Majoritarian Debate

Catherine Hafer[*] Dimitri Landa[†]

September 22, 2008

Abstract

We analyze the informational properties of debates under a wide class of debate rules in games of persuasion with agents who vary with respect to what arguments they find convincing. We show that if debate is allowed before each vote on a sequential binary agenda (that is, if the game has *open-debate voting*), then the informativeness of debates and ultimate policy choices vary with voting agendas, despite the existence of ex ante Condorcet winners; there is always a set of voting agendas under which all relevant arguments are always made without delay - i.e., before the first vote, and, under very weak sufficiency conditions, debate is fully informative without delay independent of voting agendas and debate rules. In contrast, when debate is permitted only prior to the first vote on the agenda (that is, if the game has *single-debate voting*), the informativeness of debates and ultimate policy choices are independent of voting agendas, and the relevant institutional tool to affect the informativeness of debate is the debate rule. Choosing over all pairs of voting agenda and debate rule, the most informative delay-free debate under open-debate sequential binary voting is weakly and sometimes strictly more informative than the most informative debates under single-debate majoritarian voting.

1 Introduction

Debates on the choice of a policy or the justifiability of an action are typically envisioned as contests in which each opponent marshals her best, most persuasive arguments in an attempt to sway the audience in favor of his preferred alternative and against the competition. However, with a diverse audience, if arguments are persuasive to some, they are likely to be unpersuasive to others, and the very fact of their potential persuasiveness carries with it

[*]Associate Professor of Politics, NYU, e-mail: catherine.hafer@nyu.edu
[†]Assistant Professor of Politics, NYU, e-mail: dimitri.landa@nyu.edu

the danger that the unpersuaded will turn against the alternative favored by the speaker; after all, they are receiving evidence that the reasons that appear to underlie the appeal of that alternative do not - for them - hold water. This suggests that, contrary to the popular image, making one's best arguments in debate may not always be the best course of action. What should we, then, expect from public debate? Under what institutional and political circumstances will individuals with diverse interests and reasons share their reasons with each other? When will public debate be most informative? The present paper addresses these questions by considering the properties of debates that precede the collective selection of a binding policy choice.

We model persuasion in debate as resulting from the communication of reasons or arguments that resonate (or not) with some relevant (proper) subset of the audience. Thus, an anti-abortion debater may argue that the preservation of any form of human life trumps anything else in a conflict of rights. A pro-abortion rights debater may argue that women should have autonomy in dealing with matters concerning their own bodies. Whether one of these arguments resonates for a given listener depends on whether that listener happens to share the speaker's foundational assumptions for it. If she does, then a (coherent) argument will be persuasive, and the speaker may succeed in convincing the listener to see that argument as a reason for her preferred policy choice. If not, then it may become a reason to support a different policy.

For each agent, some of the possible arguments are *active*: she knows whether or not she finds them persuasive, and they determine her perceived optimal choice. Other possible arguments are *latent*: she is initially unsure whether they are persuasive, but finds out when she hears those arguments being made. Persuasion is the effect of hearing such arguments: the "activation" of the corresponding latent reasons, which turns them into active ones.

We show that majoritarian politics creates strong incentives for the participants in debate to make their arguments. In particular, under simple majority rule with the option of restarting debate before each vote (*open-debate voting*), there is always a class of binary sequential agendas that lead to all potentially consequential informative arguments being made before the first vote regardless of the debate rule. We show, further, that the informativeness of debates and ultimate policy choices under open-debate voting are not independent of the voting agendas, despite the existence of the ex ante Condorcet winners at every point in the voting sequence and under every possible information state. Thus, the possibility of debate undermines the standard result in voting theory on the invariance of voting outcomes to agenda choice given the existence of a Condorcet winner. In contrast, when debate is permitted only prior to the first vote on the agenda (*single-debate voting*), the informativeness of debates and ultimate policy choices under open-debate voting are independent of voting agenda. The choice of debate rules can increase the informativeness of debates but, unlike the choice of voting agendas under open-debate voting, cannot guarantee maximally revealing debate. Finally, we construct examples showing that other prominent voting rules, including plurality rule and rules that designate a privileged alternative, are inferior in their potential for eliciting informative argumentation.

The remainder of the paper is organized as follows. In Section 2, we discuss the related literature on deliberation. Section 3 introduces the key elements of our model, including the formal definitions of the informational and decision-making environment. In Section

4 we present our results and some illustrative examples. The final section concludes with a discussion, including a brief comparison of simple majority to other voting rules. The Appendix gathers proofs of the formal results.

2 Relation to the Existing Literature

The existing formal literature on deliberation and debate is divided between models of cheap-talk signaling and models of persuasion in which speech can be said to be partially or fully verifiable. The latter are closer to the present model. Lipman and Seppi (1995) and Lanzi and Mathis (2004) analyze a sender-receiver model in which senders can supply partial proofs for their signals and in which the veridicality (truth content) of a message is the same for all types of receivers - that is, all types of receivers are convinced by the same messages. Common veridicality is also assumed in Glazer and Rubinstein (2006) and in Patty (forthcoming), who analyze a model of persuasion in which the messages contain "hard" or fully verifiable information. In Glazer and Rubinstein's model, this information is partial and the informativeness of messages is a function of speaker credibility. By contrast, in the model we analyze, messages are fully verifiable and complete, yet their veridicality (truth content) differs across agents - as in the abortion rights example in the Introduction. In Patty's model, unlike in the present paper, players' preferences need not be driven by the propositional content of arguments they find convincing.

