, y\}$ is written $u(z, b, \mathbf{s})$. We assume there are no dogmatic or partisan types; that is, for any bias $b \in B$ there is a nonempty subset of states $\mathbf{S}_b \subset \mathbf{S}$ such that $\mathbf{s} \in \mathbf{S}_b$ implies $u(y, b, \mathbf{s}) > u(x, b, \mathbf{s})$ and $\mathbf{s} \notin \mathbf{S}_b$ implies $u(y, b, \mathbf{s}) < u(x, b, \mathbf{s})$: in other words, all individuals' preferences over the two alternatives are subject to change. To avoid trivialities we assume that every situation occurs with positive probability and, because the concern here is with unanimity rule, that there always exist states at which all members prefer alternative $y$. The first two axioms, respectively, formalize these two essentially technical assumptions.

**Axiom 1 (Full Support)** *For all* $(\mathbf{b}, \mathbf{s}) \in \mathbf{B} \times \mathbf{S}$, $p(\mathbf{b}, \mathbf{s}) > 0$.

**Axiom 2 (Consensus)** *For all* $\mathbf{b} = (b_1, ..., b_n) \in \mathbf{B}$, $\mathbf{S}(\mathbf{b}) \equiv \cap_{i \in N} \mathbf{S}_{b_i} \neq \emptyset$.

The final, most substantive, axiom imposes some structure on the set of signals. We first state the axiom formally and then discuss its motivation. Assume that the set of signals $S$ is ordered by a binary relation, $\succ$, such that the following monotonicity condition obtains.

**Axiom 3 (Monotonicity)** *For any* $s, s' \in S$ *such that* $s \succ s'$ *and* $\mathbf{s}_- \in S^{n-1}$, *let* $\mathbf{s} = (\mathbf{s}_-, s) \in \mathbf{S}$ *and* $\mathbf{s}' = (\mathbf{s}_-, s') \in \mathbf{S}$. *Then* $u(y, b, \mathbf{s}) > u(y, b, \mathbf{s}')$ *and* $u(x, b, \mathbf{s}) < u(x, b, \mathbf{s}')$ *for any* $b \in B$.

In words, suppose there is a pair of states that differ only in that some member has observed $s \in S$ in the first state and $s' \in S$; then $s \succ s'$ implies that $s$ is stronger information than $s'$ in favor of $y$ and against $x$. And notice that this axiom also builds in a degree of symmetry: any individual's relative evaluation of the two alternatives is monotone in signals whatever the individual's bias $b$ and irrespective of exactly which committee member receives what signal.

Given that the committee is to choose from a fixed binary agenda, it is fairly natural to interpret signals $s \in S$ as constituting more or less evidence for choosing one or other of the two alternatives. In general, one could imagine that exactly what constitutes "more or less evidence" depends both on an individual's bias and on his or her signal. Although some quantitative differentiation along these lines is admitted under the monotonicity axiom, there can be no qualitative differences of interpretation. In particular, the axiom insists that if one individual considers a particular signal to be better evidence in favor of choosing $x$ over $y$ than some other signal, then all individuals share this relative evaluation, but they may disagree about how much better is the evidence. Thus, the axiom rules out the possibility that two individuals could look at a given piece of evidence, or set of signals, and draw diametrically opposed inferences about the relative merits of the available alternatives. On the one hand, this is clearly a limitation; there is, for example, considerable evidence that individuals tend to see what they hope to see when interpreting evidence (referred to as "confirmatory bias" in pyschology). On the other hand, however, the possibilities for sharing information among a group of individuals are clearly greater when all individuals share a common view of what any particular piece of information might mean. Hence, any negative result on the possibility of informative deliberation that presumes such a common view, is a stronger result than one which assumes at the outset that individuals can interpret evidence and arguments in radically different ways.

