Notes on Bias Uncertainty and Communication in Committees[1]

David Austen-Smith
MEDS, Kellogg Graduate School of Management
Northwestern University
Evanston, IL 60208

Tim Feddersen
MEDS, Kellogg Graduate School of Management
Northwestern University
Evanston, IL 60208

March 2007

[1]Prepared for presentation to the Roy Seminar, Paris, March 2007. Part of these Notes rely on Austen-Smith and Feddersen (2006).

1 Introduction

It is well understood that when committee members vote under incomplete information, the resulting committee decision need not reflect the decision that would have been made under fully shared information (eg Austen-Smith and Banks 1996; Feddersen and Pesendorfer 1998). But people in committees often talk before voting and so have an opportunity to share decision-relevant information. Over the past few years there has been a growing strategic game-theoretic literature concerned to understand better what implications such communication might have for the character and quality of collective decisions under incomplete information.[1] One issue here concerns how different voting rules for reaching a final collective choice influences the information that might be shared in any prior debate. Gerardi and Yariv (2004) suggest that the voting rule (essentially) is irrelevant, but their argument hinges on using dominated (debate-conditional) voting strategies. Assuming undominated voting changes the picture (Austen-Smith and Feddersen 2005). And in this case, it turns out that whether or not individuals' biases, that is, differences in full information preferences over available collective choices, are common knowledge plays a critical role.[2] In these Notes, we review some of the positive issues arising from bias uncertainty and suggest some intuitions regarding the welfare implications of biases being private information.

2 An example

Consider a committee of three people, $i = 1, 2, 3$, that has to choose between a fixed pair of alternatives $\{X, Y\}$. Each individual has private information $(b_i, s_i) \in \{x, y\} \times \{x, y\}$, where b_i is a preference parameter, or *bias*, and s_i is a noisy but informative *signal* regarding the

[1] For example, see: Austen-Smith 1990; Calvert and Johnson 1998; Coughlan 2000; Doraszelski et al 2003; Gerardi and Yariv 2004; Meirowitz 2006, 2007; Hafer and Landa 2007; Austen-Smith and Feddersen 2005, 2006; Caillaud and Tirole 2006.

[2] An intuition noted too in Meirowitz (2006).

alternatives. Let $(\mathbf{b}, \mathbf{s}) = ((b_1, b_2, b_3), (s_1, s_2, s_3))$ denote a profile of realized biases and signals. For this example, assume signals are uncorrelated with biases. Let $\{A_1, A_2\}$ be states such that, for all i,

$$\Pr[s_i = x|A_1] = \Pr[s_i = y|A_2] = p \in (1/2, 1)$$

with the common prior probability on A_1 being $1/2$; and, for all i, assume $\Pr[b_i = x] = r \in (0, 1)$.

Write $u(Z, b, \mathbf{s})$ for an individual's payoff from the collective choice $Z \in \{X, Y\}$, given that individual's bias is $b \in \{x, y\}$ and the profile of signals is $\mathbf{s} \in \{x, y\}^3$. Assume

$$\begin{aligned} u(X, x, \mathbf{s}) &= \begin{cases} 0 \quad if \quad \mathbf{s} \in \{\,(y, y, y)\} \\ 1 \qquad\qquad otherwise \end{cases} \\ u(Y, x, \mathbf{s}) &= 1 - u(X, x, \mathbf{s}) \\ u(X, y, \mathbf{s}) &= \begin{cases} 1 \quad if \quad \mathbf{s} \in \{\,(x, x, x)\} \\ 0 \qquad\qquad otherwise \end{cases} \\ u(Y, y, \mathbf{s}) &= 1 - u(X, y, \mathbf{s}) \end{aligned}$$