The closest models to the one developed in the present paper are Glazer and Rubinstein (2001) and Hafer and Landa (2006 and 2007). The latter analyze related models with deliberation that, like in the present model, satisfies full verification and private veridicality.

They focus on the welfare properties of the equilibria of a game in which agents choose how to allocate their deliberative resources between receiving (and processing) arguments and making arguments to others and on the aggregate consequencs of equilibrium individual choices of deliberative venues, respectively. Glazer and Rubinstein consider the informational properties of debate rules, including the simultaneous speech and one-shot finite sequential speech rules that are covered by our analysis below. Unlike the model in the present paper, their model is one of common veridicality between the speakers and the (single) listener and assumes that speakers are better informed than the listener about what the latter will find persuasive. Although they restrict speakers to making truthful arguments, their setup is closer to the standard cheap-talk model. Austen-Smith (1993) and Krishna and Morgan (2001) analyze the effects of different communication rules within a cheap talk model in which two speakers attempt to influence the actions of a third party. Battaglini (2002) examines a cheap-talk model with multiple senders and a multidimensional state variable.

The analysis of the effects of different voting rules on the informational quality of deliberation has so far focused on cheap-talk signaling in settings that allow for private values. In this literature, Gerardi and Yariv (2006) show that given unrestricted pre-vote communication, all non-unanimity voting rules generate equivalent sets of sequential equilibrium outcomes, but restrictions on communication protocols may lead to different outcomes. Austen-Smith and Feddersen (2006) show that majority rule can do better than unanimity rule in generating informative pre-vote communication and that when unanimity rule generates complete information revelation, the same can be accomplished with other voting rules, including majority rule. Meirowitz (2007) shows that fully revealing equilibria exist if and only if participants of all possible preference types believe that a majority of the group share their type.

3 The Model

3.1 The Informational Framework

3.1.1 The Formalism

We assume that policy choices can be ordered on a left-right ideological spectrum $[0, 1]$ with 0 representing the left and 1 representing the right. Agent i does not know her true ideal policy point $\theta_i \in [0, 1]$ on that spectrum, but, as discussed in the Introduction, there is a relationship between i's ideal point and the arguments that she finds persuasive. In particular, assume that i's true ideal point θ_i probabilistically determines which arguments she will find persuasive. To capture this formally, for each i let $m_i \in \{0, 1\}^n$ be an indicator vector the elements of which reflect what arguments i finds persuasive, where the value of the j^{th} element of that vector, m_i^j, $j = 1, ..., n$ can be interpreted as indicating whether a person i is convinced by a "right-leaning argument" (indicated as $m_i^j = 1$) or by "a left-leaning argument" (indicated as $m_i^j = 0$) with respect to the dimension j, with the prior probability $\Pr(m_i^j = 1) = \theta_i$.

Because i is assumed to make correct inferences about the value of m_i^j upon hearing argument j regardless of whether that argument is right- or left-wing, in this environment, right- and left-wing arguments are equally informative for any given i.[1] (We comment on

[1] This symmetry may be seen to reflect the conjunction of two assumptions: first, that a right-wing

the robustness of our results to the assumption that argument j is more informative for i when i is convinced by it in the Discussion.)

Thus, i can make inferences about her true policy ideal point from the arguments that she does or does not find persuasive (using Bayes' rule), so that $m_i^j = 1$ indicates i's greater marginal preference for a right-wing policy, that is it indicates i's posterior belief that θ_i is higher. (Note that, because the value of θ_i is agent-specific and the persuasiveness of each of i's reasons is independently drawn, different members of a society will not necessarily find the same arguments persuasive.)

Although each i is uncertain about θ_i, the distribution of θ_i, described by the probability density function $p(\theta)$, is common knowledge. Nature randomly selects and reveals some arguments to i, who may find some of them persuasive and others not. We say that argument j is *active* for i if m_i^j is revealed to i and *latent* if it is not. Let r_i^t be the number of i's active arguments at time t for which $m_i^j = 1$ and l_i^t be the number of such arguments for which $m_i^j = 0$. While the true value of the vector m_i cannot change, i may learn the values of more of its dimensions, and thus r_i^t and/or l_i^t may increase as a result of argumentation.

3.1.2 Interpretation

Before proceeding further, we note an interpretive point about the relationship between the informational model described above and the standard representation of beliefs. The

argument can be turned into a left-wing argument by adding "isn't that silly, and therefore, you must support the opposite view" at the end of it (and vice versa), and second, the assumption that there is no bias in the cognitive processing of right- vs. left- wing arguments. Such a bias could be incorporated by assuming that the true policy ideal point is some function of λ_1 and λ_2, where $\Pr(m_i^j = 1 | j \text{ is "right-wing"}) = \lambda_1$, $\Pr(m_i^j = 0 | j \text{ is "right-wing"}) = \lambda_2$, and $\lambda_1 \neq 1 - \lambda_2$ to reflect the fact that, for idiosyncratic reasons not a function of the individual's ideology, i's response to an argument depends not only on its informational content but also on its rhetorical packaging. We do not pursue this possibility in the present paper in order to focus attention on the strategic incentives for speech.

informational model we adopt is consistent with a number of analytically distinct yet empirically plausible interpretations of debate, including an interpretation that is consistent with the standard state-space representation[2] and one that cannot be captured by that representation.