The committee chooses an outcome by voting under unanimity rule. That is, $x$ is the outcome unless every member of the committee votes for $y$. Prior to voting we assume there is a deliberation phase in which every member of the committee can simultaneously send a message $m$ to every other member of the committee. The focus here is on deliberation that yields information relevant to the collective choice being shared prior to the voting stage, so there is no loss of generality in associating messages directly with the information they are presumed to report. Therefore we take the set of available messages to any individual to be a set $M$ such that $S \subseteq M$. A *message strategy* for $i \in N$ is a function, $\mu_i : B \times S \to M$. A message profile $\mathbf{m} = (m_1, m_2, ..., m_n) \in M^n \equiv \mathbf{M}$ is a *debate.*

**Definition 1** *A message strategy profile* $\boldsymbol{\mu}$ *is* fully revealing *if, for all* $i \in N$, *for all pairs of distinct signals* $s, s' \in S$, $[\cup_{b \in B} \mu_i(b, s)] \cap [\cup_{b \in B} \mu_i(b, s')] = \emptyset$.

As defined here, fully revealing message strategies may or may not reveal information about individual biases.[7] Because individuals' preferences depend only on the state and on their own bias, if a debate fully reveals the state then additional information about others' biases is decision-irrelevant. Thus the key feature of a fully revealing message strategy is that it provides full information about the speaker's signal. That is, if $\boldsymbol{\mu}$ is fully revealing then, for all individuals $i \in N$ and all bias and signal pairs $(b, s) \in B \times S$, the message $\mu_i(b, s) \in M$ unambiguously reveals that $i$'s private signal is $s$.

**Definition 2** *A committee is* minimally diverse *if and only if there exist* $b, b' \in B$ *such that* $\mathbf{S}_b \neq \mathbf{S}_{b'}$.

In words, a committee is minimally diverse if its membership exhibits preference heterogeneity at least to the extent that there is some pair of individual bias parameters that disagree about the states in which alternative $y$ should be selected. Under the full support assumption, it is possible for all individuals to exhibit the same bias and, therefore, the only committees that are not minimally diverse are committees in which there is *never* any disagreement about when alternative $y$ is the best choice ($\mathbf{S}_b = \mathbf{S}_{b'}$ for all $b, b' \in B$).

A *voting strategy* for member $i \in N$ is a function $\nu_i : B \times S \times \mathbf{M} \to \{x, y\}$ that maps every debate into a voting decision. A *fully revealing debate equilibrium* is a Perfect Bayesian

Equilibrium $(\boldsymbol{\mu}, \boldsymbol{v}) = ((\mu_1, \ldots, \mu_n), (\nu_1, \ldots, \nu_n))$ such that $\boldsymbol{\mu}$ is fully revealing and $\boldsymbol{v}$ is a profile of weakly undominated voting strategies.

**Theorem** *Assume full support, consensus and monotonicity. There exists a fully revealing debate equilibrium under unanimity rule if and only if the committee is not minimally diverse.*

Thus the circumstances under which unanimity rule promotes fully revealing deliberation are confined to those in which it is common knowledge that the committee is homogenous with respect to preferences over alternatives.[8] A proof for this result is in the Appendix (where we also confirm that the theorem extends to the case that the true bias profile $\mathbf{b} \in \mathbf{B}$ is common knowledge). In effect, the formal proof makes precise the intuition sketched in the introduction. The maintained axioms are used principally to prove that, if the committee is minimally diverse, then there must exist a state $\mathbf{s}$ and two bias-types with strictly opposing preferences over $x$ and $y$ at that state, such that $\mathbf{s}$ differs in exactly one component from a different state, say $\mathbf{s}'$, at which all bias types strictly prefer $y$ to $x$. It then follows that an individual with the bias favoring the alternative $y$ over $x$ at this state has an incentive in debate to misrepresent his or her signal under $\mathbf{s}$ and claim instead to have received the signal that distinguishes $\mathbf{s}$ from $\mathbf{s}'$. As a result, it is impossible for all of the necessary incentive compatibility conditions for truth-telling to obtain.

We close this section by recording an easy implication of the theorem; although technically obvious, the corollary is substantively consequential.

Let $q \in \{1, 2, \ldots, n\}$ and recall that a *q-rule* is a voting rule such that if at least $q \geq 1$ committee members vote for $y$ against $x$, then $y$ is the committee decision. Unanimity rule is a $q$-rule with $q = n$. Then noting that the sufficiency argument for the theorem does not depend in any substantive way on the use of unanimity rule, this argument can be applied directly to any $q$-rule to yield the following corollary.