Suppose first that the committee's choice is determined by a majority vote. In the absence of any pre-vote communication, there is no equilibrium in which all individuals vote informatively, that is, with their signals (eg Austen-Smith and Banks, 1996). As a result, the collective choice may favour a minority ex post. More formally, let $v : \{x, y\} \times \{x, y\} \to [0, 1]$ be a vote strategy, where $v(b, s)$ is the probability an individual with bias b and signal s votes for X (we assume strategies are anonymous throughout). Then under majority rule, there is no (Bayesian) equilibrium in undominated strategies in which $v(b, x) = 1 - v(b, y) = 1$ for all b. To see this suppose not and, without loss of generality, consider an individual i with bias $b_i = y$. Given the other committee members are voting with their signals, i is pivotal only in the event that one individual has a signal x and the other has a signal y; but then i's unique best response is to vote surely for Y irrespective of his signal. There is, however, a mixed strategy equilibrium when q is not too extreme: in this equilibrium, $v(x, x) = 1 - v(y, y) = 1$, $v(x, y) \in (0, 1)$ and

$v(y,x) \in (0,1)$. And clearly, the final decision may not reflect the full information majority preference.

Now suppose that the committee can deliberate before voting; specifically, suppose that voting follows one round of simultaneous cheap talk signaling. Note that in this context, and in contrast to n-person sender-receiver games (eg Ottaviani and Sorenson 2001; Glazer and Rubinstein 2001), every committee member is both a "sender" and a "receiver" and no one individual has the right to dictate the final decision. Let $\sigma_i : \{x,y\} \times \{x,y\} \to M$ be a message strategy for i, where M is in general an arbitrarily large set of messages. For current purposes, it suffices to take $M = \Delta\{x,y\}$ with the interpretation that $\sigma_i(b,s)$ is the probability an individual with bias b and signal s declares "$s = x$".[3] With talk, individuals' voting behaviour depend on the realized list of messages from debate: with an abuse of notation, write $v_i : \{x,y\} \times \{x,y\} \times \{x,y\}^3 \to \Delta\{X,Y\}$ be a vote strategy for i, where $v_i(b,s,\mathbf{m})$ is the probability an individual with bias b and signal s who hears debate $\mathbf{m} = (m_1, m_2, m_3) \in \{x,y\}^3$ votes for X. A perfect Bayesian equilibrium in undominated strategies for this setting is a profile of strategies $(\boldsymbol{\sigma}, \mathbf{v})$ such that, for every triple $(b,s,\mathbf{m})$ and all i: (1) $v_i(b,s,\mathbf{m})$ is an undominated best response to $(\boldsymbol{\sigma}, \mathbf{v}_{-i})$; (2) $\sigma_i(b,s)$ is an undominated best response voting strategy to $(\boldsymbol{\sigma}_{-i}, \mathbf{v})$; and (3) beliefs are consistent with the strategy profile and Bayes rule where defined.

Assume all individuals truthfully reveal their signals: $\sigma_i(b,x) = 1 - \sigma_i(b,y) = 1$ for all i and biases b. Then there is fully shared information at the voting stage and the unique undominated voting equilibrium conditional on these messages is for each individual i to vote 'sincerely' for their most preferred alternative as defined by $(b_i, (s_1, s_2, s_3))$. With this fact in mind, consider whether in fact truthful information sharing in debate is incentive compatible. Say that an individual is *message pivotal* in debate if the final colllective choice is sensitive to which signal the individual reveals he or she has observed. Then it is immediate that any

[3]Later, in a simpler version of this example, we explicitly consider allowing biases to be revealed in debate. Dealing with this possibility here is not essential.

x-biased individual with an x signal, or y-biased individual with y signal, has a dominant strategy to reveal their signal truthfully. So, without loss of generality, consider an x-biased individual i with a y signal: $(b_i, s_i) = (x, y)$. Then there are essentially two events in which i is message pivotal:

(1) both of the other committee members are x-biased and both have observed a y signal, which occurs with (conditional) probability

$$\frac{1}{2}r^2\left(p^3 + (1-p)^3\right);$$

(2) both of the other committee members are y-biased and both have observed an x signal, which occurs with (conditional) probability

$$\frac{1}{2}(1-r)^2\left((1-p)^2p + (1-p)p^2\right) = \frac{1}{2}(1-r)^2p(1-p).$$