Consider the latter interpretation first. Suppose that members of a group are choosing a policy that is appropriate for a fixed but uncertain state of the world and have a non-degenerate distribution over a subset of alternatives/states that includes a policy choice/state x. They are considering an argument to the effect that if A (a commonly known fact whose previous relevance to policy preference was uncertain) then the policy preference ought to be x because of B (that is, $\Pr(x|A, B) = 1$). Suppose the persuasiveness of the argument with respect to policy choice x is determined by the background beliefs B - in particular, A implies x if and only if B ($\Pr(x|A) = 1$ iff B). A listener i who shares the belief B is, thus, persuaded of the relevance of A. Following the argument, then, i would update her policy preference toward x. By contrast, a listener j who does not share beliefs B is not persuaded (for j, the argument that A is relevant was neither active nor latent). If A is relevant only if B (that is, if the argument in question is the best possible argument that could be made for the relevance of A, i.e., $\Pr(x|A) \neq \Pr(x)$ only if B), then j concludes that A is not relevant. If part of the possible appeal of x hinged on the possibility of the relevance of A (i.e., $\Pr(x|\Pr(x|A) \neq \Pr(x)) > \Pr(x|\Pr(x|A) = \Pr(x))$ then j will change her policy preference away from x as a consequence.

[2]Dekel, Lipman, and Rustichini (1997) identify "standard" state-space representation with two assumptions: first, "that all states in the model are real possibilities, as opposed to objects present only to describe the agent's perception of possibilities," and second, "that whether the agent knows a given fact at a given state is entirely determined by the set of states in which this fact is true." As the foregoing discussion makes clear, debate under the first interpetation we discuss violates the second assumption.

To see that this interpretation is not consistent with the standard state-space representation of beliefs, note that, because the argument is true if and only if B is true, if i knows B and reasons correctly, she must already have concluded that A is relevant. That is, her prior belief with respect to the relevance of A cannot be different from the posterior following the event that is the argument. Similarly, if j knows that she does not share B, she must already have concluded that A is not relevant and the argumentation in debate cannot possibly affect her beliefs. Thus, under this interpretation, debate cannot be represented as a consequential event using the standard state space representation.[3]

Consider next an example that points to a different interpretation of our informational model, this time consistent with the standard state-space representation. Let both i and j share the following beliefs: $\Pr(A) = \frac{2}{3}$, $\Pr(\neg A) = \frac{1}{3}$, $\Pr(B) = \frac{2}{3}$, and $\Pr(\neg B) = \frac{1}{3}$. However, whereas i also believes that $\Pr(A|\neg B) = 1$, j believes $\Pr(\neg A|\neg B) = 1$. Consider a publicly observed event indicating that $\neg B$. Following this event, i and j, whose prior unconditional beliefs regarding A were identical, will diverge in their posteriors. What we refer to above as the latent argument is, in this example, a *combination* of (a) publicly observable and convincing to all evidence that $\neg B$ and (b) an implication from the conditional probabilities of A given $\neg B$. The common unconditional priors are consistent with identical distributions of active arguments or with there being no active argument at all. For consistency and because the first interpretation fits with a natural language description of debates somewhat better, we adopt this interpretation throughout (though recognizing that it is but one of a

[3] The relatively recent work on non-standard state-space representation of beliefs includes Li 2006; Heifetz, Meier, and Schipper 2006; and others. Although this work is clearly relevant to applied literature on strategic communication, it is outside the focus of the present paper.

number of possible interpretations).

3.2 The Game Form

We assume that the population is finite and divided into speakers and listeners. Let $\mathcal{S}$ be the set of speakers and $\mathcal{R}$ the set of listeners.

The game form is defined by a combination of a voting rule and a debate protocol. Unless specified otherwise, we consider debate in relation to majoritarian preference aggregation with a finite sequential binary voting agenda. The set of voters is $\mathcal{R}$.[4] Let Π be the finite set of alternatives from which the collective choices are being made and let $z := |\Pi| - 1$. A sequential binary agenda is an agenda of the following form: index the alternatives in the set Π as $a_0, a_1, ..., a_z$ such that the listeners first vote over $\{a_0, a_1\}$, then over $\{w_1, a_2\}$, then over $\{w_2, a_3\}$, and so on, where w_k is the majority rule winner of the k^{th} vote, and w_z (the majority choice in the z^{th} vote) is the ultimate collective choice. Say that a game has *open-debate voting* if a debate following the given debate protocol (formally defined below) is permitted before each vote. In contrast, a game has *single-debate voting* if it allows debate only before the first vote. We use the terms *debate* j and the j^{th} *debate* interchangeably to refer to the debate that would take place between the $(j-1)^{\text{th}}$ and j^{th} votes. Further, we refer to the sequence of debates $\{1, ..., z\}$ under open-debate voting or the single-debate under single-debate voting as the *grand debate*.