**Corollary** *Assume full support, consensus and monotonicity. If there exists a fully revealing debate equilibrium under unanimity rule then there exists a fully revealing debate equilibrium under all q-rules.*

In other words, because committees that are not minimally diverse unanimously agree on the preferred alternative in every possible situation, such committees can always support fully revealing deliberation whatever voting rule is used to finalize a decision.

## Conclusion

The properties of voting rules have long received attention, most of which has focused on the aggregation of preferences. More recently, the literature has concerned the extent to which various voting procedures efficiently aggregate information. This paper, too, is concerned with information aggregation and voting rules, but rather than ask how different rules aggregate information through the aggregation of votes, it asks how different rules provide incentives for voters to share information prior to taking a vote. Coughlan (2000) finds necessary and sufficient conditions for a fully revealing debate among jurors under any voting rule when preferences are commonly known. In this paper, we show that Coughlan's result does not extend to the case of preference uncertainty for non-unanimous rules: majority rule at least can support full information sharing in debate in some environments when preferences are uncertain; in contrast, unanimity rule quite generally provides strategic incentives for at least some individuals to conceal their private information in debate.

A set of important questions of remain unanswered. For example: What are the comparative properties of voting rules with respect to providing incentives for informative but not fully informative deliberation? What are the welfare properties associated with choosing one rule over another when deliberation is feasible? Such questions must be left to subsequent research.

**Appendix**

For all $\mathbf{b} \in \mathbf{B}$, let $\mathbf{T}^0(\mathbf{b}) \equiv \mathbf{S}(\mathbf{b})$ and, for any $k = 1, 2, \ldots$, recursively define the sets

$$\mathbf{T}^k(\mathbf{b}) = \{\mathbf{s} \notin \cup_{l=1}^{l=k} \mathbf{T}^{k-l}(\mathbf{b}) | \exists s, s' \in S : s' \succ s, \, (\mathbf{s}_-, s) = \mathbf{s}, \, (\mathbf{s}_-, s') = \mathbf{s}' \text{ and } \mathbf{s}' \in \mathbf{T}^{k-1}(\mathbf{b})\}.$$

Thus $\mathbf{T}^1(\mathbf{b})$ is the set of states not in $\mathbf{T}^0(\mathbf{b})$ such that, given the realized bias profile $\mathbf{b}$, changing any one person's information from $s$ to $s'$ results in a state in $\mathbf{T}^0(\mathbf{b}) \equiv \mathbf{S}(\mathbf{b})$; $\mathbf{T}^2(\mathbf{b})$ is the set of states not in $\mathbf{T}^1(\mathbf{b})$ such that changing any one person's information from $s$ to $s'$ results in a state in ; and so on. Informally, the set $\mathbf{T}^k(\mathbf{b})$ is the set of states such that there is a path of $k$ single coordinate changes of information that lead to a state at which $y$ is preferred unanimously. Since $S$ and $N$ are finite it follows that

$$\cup_{k=0,1,\ldots,n} \mathbf{T}^k(\mathbf{b}) = \mathbf{S}.$$

For example, suppose $n = 3$, $S = \{0, 1\}^3$ and $\mathbf{S}(\mathbf{b}) = \{(1, 1, 1)\}$. Then

$$\begin{aligned} T^0(\mathbf{b}) &= \{(1, 1, 1)\} \\ \mathbf{T}^1(\mathbf{b}) &= \{(0, 1, 1), (1, 0, 1), (1, 1, 0)\} \\ \mathbf{T}^2(\mathbf{b}) &= \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \\ \mathbf{T}^3(\mathbf{b}) &= \{(0, 0, 0)\}. \end{aligned}$$

The following property of minimally diverse committees in environments satisfying the three axioms is useful for proving the main theorem. The lemma insures that in minimally diverse committees, there must exist two bias types and a state such that, first, the two bias types have strictly opposing preferences at the state and, second, that the state differs in only one component from another state at which all bias types strictly prefer $y$ to $x$.

**Lemma** *Assume full support, consensus and monotonicity. In a minimally diverse committee there exists a bias profile* $\mathbf{b} = (\mathbf{b}_{-}, b, b') \in \mathbf{B}$ *and a state* $\mathbf{s} \in \mathbf{T}^1(\mathbf{b})$ *such that* $\mathbf{s} \notin \mathbf{S}_b$ *but* $\mathbf{s} \in \mathbf{S}_{b'}$.