In event (1), individual i strictly prefers to tell the truth while, in event (2), i strictly prefers to dissemble. Making the requisite calculations, telling the truth in debate conditional on all others fully revealing their signals is incentive compatible for an x-biased individual with a y signal if and only if

$$\left(\frac{r}{1-r}\right)^2 \geq \frac{p(1-p)}{(1-p)^3 + p^3}.$$

Similarly, incentive compatibility of truthtelling for a y-biased individual with an x signal is insured if

$$\left(\frac{1-r}{r}\right)^2 \geq \frac{p(1-p)}{(1-p)^3 + p^3}.$$

For every $r \in (0,1)$, therefore, there exists some $p < 1$ for which there exists an equilibrium in which all individuals reveal their signals truthfully in debate and subsequently vote under full information to yield the ex post first best de facto majority preferred outcome.

It is important to note that the left side of these inequalities describes the relative likelihood of the speaker's bias-type being the majority bias in the committee. Moreover, if $r \in \{0,1\}$ then truthtelling in debate cannot be an equilibrium for any $p \in (1/2, 1)$: bias uncertainty is essential for fully revealing debate. In particular, with bias uncertainty there are multiple message pivotal events under majority rule; whether or not any individual tells the truth in debate thus depends on the relative likelihood of those events in which truthtelling is a best response; that is, on those events in which the individual is a member of the full information winning coalition. And because message pivotal events are defined both in terms of a particular pairs of bias and signal profiles, the fact that signals are informative, coupled with the event of being message pivotal, provides the individual with information regarding the realized bias profile, even when biases and signals are statistically independent.

Now suppose committee decision making is by unanimity rule with status quo X: that is, X is the outcome unless all committee members vote for Y. Without debate, individuals condition their vote in equilibrium on being vote pivotal, which occurs only if both of the other committee members are voting for the alternative, Y. Given all others are voting with their signal, a y-biased individual's best response is therefore to vote for Y irrespective of her signal. Hence, as with majority rule, the ex post committee choice may not be the decision adopted under full information. Suppose there is a debate stage in which all individuals are supposed to report their signals truthfully and consider a y-biased individual with an x signal. Under unanimity rule, any individual's best voting decision is independent of her own message in debate. Thus the only message pivotal event is when at least one of the other two individuals is x-biased and both have y signals (any b-biased individual with an $s = b$ signal surely votes with his signal). But then the unique best response message is to lie and report "y": bias uncertainty is insufficient to promote fully revealing debate under unanimity rule. In other words, there can only be a single message pivotal event under unanimity rule and this event induces dissembling as a best response for at least some types of committee member.

In sum, the example suffices to show that bias uncertainty can promote full information sharing in debate under majority rule, but does not demonstrate that the same is false under unanimity rule. Furthermore, while the presence of bias uncertainty can affect the extent of information sharing in debate, this in itself does not say anything about the welfare properties of committee decision making under bias uncertainty. In the next section we address the first issue regarding bias uncertainty and unanimity rule quite generally and, in the subsequent section, use a considerably less general model to consider some welfare implications of bias uncertainty.

3 Model and theorems

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 $(\mathbf{b}, \mathbf{s}) \in \mathbf{B}l$ 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.

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

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.

Monotonicity *For any $s, s' \in S$ such that $s \succ s'$ and $\mathbf{s}_- \in S^{n-1}$, let $\mathbf{s} = (\mathbf{s}_-, s)$ and $\mathbf{s}' = (\mathbf{s}_-, s') \in \mathbf{S}$. Then $u(Y, b, s) > u(Y, b, s')$ and $u(X, b, s) < u(X, b, s')$ 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. Thus, 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. Although a limitation, the axiom is conservative for proving a negative result, as the possibilities for information sharing are clearly greater when all individuals share a common view of what any particular piece of information might mean.