A *debate protocol* is a triple consisting of a set of speakers $\mathcal{S}$, a set of speaker messages, and a debate rule. For each speaker i, her speech at time t is the set of arguments that

[4] Including $\mathcal{S}$ in the set of voters complicates the expression of the proofs without altering the substance of the results.

she makes, $D_i^t \subseteq \{1, ..., n\}$. On dimensions $\{1, ..., n\} \backslash D_i^t$, she may be thought to be silent or making content-free speech (e.g., "this is a great country"). When it is necessary to distinguish between different debates, we represent i's speech at time t in debate T as $D_i^{T,t}$.

A *debate rule* is defined to be a pair of a speaking order and a debate termination protocol. A *speaking order* O is an ordered list of sets of speakers, o elements long, such that $O^t \subseteq \mathcal{S}$ for $t = 1, ..., o$ and $\bigcup_{t=1}^{o} O^t = \mathcal{S}$. Thus, the definition of the speaking order requires that each member of $\mathcal{S}$ has an opportunity to speak at least once. At every time $t = 1, ..., o$, every $i \in O^t$ may choose any $D_i^t \subseteq \{1, ..., n\}$, and every listener $k \in \mathcal{R}$ learns the persuasiveness to her, m_k^j, for every $j \in D_i^t$, upon hearing D_i^t. Note that the speaking order O encompasses sequential and simultaneous speech. The *debate termination protocol* specifies how the debate ends relative to the speaking order. An exogenously imposed termination corresponds to a single cycle of O. An endogenous termination corresponds to repeating O indefinitely until no new speech is made for one complete cycle. In this case, for every natural number k, $i \in O^{kt}$ may choose any $D_i^{kt} \subseteq \{1, ..., n\}$ at time kt. After each speech (or, for t s.t. $|O^t| > 1$, each set of speeches), listeners update their beliefs about their types consistent with Bayes' Rule.

Let π_j be speaker j's most preferred alternative in Π and let $C(X)$ be the convex hull of X. We make the following intuitive restriction on the set of alternatives Π:

$$\text{For all } a \in \Pi, a \in C(\{\pi : \exists j \in \mathcal{S} \text{ s.t. } \pi_j = \pi\}).$$

In other words, the most extreme left and right elements of that set are such that the set of speakers $\mathcal{S}$ includes at least one speaker for whom the left-most extreme alternative is the most preferred one in the set, and at least one speaker for whom the right-most extreme alternative is the most preferred one in the set.[5] In fact, for our results, it is equivalent to posit that $\mathcal{S}$ has two elements, $\{L, R\}$; we adopt this simplification for the remainder of the paper and will, for convenience, refer to the left and right speakers.

3.3 Preferences and Beliefs

Given agent i's ideal policy θ_i and the collective choice w_z, we assume that i's utility, $\hat{u}_i$, is strictly concave in the distance between the collective choice and θ_i - in particular, $\frac{\partial \hat{u}_i}{\partial(|w_z-\theta_i|)} < 0$ and $\frac{\partial_i^2 \hat{u}}{\partial(|w_z-\theta_i|)^2} < 0$. This guarantees a single-peaked preference profile. Thus, the majority-preferred alternative is determined by the preferences of the listener who is the median voter *given the information available at that point in the game.* Because different voters may respond differently to the same messages, the identity of the median voter may change as new information becomes available.

To keep things simple, we assume that it is common knowledge that $\theta_i \sim U[0,1]$ $\forall i$. We assume, moreover, that θ_j $\forall j \in \mathcal{S}$ is common knowledge and that each speaker $i \in \mathcal{S}$ knows the biases r_j^t, l_j^t $\forall j \in \mathcal{R}$.[6]

Agent i's complete knowledge of the persuasiveness of arguments does not imply certain knowledge of her preferred policy position θ_i. Accordingly, listener i's beliefs at time t about

[5]This assumption can be straightforwardly justified in the context of endogenous choices of speaking vs. listening roles in deliberation - see the model of such choices in Hafer and Landa (2006), which yields it as an inference.

[6]As a matter of interpretation, we suppose that speakers' reasons are active, i.e., that $\mathcal{A}_i = \{1, 2, ..., n\}$ and thus $l_i^0 + m_i^0 = n$ $\forall i \in \mathcal{S}$. The speakers' ability to make potentially persuasive arguments on every issue dimension follows naturally from such a supposition. However, in the technical sense it plays no role in establishing the results that follow and thus is, strictly speaking, unnecessary. Similarly, relaxing the assumption that $i \in \mathcal{S}$ knows θ_j $\forall j \in \mathcal{S}$ has no effect on our substantive results.

her own type θ_i can be characterized by the probability density function

$$p(\theta|r_i^t, l_i^t) = \frac{\theta^{r_i^t}(1-\theta)^{l_i^t}}{\int_0^1 \hat{\theta}^{r_i^t}(1-\hat{\theta})^{l_i^t} d\hat{\theta}}, \tag{1}$$

where, as before, r_i^t and l_i^t represent the number of right and left-wing arguments, respectively, known to i at time t. Let $t = 0$ indicate the initial information state. When it is necessary to distinguish between different debates, we use $r_i^{T,t}$ and $l_i^{T,t}$ to represent the numbers of effective right- and left-wing arguments known to i at time t in debate T. Efficient updating upon receiving a message requires agent i to update her beliefs about θ_i using Bayes' Rule. Let $u_i(\pi|r_i^t, l_i^t)$ be i's expected utility from policy π given $p(\theta|r_i^t, l_i^t)$.