**Proof** Let $\mathbf{b} = (\mathbf{b}_{-}, b, b') \in \mathbf{B}$ (where, by an abuse of notation, it is understood that $\mathbf{b}_{-} \in B^{n-2}$); by consensus, $\mathbf{S}(\mathbf{b}) \neq \emptyset$. First assume there is a state $\mathbf{s} \in \mathbf{S}_b \cap \mathbf{T}^{k+1}(\mathbf{b})$. By full support and definition of $\mathbf{T}^k(\mathbf{b})$, there exists a signal $s' \succ s$ such that $(\mathbf{s}_{-}, s') = \mathbf{s}' \in \mathbf{T}^k(\mathbf{b})$; moreover, by monotonicity, $\mathbf{s}' \in \mathbf{S}_b$. Hence, $\mathbf{s} \in \mathbf{S}_b \cap \mathbf{T}^{k+1}(\mathbf{b})$ implies there exists a state $\mathbf{s}' \in \mathbf{S}_b \cap \mathbf{T}^k(\mathbf{b})$. Now suppose $b$ is such that, for any $\mathbf{s} \in \mathbf{T}^k(\mathbf{b})$, $\mathbf{s} \notin \mathbf{S}_b$. Then by the previous argument, there can be no $\mathbf{s} \in \mathbf{T}^{k+1}(\mathbf{b})$ such that $\mathbf{s} \in \mathbf{S}_b$. Hence, $\mathbf{S}_b \cap \mathbf{T}^1(\mathbf{b}) = \emptyset$ implies $\mathbf{S}_b \cap \mathbf{T}^k(\mathbf{b}) = \emptyset$ for all $k > 1$ in which case, because $\cup_{k=0,1,\dots,n}\mathbf{T}^k(\mathbf{b}) = \mathbf{S}$, it must be that $\mathbf{S}_b = \mathbf{S}(\mathbf{b})$. It follows that if, contrary to the lemma, for all $\mathbf{b} \in \mathbf{B}$ there exists no $\mathbf{s} \in \mathbf{T}^1(\mathbf{b})$ and components $b, b'$ of $\mathbf{b}$ such that $\mathbf{s} \notin \mathbf{S}_b$ but $\mathbf{s} \in \mathbf{S}_{b'}$, then $\mathbf{S}_b = \mathbf{S}(\mathbf{b})$ for all components of $\mathbf{b}$, violating minimal diversity. $\square$

**Proof of Theorem** (Necessity) In any fully revealing debate equilibrium, the restriction to weakly undominated voting strategies implies $\nu_i(b, s, \mathbf{m}) = y$ if and only if $(\mathbf{s}_{-i}, s) \in \mathbf{S}_b$, where $\mathbf{s}_{-i} = \mathbf{m}_{-i}$ for every $i \in N$ and $b \in B$. It follows that a member's voting strategy does not depend on the message she sends in debate. Consider the deliberation stage and, by way of contradiction, suppose $\boldsymbol{\mu}$ is fully revealing yet the committee is minimally diverse. Then, given the behavior at the voting stage, fully revealing message strategies constitute an equilibrium if and only if, for every $i \in N$ and every $(b_i, s_i) \in B \times S$, it is the case that

$$EU(m_i = s_i, b_i, s_i) - EU_i(m_i = s', b_i, s_i) \geq 0 \text{ for any } s' \in M \backslash \{s_i\} \tag{1}$$

where $EU(m_i, b_i, s_i) =$

$$\sum_{\mathbf{b}_{-i} \in B^{n-1}} \sum_{\mathbf{s}_{-i} \in S^{n-1}} p(\mathbf{b}_{-i}, \mathbf{s}_{-i} | b_i, s_i) \left[ \Pr(x | \mathbf{b}, \mathbf{s}, m_i) u(x, b_i, \mathbf{s}) + \Pr(y | \mathbf{b}, \mathbf{s}, m_i) u(y, b_i, \mathbf{s}) \right]$$