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, $\sigma_i : B \times S \to M$. 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 $\boldsymbol{\sigma}$ is* fully revealing *if, for all $i \in N$, for all pairs of distinct signals $s, s' \in S$, $[\cup_{b \in B} \sigma_i(b, s)] \cap [\cup_{b \in B} \sigma_i(b, s')] = \emptyset$.*

As defined here, fully revealing message strategies may or may not reveal information about individual biases.[4] 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 σ is fully revealing then, for all individuals $i \in N$ and all bias and signal pairs $(b, s) \in B \times S$, the message $\sigma_i(b, s) \in M$ 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* $\mathbf{S}_b \neq \mathbf{S}_{b'}$.

Note that 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 $v_i : B \times S \times \mathbf{M} \to \{x, y\}$ that maps every debate into a voting decision. A *fully revealing debate equilibrium* is a Perfect Bayesian Equilibrium $(\sigma, \mathbf{v}) = ((\sigma_1, \ldots, \sigma_n), (v_1, \ldots, v_n))$ such that σ is fully revealing and $\mathbf{v}$ is a profile of weakly undominated voting strategies.

Before presenting the result on unanimity rule, it is useful to recall a (slight generalization of a) result due to Coughlan (2000). For any $q \in \{1, 2, \ldots, n\}$, 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 $q = n$.

Theorem 1 (Coughlan 2000) *Assume the realized bias profile is common knowledge and that every state* $\mathbf{s} \in \mathbf{S}$ *occurs with positive probability. Assume consensus and monotonicity. Then,*

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

for all q–rules, $n/2 < q < n$, there exists a FRDE if and only if the committee is not minimally diverse.

Therefore, bias uncertainty is typically necessary for full information sharing in debate under any rule shy of unanimity. And while the example showed that bias uncertainty can support a FRDE, this is by no means guaranteed. That bias uncertainty can never help with unanimity, however, is the content of the next result.

Theorem 2 *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, whether or not there is bias uncertainty. 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).

4 Welfare

An important problem is to identify the optimal rule for committee choice under incomplete information with debate.[5] This problem is particularly challenging if we insist that individuals' voting strategies are undominated conditional on the realized debate, a requirement that opens up the possibility, as illustrated by preceding theorems, that voting rules affect the incentives for deliberation.[6] A more limited welfare question concerns the extent to which the de facto

[5]Chwe (2006) derives the optimal voting rule when there is no debate. Interestingly, this rule turns out to be nonmonotonic in votes and is thus not a q-rule.

[6]Gerardi and Yariv (2004) show, first, that all non-unanimous q rules 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 (not necessarily proper)

committee decision under a given voting rule (typically, a q-rule) with incomplete information and debate reflects the committee's decision conditional on all private information being common knowledge at the time of the vote (eg McLennan 1998; Federsen and Pesendorfer 1997). Equivalently, given the rule, the goal is to maximize the aggregate expected payoff of those in the full information winning coalition. And the intuition suggested by the results above is that, at least for nonunanimous q-rules, bias uncertainty can improve the welfare of the full information winning coalition by facilitating more information sharing in debate: when individuals are not sure ex ante whether they are members of the full information winning coalition, then they have an incentive to reveal private information in debate that is absent when they are confident they are not members of this coalition.

Unfortunately, even restricting attention to q-rules, addressing the latter welfare issue directly is complicated by having to describe the full equilibrium set to any more or less general incomplete information (debate and voting) game induced by the rule (see for example Austen-Smith and Feddersen (2005) on this issue). At this stage, no such characterization is available.[7] So rather than attempt a general result here, we instead illustrate the intuition above with a simpler variant of the opening example.

In the variant of the example, we assume two individuals in a committee in which one fixed individual has the right to make the committee decision. Individuals' preferences and

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 do not require voting strategies to be undominated conditional on the realized debate. Instead, their 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.

[7] Likewise, the optimal mechanism here is as yet unavailable in general. Chwe's (2006) result for majority voting with incomplete information suggests such a mechanism will be complex.

the informational assumptions on signals $s \in \{x, y\}$ are exactly as specified above. However, we perturb the assumptions on the distribution of biases: for each $i = 1, 2$, assume that while the prior probability that $b_i = x$ is 1/2, realized bias types are correlated; that is, for distinct $i, j \in \{1, 2\}$

$$\Pr[b_i = b_j | b_i] = \rho \in [0, 1].$$

If $\rho = 0$ they have conflicting preferences with probability 1 whereas if $\rho = 1$ both have identical preferences. The idea here, is that the decision-maker is a proxy for the pivotal voter defining the winning coalition under a q-rule in a larger committee, and the other individual (the advisor) is unsure whether she too is a member of that winning coalition. Hereafter, let individual $i = 1$ be the advisor and individual $i = 2$ be the decision-maker.