4 Analysis

4.1 Equilibrium

The equilibrium concept is Perfect Bayesian Equilibrium in undominated behavioral strategies. In particular, at every point in the game, players' beliefs are derived from their common prior and their active arguments via Bayes Rule; i votes for the alternative $v_i^T \in \{w_{T-1}, a_T\}$ at vote T according to strategy that maximizes her expected payoffs given other players' future voting and speaking strategies and expected responses to speech; and i's equilibrium speaking strategy at time t in debate T, D_i^{T,t^*}, maximizes her expected payoff given other players' speaking strategies D_j^{T,t^*}, their future voting and speaking strategies and expected responses to speech. The formal definition of the equilibrium conditions is in the Appendix.

4.2 Voting Agendas and the Informativeness of Debate

Throughout the paper, we restrict our attention to situations in which it is in principle possible, in expectation from the speakers' standpoint, for revelation of arguments to change the preferences over alternatives (and thus the behavior) of enough voters to produce a policy choice different than the ex ante Condorcet winner.[7]

We make use of the following definitions. We say that an argument is *potentially consequential* (at time t in debate j) if there is a positive probability that making it will change the policy outcome relative to the Condorcet winner under the current information state. *Debate j reveals a new argument* if in the course of that debate one of the senders makes an argument that was not made in an earlier debate. *Debate j is fully revealing* if and only if there is no subsequent debate in which a new argument that is potentially consequential for the policy outcome could be revealed. The *grand debate is fully revealing* if and only if it reveals all potentially consequential arguments. The *grand debate is fully revealing without delay* if and only if debate 1 is fully revealing.

We begin our analysis with an example, which provides some of the intuition for our main results.

Example 1. Suppose the set of listeners/voters $\mathcal{R} = \{1, 2, 3, 4, 5, 6, 7\}$, $n = 5$ possible arguments, and $\hat{u}_i(\pi) = -(\theta_i - \pi)^2$. Suppose the ideal points of the left and right speakers are $\theta_L = 0$ and $\theta_R = 1$; let the initial information of the listeners be $(r_1, l_1) = (r_2, l_2) = (0, 4)$; $(r_3, l_3) = (0, 1)$; $(r_4, l_4) = (1, 1)$; $(r_5, l_5) = (1, 0)$; and $(r_6, l_6) = (r_7, l_7) = (4, 0)$. The set

[7] Formally, suppose the median voter prefers y, and let x be the largest alternative less than y and z be the smallest alternative greater than x (so that $x < y < z$). Then this condition amounts to restricting attention to the cases in which either $|\{j \in \mathcal{R} : u_j(z|l_j^0 + |\mathcal{L}_j^0|, m_j^0) \geq u_j(y|l_j^0 + |\mathcal{L}_j^0|, m_j^0)\}| \geq \frac{|\mathcal{R}|}{2}$ or $|\{j \in \mathcal{R} : u_j(x|l_j^0, m_j^0 + |\mathcal{L}_j^0|) \geq u_j(y|l_j^0, m_j^0 + |\mathcal{L}_j^0|)\}| \geq \frac{|\mathcal{R}|}{2}$ (or both).

of alternatives is $\{\frac{1}{7}, \frac{1}{2}, \frac{6}{7}\}$. Based on the listeners' initial information, their expected ideal points are, respectively, $\frac{1}{6}; \frac{1}{6}; \frac{1}{3}; \frac{1}{2}; \frac{2}{3}; \frac{5}{6}; \frac{5}{6}$. Thus, in the absence of additional information, 1 and 2 prefer $\pi = \frac{1}{7}$; $3, 4$, and 5 prefer $\pi = \frac{1}{2}$; and 6 and 7 prefer $\pi = \frac{6}{7}$. Note that no additional learning could change the preferences of $1, 2, 6$, and 7; but $3, 4$, and 5 could each be moved toward either of the other alternatives.

1a. Consider the agenda $a_0 = \frac{1}{7}, a_1 = \frac{6}{7}, a_2 = \frac{1}{2}$ under open-debate voting. The alternatives, the preference profile, and the information state are symmetric. Hence, the probability that $\frac{6}{7}$ beats $\frac{1}{2}$ under full information is equal to the probability that $\frac{1}{7}$ beats $\frac{1}{2}$ under full information. Thus, full revelation before the first vote leads to $\frac{1}{7}$ with probability p, $\frac{6}{7}$ with probability p, and $\frac{1}{2}$ with probability $1 - 2p$. If there is no information transmission before the first vote, the median voter (4) is indifferent and will choose each alternative with probability $\frac{1}{2}$. If $\frac{1}{7}$ wins the first vote, then the second-vote agenda is $\{\frac{1}{7}, \frac{1}{2}\}$ and, unless there is additional speech, 4 will prefer in the second vote her expected ideal point $\frac{1}{2}$. But then L will prefer to make all arguments before the second vote. If $\frac{6}{7}$ wins the first vote, then the second-vote agenda is $\{\frac{6}{7}, \frac{1}{2}\}$ and again, unless there is additional speech, 4 will prefer in the second vote her expected ideal point $\frac{1}{2}$. But then R will prefer to make all arguments before the second vote. If there is no information transmission before the first vote, then, assuming symmetric equilibrium in which the median voter tosses a fair coin, $\frac{1}{7}$ is selected with probability $\frac{1}{2}p$, $\frac{6}{7}$ with probability $\frac{1}{2}p$, and $\frac{1}{2}$ with probability $1 - p$. Concavity of speaker preferences insures that both L and R prefer the latter lottery (corresponding to revelation with delay) to the former one.