and $\Pr(z|\mathbf{b}, \mathbf{s}, m_i)$ is the probability that $z \in \{x, y\}$ is the committee decision given bias profile $\mathbf{b} = (\mathbf{b}_{-i}, b_i)$, state $\mathbf{s} = (\mathbf{s}_{-i}, s_i)$ and debate $(\mathbf{m}_{-i}, m_i) = (\mathbf{s}_{-i}, m_i)$. Fix $i \in N$ and let $(b_i, s_i) = (b, s)$; for any $s' \in M \backslash \{s\}$, define the function

$$\varphi_{(b,s)}(s, s'; \mathbf{b}_{-i}, \mathbf{s}_{-i}) \equiv \left[ \Pr(x|\mathbf{b}, \mathbf{s}, s) - \Pr(x|\mathbf{b}, \mathbf{s}, s') \right] \left[ u(x, b, \mathbf{s}) - u(y, b, \mathbf{s}) \right]$$

with $\mathbf{b} = (\mathbf{b}_{-i}, b)$ and $\mathbf{s} = (\mathbf{s}_{-i}, s)$. Then we can rewrite (1) equivalently as requiring that for all $(b, s) \in B \times S$ and all $s' \in M \backslash \{s\}$,

$$\sum_{\mathbf{b}_{-i} \in B^{n-1}} \sum_{\mathbf{s}_{-i} \in S^{n-1}} p(\mathbf{b}_{-i}, \mathbf{s}_{-i} | b, s) \varphi_{(b,s)}(s, s'; \mathbf{b}_{-i}, \mathbf{s}_{-i}) \geq 0. \tag{2}$$

By assumption, $\boldsymbol{\mu}_{-i}$ is fully revealing of all others' signals and, by the preceding argument on $\nu_i$, for all messages $m_i \in M$ and all bias profiles $(\mathbf{b}_{-i}, b) \in \mathbf{B}$, $(\mathbf{s}_{-i}, s) \in \mathbf{S} \backslash \mathbf{S}_b$ implies $\Pr(x|(\mathbf{b}_{-i}, b), (\mathbf{s}_{-i}, s), m_i) = 1$. Similarly, for any state $(\mathbf{s}_{-i}, s) \in \mathbf{S} \backslash (\mathbf{S}(\mathbf{b}) \cup \mathbf{T}^1(\mathbf{b}))$ it must be that $\Pr(x|(\mathbf{b}_{-i}, b), (\mathbf{s}_{-i}, s), m_i) = 1$. Given $(b_i, s_i) = (b, s)$, therefore, for all $s' \in M \backslash \{s\}$ and all

$\mathbf{b}_{-i} \in B^{n-1}$,

$$(\mathbf{s}_{-i}, s) \in \mathbf{S} \backslash \left[\mathbf{S}(\mathbf{b}) \cup \mathbf{T}^1(\mathbf{b}) \cup \mathbf{S}_b\right] \Rightarrow \varphi_{(b,s)}(s, s'; \mathbf{b}_{-i}, \mathbf{s}_{-i}) = 0. \tag{3}$$

The preceding argument implies that an individual $i$ with bias $b$ can change the outcome by switching from message $s$ to some $s' \neq s$ only in situations $(\mathbf{b}, \mathbf{s})$ such that $(\mathbf{s}_{-i}, s) \in \mathbf{S}(\mathbf{b}) \cup \left(\mathbf{T}^1(\mathbf{b}) \cap \mathbf{S}_b\right)$. For all $\mathbf{b} \in \mathbf{B}$, define

$$x_i(\mathbf{b}, s, s') = \{(\mathbf{s}_{-i}, s) \in \mathbf{S}(\mathbf{b}) | (\mathbf{s}_{-i}, s') \notin \mathbf{S}(\mathbf{b})\}$$

to be the set of states such that if an individual $i$ who is supposed to report $s$ instead reports $s'$ then, conditional on $\mathbf{b}$, the outcome changes from $y$ to $x$. Similarly, define

$$y_i(\mathbf{b}, s, s') = \{(\mathbf{s}_{-i}, s) \in \left(\mathbf{T}^1(\mathbf{b}) \cap \mathbf{S}_b\right) | (\mathbf{s}_{-i}, s') \in \mathbf{S}(\mathbf{b})\}$$