We consider two communication protocols. In the first protocol there is a single debate stage exactly as hitherto, in which both individuals make cheap talk speeches regarding their signals, following which the decision-maker chooses X or Y. For the second protocol, we add an earlier debate stage in which both players can declare their biases following which they both send messages regarding their signals and, finally, the decision-maker makes a decision. Equilibria are again perfect Bayesian with undominated strategies. Note that, because an individual's payoffs, given her bias, depend exclusively on the final decision and the profile of realized signals $\mathbf{s}$, there can be no value (in in signaling both bias and signal simultaneously or of having a signal debate prior to making any statements regarding biases. Thus the role of permitting bias revelation early in any committee discussion is to improve the possibility of coordination between committee members.

Begin with assuming only a single debate stage in which individuals make speeches about their signals. Strategies are as before: $\sigma_i : \{x, y\} \times \{x, y\} \to \Delta\{x, y\}$ is a debate strategy with $\sigma_i(b, s)$ being the probability that individual i with bias b and signal s declares "$s = x$"; and $v : \{x, y\} \times \{x, y\} \times \{x, y\}^2 \to \Delta\{X, Y\}$ is the decision-maker's decision (vote) strategy with $v(b, s, \mathbf{m})$ being the probability that the decision-maker (individual 2) with bias b and signal

s who hears debate $\mathbf{m} = (m_1, m_2) \in \{x, y\}^2$ chooses X. Hereon, we economize on notation by writing $\sigma_i^{b,s}$ for the probability an advisor with bias type b and signal s sends message x; and writing $v_s^b(n)$ for the probability the decision-maker with bias b, signal s who hears message n chooses X.

Proposition 1 *There exists a FRDE iff $\rho \geq (1-p)$. If $\rho < (1-p)$ the advisor cannot reveal any information in equilibrium and the decision-maker chooses with his signal.*

Proof: In a fully revealing equilibrium we have, for all i, $\sigma_i^{b,x} = 1 - \sigma_i^{b,y} = 1$ for every $b \in \{x, y\}$ and $i = 1, 2$. In this case the best response strategy for a decision-maker with bias b who observes signal s' and hears message s from the advisor is

$$v_{s'}^b(s) \begin{cases} 1 \quad if & (s, s') \in \{x, x\} \ \forall b \\ 1 \quad if & (s, s') \in \{\{x, y\}, \{y, x\}\} \text{ and } b = x \\ 0 & otherwise \end{cases}$$

The incentive compatibility conditions for a FRDE are obviously satisfied for the decision-maker since he is correctly choosing as if he were fully informed. So, we only need to check the IC constraints for the advisor.

Without loss of generality, assume the sender has bias $b_1 = x$. Then it is trivially the case (mod inversions of natural language) that she surely reveals a signal $s_1 = x$. So suppose she has observed $s_1 = y$. There are only two pivotal events in which the sender's message affects the outcome here: either the receiver shares the sender's bias (x) and has observed a signal $s_2 = y$; or the receiver has the opposite bias (y) and has observed a signal $s_2 = x$. In the former event, the sender's best response is to tell the truth and reveal $s_1 = y$; in the latter event, the sender's best response is to lie and claim $s_1 = x$. The probability of these two events is, respectively, ρp and $(1-\rho)(1-p)$. Therefore the sender is willing to tell the truth when she has signal y iff

$$\rho p \geq (1-\rho)(1-p);$$

that is, iff $\rho \geq (1-p)$. If $\rho < (1-p)$, then clearly the advisor has a strict best response to announce "$s = b$" irrespective of her signal; hence no information is credibly conveyed in debate (the decision-maker can always reveal his signal in debate, but this is of no consequence for the final outcome). Hence the decision-maker must choose on the basis of his own information. Suppose, without loss of generality, that $b_2 = x$. Then player 2 has a dominant strategy to choose X conditional on $s_2 = x$. If $s_2 = y$, however, his payoff from choosing X is $(1-p)/2$ whereas that from choosing Y is $p/2$. Since $p > 1/2$, the best response is to choose Y. $\square$