1b. Suppose, though, that the agenda is $a_0 = \frac{1}{7}, a_1 = \frac{1}{2}, a_2 = \frac{6}{7}$, still under open-debate voting. Then all arguments are made without delay - that is, before the first vote. To see this, suppose first that $\frac{1}{2}$ wins the first vote and is preferred to $\frac{6}{7}$ before the second debate. Then R will prefer in that debate to make all possible arguments. But then, anticipating it, L prefers to make all arguments in the first debate. If $\frac{1}{2}$ wins the first vote but $\frac{6}{7}$ is majority preferred going into the second debate, then L will prefer to make all arguments, but will also prefer to have done so in the first debate. Suppose now that $\frac{1}{7}$ wins the first vote and is preferred by the majority to $\frac{6}{7}$ going into the second debate. Then R will want to make all possible arguments, and will prefer to have done so in the first debate. (It is not possible for $\frac{1}{7}$ to win the first vote but for $\frac{6}{7}$ to be majority preferred going into the second debate.) Thus, given open-debate voting, there is a fully revealing debate before the first vote.

These examples develop the intuition that sometimes we may be able to construct a voting agenda that leads to fully revealing debate without delay. Our first two results address this intuition. In the environment with a binary set of alternatives (note that the distinction between single- and open-debate voting is moot here), we have the following result:

Theorem 4.1 *Let $|\Pi| = 2$. Then, regardless of the debate rule in every debate equilibrium the debate is fully revealing if at least one of the following conditions does not hold:*

(a) in the initial information state, the ex ante median voter's most preferred alternative is $\frac{1}{2}$;

(b) $\{E[\theta_i | r_i^0, l_i^0]\}_{i \in \mathcal{R}}$ are symmetric around $\frac{1}{2}$; and

(c) $\{\pi^1, \pi^2\}$ is symmetric around $\frac{1}{2}$.

Otherwise, the speakers are indifferent between making arguments and not.

Our next result considers the prospect of a fully revealing grand debate under open-debate voting for $|\Pi| > 2$.

Theorem 4.2 *Let $|\Pi| > 2$ and suppose the game has open-debate voting. Then:*

1. *in expectation, the grand debate is always fully revealing;*
2. *there exists a non-empty set of agendas that yield fully informative grand debate without delay.*
3. *the grand debate is fully revealing without delay if either:*

   (a) there exist no $x, y \in \Pi$ such that the ex ante median voter is indifferent between x and y; or

   (b) if there exist $x, y \in \Pi$ such that the ex ante median voter is indifferent between x and y, then for any such pair, either x or y is the last alternative in the agenda or there exists $z \in \Pi$ such that $z \in (x, y)$ and z appears on the agenda before either x or y.

Thus, although open-debate voting permits debate before each binary vote, when conditions in part 3 of the Theorem are satisfied, the debate that takes place before the first vote contains all the information relevant to the policy choice. In effect, debates before the subsequent votes need never take place, so long as they are permitted - that is, so long as any speaker can insist on having them. Note that condition (a) in part 3 is a condition on the set of alternatives, while condition (b) is on the agenda. Part 2 of Theorem 4.2 shows that we can always construct an agenda that yields full revelation without delay under open-debate voting. Thus, the set of agendas that generate fully revealing grand debate without delay is always non-empty, and possibly unconstrained - when condition (a) in part 3 of Theorem 4.2 holds.

Example 1 underscores the next theorem and its corollary, which show how the possibility of debate affects the robustness of a key result in voting theory.

Theorem 4.3 *The informativeness of debate is ex ante invariant with respect to the choice of voting agenda under single-debate voting but not under open-debate voting.*

As is well known, if, in a standard voting game, there exists a (ex ante) Condorcet winner, then the choice of the agenda is (ex ante) immaterial: the (ex ante) outcome will always be that (ex ante) Condorcet winner. However, given the presence of debate, we have the following corollary to Theorem 4.3:

Corollary 4.1 *The existence of the (ex ante) Condorcet winner does not guarantee agenda-independence of outcomes.*

The intuition for these results is as follows. Given our assumptions on voters' preferences, the ex ante Condorcet winner always exists both under single-debate or at any point in the voting sequence under the open-debate voting. Debate may change the information state, and changing the information state may, in turn, change the Condorcet winner. This alone does not imply the agenda-dependence of outcomes. Because under single-debate voting, all debate must occur before voting begins, the identity of the Condorcet winner is fixed for the entire sequence of votes. As a result, the sequence of the votes is irrelevant for the outcome. However, under open-debate voting, the possibility of debate after each pairwise vote means that speakers may prefer to delay revelation (and thus delay changing the Condorcet winner) until some alternatives have been eliminated in earlier votes. Whether or not such delay occurs depends on which alternatives are likely to be eliminated in this manner, i.e. on the voting agenda.

4.3 Single- vs. Open-Debate Voting

Theorem 4.2 suggests that under open-debate voting, voting agenda choice is an effective tool of achieving full revelation without delay. In contrast, Theorem 4.3 suggests that under single-debate voting, voting agenda choice is not an effective tool at all. The following example shows, however, that the informational content of debate under the single-debate voting is affected by debate rules.