to be the set of states in which $i$ prefers $y$ and, if $i$ is supposed to report $s$ but instead reports $s'$ at $\mathbf{b}$, the outcome changes from $x$ to $y$. Note that, by monotonicity, if $y_i(\mathbf{b}, s, s') \neq \emptyset$ for some $\mathbf{b} \in \mathbf{B}$, then $x_i(\mathbf{b}, s, s') = \emptyset$ for all $\mathbf{b} \in \mathbf{B}$ and, if $x_i(\mathbf{b}, s, s') \neq \emptyset$ for some $\mathbf{b} \in \mathbf{B}$, then $y_i(\mathbf{b}, s, s') = \emptyset$ for all $\mathbf{b} \in \mathbf{B}$. That is, $y_i(\mathbf{b}, s, s') \neq \emptyset$ for some $\mathbf{b} \in \mathbf{B}$ implies that $s'$ is stronger evidence for $y$ than $s$, whereas $x_i(\mathbf{b}, s, s') \neq \emptyset$ for some $\mathbf{b} \in \mathbf{B}$ implies $s'$ is weaker evidence for $y$ than $s$. By monotonicity both statements cannot be true. For any $\mathbf{b} \in \mathbf{B}$ and $s, s' \in S$, let

$$Z_i^-(\mathbf{b}, s, s') \equiv \{\mathbf{s}_{-i} \in S^{n-1} | (\mathbf{s}_{-i}, s) \in \left[y_i(\mathbf{b}, s, s') \cup x_i(\mathbf{b}, s, s')\right]\}.$$

Collecting terms and using (3), we can rewrite the incentive compatibility constraint (2) as requiring, for all $i \in N$, $(b,s) \in B \times S$ and $s' \in M \backslash \{s\}$,

$$\sum_{\mathbf{b}_{-i} \in \mathbf{B}^-} \sum_{\mathbf{s}_{-i} \in Z_i^-(\mathbf{b},s,s')} p(\mathbf{b}_{-i}, \mathbf{s}_{-i}, |b,s) \varphi_{(b,s)}(s,s';\mathbf{b}_{-i},\mathbf{s}_{-i}) \geq 0. \tag{4}$$

By the Lemma and full support, minimal diversity implies there is a $(\mathbf{b}_{-i}, b) \in \mathbf{B}$ and a pair of signals $s, s' \in S$ such that $y_i((\mathbf{b}_{-i}, b), s, s') \neq \emptyset$ and $x_i((\mathbf{b}_{-i}, b), s, s') = \emptyset$. By definition, $(\mathbf{s}_{-i}, s) \in y_i((\mathbf{b}_{-i}, b), s, s')$ implies $u(x, b, (\mathbf{s}_{-i}, s)) < u(y, b, (\mathbf{s}_{-i}, s))$ and $\Pr(x|(\mathbf{b}_{-i}, b), \mathbf{s}, s) - \Pr(x|(\mathbf{b}_{-i}, b), \mathbf{s}, s') = 1$. Hence, for all $(\mathbf{b}_{-i}, b) \in \mathbf{B}$,

$$\mathbf{s}_{-i} \in Z_i^-(\mathbf{b}, s, s') \Rightarrow \varphi_{(b,s)}(s, s'; \mathbf{b}_{-i}, \mathbf{s}_{-i}) < 0.$$

But then the incentive compatibility conditions are surely violated, contradicting the existence of a fully revealing debate equilibrium in any minimally diverse committee. This proves necessity.

(Sufficiency) Assume the committee is not minimally diverse. Then for all $b \in B$ and all $\mathbf{b} = (\mathbf{b}_-, b) \in \mathbf{B}$, $\mathbf{S}_b = \mathbf{S}(\mathbf{b})$. In this case there is no $\mathbf{b} \in \mathbf{B}$ and pair of signals $s, s' \in S$ such that $y_i(\mathbf{b}, s, s') \neq \emptyset$ for any $i \in N$. Since incentive compatibility is assured for any $i \in N$, $b \in B$ and pair of signals $s, s' \in S$ such that $x_i(\mathbf{b}, s, s') \neq \emptyset$ and $y_i(\mathbf{b}, s, s') = \emptyset$, full revelation is an equilibrium strategy. This completes the proof. $\square$

Finally, to see that the theorem goes through under complete information regarding individuals' biases, fix a bias profile $\mathbf{b} = (b_1, \ldots, b_n)$, suppose $\mathbf{b}$ is common knowledge and let $\mathbf{B} = \{\mathbf{b}\}$. Then the definitions and the argument directly apply on replacing references to "biases $b, b' \in B$" with references to "individuals $i, j \in N$ with biases $b_i, b_j$", and so on.