It is worth emphasising that a FRDE can exist here for very small values of ρ, depending on just how informative is the signal, p: the more information that any given signal provides, the less concern the advisor exhibits about the probability of not being a member of the de facto 'winning coalition'. This result is clearly a direct analogue of the conditions defining existence of a FRDE for the majority rule example with which we began. The analytical advantage here, however, is that the proposition completely describes the equilibrium set up to the mixed equilibrium for the nongeneric event $\rho = (1-p)$. The welfare implications of bias uncertainty here are therefore easy to describe.

Proposition 2 *Assume the most informative equilibrium is played for every parameterization (p, ρ). (a) If $\rho > (1-p)$ then, ex ante, the decision-maker (respectively, advisor) strictly prefers biases to remain secret (respectively, revealed) before the game is played. (b) If $\rho < (1-p)$ then, ex ante, both players strictly prefer biases to be revealed prior to the game being played.*

Proof: Suppose that both players' biases are made common knowledge prior to the game being played. Then either they share the same bias, in which case the most informative equilibrium is a FRDE, or they have opposing biases in which case (as is easy to confirm) no information is revealed by the advisor, player 1. The payoffs to the two players are therefore:

$$EU_1^{rev} = EU_2^{rev} = \rho + (1-\rho)\left(\frac{1+p}{2}\right).$$

To see this, first note that if biases are the same, all information is revealed by player 1 and both players are assured a payoff of one. Next, suppose the biases are revealed to be distinct. Then no information is revealed and the decision-maker can do no better than choose according to his signal; that is, choose X iff $s_2 = x$. In this case the sender (player 1) receives one iff either $s_1 = s_2$, which occurs with probability p, or her bias is x (respectively, y) and $s_2 = x$ (respectively, y), which occurs with probability $(1 - p)/2$. These facts yield the expression for EU_1^{rev}. Similarly, given the biases are distinct, the decision-maker obtains a payoff equal to one iff either $s_1 = s_2$ or her bias is x (respectively, y) and $s_2 = x$ (respectively, y). This justifies EU_2^{rev}.

Now suppose both players' biases remain secret. If $\rho > (1 - p)$ then there exists a FRDE and the decision-maker obtains a payoff $EU_2^{pvt} = 1$ surely. The advisor, however, receives payoff one surely only if the biases are the same; if biases are different, then the advisor reveals her signal and receives payoff one iff $s_1 = s_2$. Thus $EU_1^{pvt} = \rho + (1 - \rho)p$. On the other hand, if $\rho < (1-p)$ then the only equilibrium involves no information revelation and the decision-maker chooses according to his signal. Hence, following the reasoning for the case in which biases are revealed, we obtain

$$EU_1^{pvt} = EU_2^{pvt} = \frac{1 + p}{2}.$$

Therefore, if $\rho > (1 - p)$ the decision-maker gains $(1 - \rho)(1 + p)/2$ when biases are secret but the advisor loses $(1 - \rho)(1 - p)/2$. And when $\rho < (1 - p)$, both players strictly prefer biases to be revealed. □

If the likelihood that the advisor is in fact a member of the 'winning coalition' in that she shares a common bias with the decision-maker, then it the presence of bias uncertainty strictly improves the welfare of that coalition (i.e. the decision-maker) since all decision-relevant information can be shared in equilibrium. Because the two players agree about the desirability of bias-revelation when ρ is sufficiently low relative to p, however, it is interesting to ask whether providing an opportunity to coordinate directly by revealing biases prior to debating signals can improve committee performance. To address this issue, we turn to the second protocol described earlier, whereby there is a prior debate stage in which (effectively) individuals can reveal their biases.[8]