Example 2. Suppose the set of listener-voters $\mathcal{R} = \{1, 2, 3\}$, $n = 2$ possible arguments, and that for all $i \in \mathcal{R}$, $r = l = 0$ initially - that is, the listeners do not know the persuasiveness of any arguments. The set of alternatives is $\{\frac{1}{4}, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}\}$. Let the utilities of the left speaker (L) be $\hat{u}_L(\frac{1}{4}) = 1$, $\hat{u}_L(\frac{1}{2}) = \frac{8}{9}$, $\hat{u}_L(\frac{2}{3}) = \frac{2}{3}$, $\hat{u}_L(\frac{3}{4}) = 0$. The ex-ante Condorcet winner is $\frac{1}{2}$.

2a. Suppose single-debate voting with an exogenous termination rule; first L speaks, then R speaks. R prefers making arguments on one dimension to making no arguments, conditional on L having chosen not to make arguments. To see this, note that if after speech, a majority has $r = 1$, the outcome is $\frac{2}{3}$, but if a majority has $l = 1$, the outcome is still $\frac{1}{2}$. (In the latter case, their ideal points are $\frac{1}{3}$, but $\frac{1}{3}$ is not an alternative and $\frac{1}{2}$ is closer than $\frac{1}{4}$ to $\frac{1}{3}$.) R also prefers arguing on one dimension to arguing on two - by concavity, given that possible ideal points are $\{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$. Anticipating that R will argue on one dimension if L says nothing, L prefers saying nothing to making arguments on both dimensions:

$$\frac{1}{2}\hat{u}_L(\frac{1}{2}) + \frac{1}{2}\hat{u}_L(\frac{2}{3}) > \frac{10}{64}\hat{u}_L(\frac{1}{4}) + \frac{44}{64}\hat{u}_L(\frac{1}{2}) + \frac{10}{64}\hat{u}_L(\frac{3}{4}).$$

Further, speaking on one dimension makes L worse off because if a majority learns $r = 1$, R may prefer to respond by speaking on the other dimension (in which case, L is in expectation worse off: $\frac{4}{9}\hat{u}_L(\frac{3}{4}) + \frac{5}{9}\hat{u}_L(\frac{1}{2}) < \hat{u}_L(\frac{2}{3})$), whereas if a majority learn $l = 1$, R says nothing. Thus, in equilibrium, L prefers not to make arguments and R prefers making arguments on one dimension.

2b. Suppose next an endogenous termination rule, again under single-debate voting. Ex ante, both L and R prefer no revelation to revelation on both dimensions. L prefers no revelation to revelation on one dimension. R prefers revelation on one dimension to revelation on none, but if R speaks on one dimension and a majority of voters learn $l = 1$ from her speech, then, given the updated beliefs of the voters, L will prefer to speak on the other dimension in response. If, on the other hand, a majority learn $r = 1$ from R's speech, then, given the updated beliefs of the voters, R subsequently may prefer to make arguments on the other dimension. Thus, R anticipates that revelation on one dimension will lead to revelation on the other dimension at least when the outcome is not in his favor, and possibly even when it is. Thus, R prefers making no arguments to making arguments on one dimension, and so no arguments are made in equilibrium under an endogenous termination rule.

Our next result provides a welfare comparison of single- and open-debate voting, comparing across all possible combinations of agenda and debate rules for any given set of preferences and alternatives.

Theorem 4.4 *Holding fixed agents' preferences and the set of alternatives, the most informative equilibrium grand debate that can be induced by any agenda and debate rule under single-debate voting is always weakly and sometimes strictly less informative than the most informative equilibrium grand debate that can be induced by any agenda and debate rule under open-debate voting.*

5 Discussion

5.1 Informational Properties of Debate with Simple Majority Rule

Our analysis identifies three conditions that affect the informational consequences of debate under majoritarian agendas: the possibility of starting a debate before each vote (that is, single- vs. open-debate voting), the choice of a debate rule, and the voting agenda (the sequence in which the alternatives are voted on). The conjunction of Theorems 4.2 and 4.4 shows that the very possibility of starting a debate before each vote - without that possibility necessarily realized on the path of play - matters for the equilibrium informational properties of debate. That possibility alone serves to substantiate the threat of retaliatory speech; when it is absent, debate is weakly less informative.

The weak sufficiency conditions in Theorems 4.1 and 4.2 suggest that under the open-debate voting, majority rule is highly conducive to eliciting fully informative debates without delay; when such first-best outcomes cannot be guaranteed regardless of voting agenda or debate rule, the choice of a voting agenda is an effective tool of achieving a first-best. Under single-debate voting, however, the relevant institutional lever is the debate rules, and their maximal effectiveness cannot be as high as that of the agendas under open-debate voting.

Finally, it is worth noting that our results are robust to relaxing the assumptions on agency maintained above to allow for inefficient updating following an "unconvincing" argument. Following such argument, a Bayesian receiver updates away from the policy position that is supported by that argument or by the possibility that she might be convinced by it. Experimental evidence reported in Dickson, Hafer, and Landa (2008) suggests that subjects may fail to do so in the deliberation games set in a laboratory environment: their prior policy positions are "sticky" or subject to a version of confirmatory bias. In the presence of a single debater in the present environment, such inefficient updating would alter the speaker's strategy because the downside of making an argument is decreasing in the proportion of receivers who update that way. With multiple speakers, however, that effect is neutralized: if a given speaker chooses to make an argument on a given dimension, or is expected to do so, then a speaker with an opposite preference will have a dominant strategy to speak on that dimension, sinch doing so could only bring on board the inefficient updaters who were unconvinced by the competition.