**References**


Aragones, Enriqueta, Itzhak Gilboa, Andrew W. Postlewaite and David Schmeidler. 2001. "Rhetoric and Analogies." University of Pennsylvania. Typescript.

Austen-Smith, David. 1990a. "Information transmission in debate." *American Journal of Political Science* 34 (February): 124-52.

Austen-Smith, David. 1990b. "Credible debate equilibria." *Social Choice and Welfare* 7 (January): 75-93.

Austen-Smith, David. 1993. "Interested experts and policy advice: multiple referrals under open rule." *Games and Economic Behavior* 5 (January): 3-43.

Austen-Smith, David and Jeffrey S. Banks. 1996. "Information aggregation, rationality, and the Condorcet Jury Theorem." *American Political Science Review* 90 (March): 34-45.

Austen-Smith, David and Timothy J. Feddersen. 2005. "Deliberation and voting rules." In *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks,* eds. David Austen-Smith and John Duggan. Heidelberg:Springer.

Bohman, James and William Rehg, eds. 1997. *Deliberative Democracy: Essays on Reason and Politics.* Cambridge MA:MIT Press.

Calvert, Randall L. and James Johnson. 1998. "Rational argument, political argument and democratic deliberation." Presented at the Annual Meeting of the American Political Science Association.

Cohen, Joshua. 1989. "Deliberation and Democratic Legitimacy." In *The Good Polity*, eds. Alan Hamlin and Philip Pettit. Oxford:Basil Blackwell.

Coughlan, Peter J. 2000. "In defense of unanimous jury verdicts: mistrials, communication and strategic voting." *American Political Science Review* 94 (June): 375-93.

Doraszelski, Ulrich, Dino Gerardi and Francesco Squintani. 2003. "Communication and voting with double-sided information." *Contributions to Theoretical Economics* 3(1), Article 6.

Feddersen, Timothy J. and Wolfgang Pesendorfer. 1998. "Convicting the innoncent: the inferiority of unanimous jury verdicts." *American Political Science Review* 92 (March): 23-35.

Fishkin, James. 1993. *Democracy and Deliberation: New Directions for Democratic Reform.* New Haven: Yale University Press.

Gauss, Gerald F. 1997. "Reason, justification and consensus: why democracy can't have it all." In *Deliberative Democracy*, eds. James Bohman and William Rehg. Cambridge MA:MIT Press.

Gerardi, Dino and Leeat Yariv. 2004. "Putting your ballot where your mouth is: an analysis of collective choice with communication." Yale University. Typescript.

Gutmann, Amy and Dennis Thompson. 1995. "Moral disagreement in a democracy." S*ocial Philosophy and Policy* 12 (Winter): 87-110.

Hafer, Catherine and Dimitri Landa. 2003. "Deliberation as self-discovery." New York University. Typescript.

Lipman, Barton L. and Duane J. Seppi. 1995. "Robust inference in communication games with partial provability." *Journal of Economic Theory* 66 (August): 370-405.

McLennan, Andrew. 1998. "Consequences of the Condorcet Jury Theorem for beneficial information aggregation by rational players." *American Political Science Review.* 92 (June): 413-18.

Meirowitz, Adam. 2004. "In defense of exclusionary deliberation: communication and voting with private beliefs and values." Princeton University. Typescript.

Meirowitz, Adam. 2005. "Deliberative democracy or market democracy: designing institutions to aggregate preferences and information." Princeton University. Typescript.

Ottaviani, Marco and Peter Sorensen. 2001. "Information aggregation in debate: who should speak first?" *Journal of Public Economics* 81 (September): 393-422.

Rawls, John. 1993. *Political Liberalism.* New York: Columbia University Press.

Walker, John and Desmond Lane. 1994. "The jury system."
http://libertyvictoria.org.au/topics.asp?iss=2&topic=49 (February 25, 1994).

**Endnotes**

* Earlier versions of this paper were titled "The Inferiority of Deliberation under Unanimity Rule". We appreciate the comments of several anonymous referees and the financial support of the NSF under grant SES-0505818.