Strategies for this protocol are as follows. Let $\beta_i^{b,s} \in [0,1]$ denote the probability that individual i with bias b and signal s announces her bias is x; given bias-stage messages $(m, m') \in \{x, y\}^2$; let $\sigma_i^{b,s}(m, m') \in [0,1]$ is the probability individual i with bias b and signal s who has announced her bias is m and heard a message that the other player's bias is m', announces that her signal is x; and let $v_s^b((m, m'), (n, n'))$ be the probability the decision-maker with bias b and signal s who announces bias m and signal n, hears the advisor's bias message m' and signal message n', chooses outcome X.

Proposition 3 *Assume each player has observed his or her particular bias and signal. If $\rho \geq 1/2$, there exists an equilibrium in which each player truthfully reveals their bias in the first message round and (1) if the biases are the same, players truthfully reveal their signal in the second round and the receiver chooses on the basis of full information; (2) if the biases are different, players simply announce their true bias independently of their signal and the decision-maker chooses with his signal.*

Proof: Consider the following strategies (specified only for an x-biased individual; those for a y-biased individual are symmetric):

$$\beta_i^{x,x} = \beta_i^{x,y} = 1, \; i = 1, 2$$

$$\begin{aligned}\sigma_i^{x,x}(m, m') &= 1, \forall (m, m'), \; i = 1, 2 \\ \sigma_i^{x,y}(m, x) &= 1 - \sigma_i^{x,y}(m, y) = 0, \forall m, \; i = 1, 2\end{aligned}$$

[8] Of course individuals can talk about anything at any stage. However, given messages are cheap talk, it is not hard to see that assuming that biases are the subject of the first communication round and signals the subject of second is, at least in equilibrium, without loss of generality.

$$
\begin{aligned}
v_x^x(\left(m,m'\right),\left(n,n'\right)) &= 1, \forall(\left(m,m'\right),\left(n,n'\right)) \\
v_y^x((m,m),(\cdot,x)) &= v_y^x((x,y),(x,x)) = 1, \forall m \\
v_y^x(\left(m,m'\right),\left(n,n'\right)) &= 0 \ \textit{otherwise}
\end{aligned}
$$

To confirm that this strategy profile constitutes an equilibrium if $\rho \geq 1/2$, we calculate

$$
\begin{aligned}
E[U_i^{x,x}|\beta_i^{x,x} &= 1,\cdot] = E[U_i^{x,x}|\beta_i^{x,x} = 0,\cdot] = \rho + (1-\rho)\frac{(1+p)}{2} \\
E[U_i^{x,y}|\beta_i^{x,y} &= 1,\cdot] = \rho + (1-\rho)\frac{(1+p)}{2} \\
E[U_i^{x,y}|\beta_i^{x,y} &= 0,\cdot] = \rho\frac{(1+p)}{2} + (1-\rho)
\end{aligned}
$$

where $E[U_i^{b,s}|\beta_i^{b,s},\cdot]$ denotes the equilibrium expected payoff for individual i with bias b and signal s from adopting bias-debate strategy $\beta_i^{b,s}$. Hence telling the truth at the bias revelation stage is incentive compatible iff $E[U_i^{x,y}|\beta_i^{x,y} = 1,\cdot] \geq E[U_i^{x,y}|\beta_i^{x,y} = 0,\cdot]$, which obtains if $\rho \geq 1/2$. And given biases are revealed truthfully, the subsequent debate and decision strategies are easily checked to be best responses. □

Thus allowing an opportunity to coordinate through sharing bias information before revealing anything about decision-relevant signals, cannot improve the welfare properties for the decision-maker when $\rho < 1-p$.[9]

Note that in the identified equilibrium (say, the *bias revelation* equilibrium), so long as $\rho \geq 1/2$, both individuals reveal their biases in the first round of talk; if their biases are