1. Communication is called “cheap talk” if the particular message or speech made is unverifiable by others and imposes no direct costs on the speaker. Any influence such speech might have is exclusively through its influence on the decision-relevant beliefs of others. Although cheap talk is not the only model of communication one might use for thinking about deliberation (for instance, not all claims in debate are unverifiable), it is a natural benchmark case to consider. If there is credible communication under cheap talk then there is credible communication when, say, lying is costly, however small the cost; but credible communication when lying has a cost does not imply such communication is possible when lying is a little less costly.

2. This stands in stark contrast to a common intuition that requiring a unanimous vote provides the strongest incentives for those involved to share their opinions and decision-relevant information most fully. For example, writing in support of a legal decision, the South Australian High Court (1993) argued that “The necessity of a consensus of all jurors which flows from the requirement of unanimity, promotes deliberation and provides some insurance that the opinions of each of the jurors will be heard and discussed” (quoted in Walker and Lane, 1994:2)

3. To the extent that unanimity rule is empirically uninteresting, our results on the merits of the rule are of at most academic concern. Yet unanimity rule is used in several important decision-making environments. For instance, the European Council (comprising the heads of state) and the Council of Ministers (consisting of national ministers) within the European Union use unanimity rule for decisions on, among a variety of other issues, tax harmonization, expansion of the union and constitutional reform; a unanimous vote is required for conviction in jury trials for particular sorts of criminal offence; unanimous consent is typically required for trade agreements in the WTO; and US law requires that a proposal for debt restructuring of a company not in bankruptcy can be accepted only by a unanimous vote of the creditors.

4. There is a complementary formal literature concerned with $n$-person debate aimed at influencing an uninformed monopolistic decision-maker. Examples include Ottaviani and Sorensen (2001), Lipman and Seppi (1995), and Austen-Smith (1993). Key differences between the papers cited in the text and those falling within this complementary class are that, in the latter, the set of individuals deliberating does not coincide with the set of individuals responsible for making a decision and there is no explicit concern with strategic voting.

5. It is worth noting that Gerardi's and Yariv's (2004) main result provides a somewhat different argument than Coughlan's (2000) theorem to support the view that voting rules are largely unimportant for deliberative democracy. They show, first, that all non-unanimous $q$ rules (that is, rules requiring at least $q \geq 1$ committee members to vote for an alternative for that alternative to be adopted) are equivalent in that the sets of sequential equilibrium

outcomes induced by use of any $q$ rules with $q \neq 1, n$ are identical once voting is preceded by deliberation and, second, that those outcomes induced by unanimity rule are a subset of those of induced by any non-unanimous rule. Thus any non-unanimous $q$ rule can be chosen without affecting what is possible in a pre-vote debate. However, Gerardi and Yariv's result exploits the fact that the rules governing debate are unconstrained and that (at least on the equilibrium path) all voters voting unanimously is consistent with sequential rationality for non-unanimous $q$ rules. Such consistency is obtained in their analysis either by admitting weakly dominated strategies or, with a mild domain restriction, precluding dominated strategies defined in terms of (*ex ante* or *interim*) expectations formed prior to any individual hearing any debate. Thus a strategy pair (message and voting) can be undominated in expectation but, at the same time, the specified voting strategy can be dominated conditional on the realized debate. Requiring instead that voting must be undominated conditional on the realized debate opens up the possibility that voting rules affect the incentives for deliberation (see Austen-Smith and Feddersen, 2005).

6. This conclusion does not depend on the assumption that utility values can take at most two values. A similar result goes through when preferences are assumed strictly monotonic in the number of guilty signals, although in this case there is a more demanding lower bound on the quality of information, $p$.

7. At first glance, the definition here might seem unnecessarily awkward with a simpler version being to require only that if, for all $i$, all $b$ and any $s \neq s'$, $\mu_i(b, s) \neq \mu_i(b, s')$. But this

does not work, as it admits the possibility that $\mu_i(b', s) = \mu_i(b, s')$ for some $b' \neq b$, in which case $i$'s signal regarding the state is not revealed.

8. Note that the example of the previous section satisfies the premises of the theorem and involves a minimally diverse committee. So the lack of a fully revealing debate equilibrium observed there also follows directly from the theorem.