[9] For completeness, we note that there is an essentially identical equilibrium with the three person majority rule committee introduced at the start of these Notes. Suppose in that setting too we introduced two rounds of debate, in which individuals simultaneously send messages regarding their biases in the first round of talk. Then there is an equilibrium in which all biases are revealed in the first stage; those individuals in the majority then reveal their signals truthfully in the second round of debate, and the minority player simply announces his bias again; and finally, the majority votes together for the alternative that maximizes its expected welfare given the information revealed in debate. The minority player votes sincerely under full information. Clearly, just as with the current two player example, the de facto decision does not always reflect the full information welfare maximizing decision for the majority coalition.

the same then all information is revealed in the second round whereas, if biases are different, no further decision-relevant information sharing takes place. It is immediate, therefore, that the expected payoff to any (b, s)-type individual from playing the bias revelation equilibrium identified in Proposition 3 equals the expected payoff they achieve if biases could be revealed ex ante by some external agent. Specifically, for all i, b, s,

$$E[U_i^{b,s} | \text{ biases revealed}] = \rho + (1 - \rho)\,\frac{(1+p)}{2}.$$

Furthermore, because $p > 1/2$, Propositions 1 and 3 together imply there is also a FRDE with a bias debate stage when $\rho \geq 1 - p$: both individuals babble during the first round of talk and, subsequently, all decision-relevant information is revealed during the second round of communication exactly as described in Proposition 1; that is, for all i, b, s, $\beta_i^{b,s} = 1/2$ and $\sigma_i^{b,x} = 1 - \sigma_i^{b,y} = 1$. Hence we have the same welfare comparisons as described in Proposition 2, with the further observation that, when $\rho \geq 1/2 > 1 - p$, the advisor strictly prefers to play the bias revelation equilibrium to playing the FRDE with no bias revelation (of course, the decision-maker has the opposite preferences).

In sum, bias uncertainty can improve the welfare of the de facto full information 'winning coalition'. And even in those circumstances in which bias revelation is available in equilibrium, there exists a superior (from the decision-maker's perspective) equilibrium in which biases remain private information.

5 Conclusion

There is clearly a great deal left to be learned regarding communication in committees. In particular, identifying the optimal voting rule in the presence of debate is an important and open issue.

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 $\cup_{k=0,1,\ldots,n} \mathbf{T}^k(\mathbf{b}) = \mathbf{S}$, 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 2 (Necessity) In any fully revealing debate equilibrium, the restriction to weakly undominated voting strategies implies $v_i(b, s, \mathbf{m}) = y$ 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 σ 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, σ_{-i} is fully revealing of all others' signals and, by the preceding argument on v_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 $s' \in M \backslash \{s\}$,

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

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. □

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

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

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.

Austen-Smith, David and Timothy J. Feddersen. 2006. “Deliberation, preference uncertainty and voting rules.” *American Political Science Review* 100 (May): 209-17.

Caillaud, Bernard and Jean Tirole. 2006. "Consensus building: how to persuade a group" CNRS Working Paper.

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.

Chwe, Michael. 2006. "A robust and optimal anonymous procedure for Condorcet’s model." UCLA, Working Paper.

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. 1997. “Voting behavior and information aggregation in large elections with private information.” *Econometrica* 65(5): 1029-58.

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

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

Glazer, Jacob and Ariel Rubinstein. 2001. "Debates and decisions: on a rationale of argumentation rules." *Games and Economic Behavior* 36(2):158-73.

Hafer, Catherine and Dimitri Landa. 2007. "Deliberation as self-discovery and the institutions for political speech." forthcoming in *Journal of Theoretical Politics.*

McLennan, Andrew. 1998. "Consequences of the Condorcet Jury Theorem for beneficial information aggregation by rational players." *American Political Science Review.* 92(2):413-18.

Meirowitz, Adam. 2004. "In defense of exclusionary deliberation: communication and voting with private beliefs and values" forthcoming in *Journal of Theoretical Politics.*

Meirowitz, Adam. 2007. "Deliberative democracy or market democracy: designing institutions to aggregate preferences and information." *Quarterly Journal of Political Science* 1(4): 373-92.

Ottaviani, Marco and Peter Sorensen. 2001. "Information aggregation in debate: who should speak first?" *Journal of Public Economics* 81(3):393-422.