5.2 Simple Majority vs. Other Voting Rules

Although our primary focus is on the properties of debates with simple majority voting, it is instructive to consider briefly whether the properties we have identified distinguish environments with simple majority voting from those with other common voting rules. In this subsection, we provide some evidence that these properties are not universal.

Consider first the following example, which shows that sequential-agenda majority rule can do strictly better than plurality rule.[8]

Example 3. Assume the environment in Example 1 except that now, let $(r_4, l_4) = (3, 2)$. Based on the listeners' initial information, their expected ideal points are, respectively, $\frac{1}{6}; \frac{1}{6}; \frac{1}{3}; \frac{4}{7}; \frac{2}{3}; \frac{5}{6}; \frac{5}{6}$. As in Example 1, in the absence of additional information, 1 and 2 prefer $\pi = \frac{1}{7}$; 3, 4, and 5 prefer $\pi = \frac{1}{2}$; and 6 and 7 prefer $\pi = \frac{6}{7}$. Thus $\pi = \frac{1}{2}$ is the ex ante plurality winner. Note that 4 is now fully informed, so her preference is fixed, and so 3 and 5 are the only movable voters. Because their positions are symmetric, the probability distributions over their possible responses to arguments are also symmetric. It follows that the probability that informative speech creates a plurality for $\frac{1}{7}$ is equal to the probability that it creates one for $\frac{6}{7}$. L's expected loss (relative to $\pi = \frac{1}{2}$) from $\pi = \frac{6}{7}$ is greater than her expected gain (relative to $\pi = \frac{1}{2}$) from $\pi = \frac{1}{7}$. Thus, given the symmetry of the distribution of possible outcomes in response to informative speech, L prefers that there be no informative speech, and thus, her weakly dominant strategy is not to speak. R's position is symmetric, so no informative speech occurs in equilibrium under plurality rule.[9] Note, though, that, as argued above, there exists a binary voting agenda such that it will generate full debate without delay under open-debate voting.

The second comparison is to preference aggregation rules that designate a privileged alternative. An important property of simple majority rule is its neutrality with respect to the alternatives under consideration. In particular, given a set of alternatives, it (1) allows for any possible binary agenda on that set and (2) requires a simple majority to defeat any

[8] Note that although this need not be the case for plurality rule generally, symmetric "strategic voting" in this example is behaviorally equivalent to the "sincere voting" (for one's first best alternative).

[9] More generally, given a set of three alternatives $\{A, B, C\}$, let p be the probability that A wins and q be the probability that C wins after debate. There will be no debate under pluralty rule if and only if $\forall i \in \mathcal{S}$, $u_i(B) > \frac{p}{p+q} u_i(A) + \frac{q}{p+q} u_i(C)$. It is clear that this condition is not knife-edge.

alternative in it. The upper bound on the informational efficiency of rules that designate a privileged alternative (e.g., a "status quo"), such as, for example, super-majority rules, can be shown to be lower than for the majority rule. Importantly, this effect in our analysis is due to the possibility of *any* binary agenda rather than to the fact (responsible for a parallel result due to Austen-Smith and Feddersen (2006) in the standard cheap-talk deliberative environment) that only a simple majority is needed to defeat any alternative. The intuition is as follows. A natural restriction on the agenda that is typically implied by the existence of a status quo alternative is that votes on any replacement alternatives take place first, and then the winner amongst them is pitted against the status quo - with or without the requirement that the challenger to the status quo obtain super-majority support. The effect of this restriction on the informational quality of debate is evident in the environment of Example 1 above. If $\pi = \frac{1}{2}$ is the status quo alternative, then the only possible agenda that places it last is $a_0 = \frac{1}{7}, a_1 = \frac{6}{7}, a_2 = \frac{1}{2}$. Consider then open-debate voting. Before the first debate, the median voter (4) is indifferent between $\frac{1}{7}$ and $\frac{6}{7}$, and, since $\frac{1}{7}$ and $\frac{6}{7}$ each have an equal chance of garnering the support of 5 of the 7 voters against the status quo, all the possible speakers prefer not to speak until after the vote over $\{\frac{1}{7}, \frac{6}{7}\}$. Note that this result is independent of whether a simple majority or a super-majority is required to upset the status quo.

6 Conclusion

This paper analyzes informational properties of debate under different institutional assumptions. Our key results suggest that insitutional details of the decision-making process, such single- vs. open-debate voting, the nature of the voting agenda, and the debate rules all affect the expected outcome of collective decision-making following debate. However, they do so in relatively subtle ways that point to the importance of considering interactions between them, as we have attempted to do in our analysis. While this analysis focused on majoritarian decision-making, it also suggests the value of a more focused comparison across common voting rules, which we have left to future work.

7 Appendix

7.1 Formal Definition of the Equilibrium

The equilibrium requires that at every point in the game, players' beliefs are derived from their common prior and their active arguments via Bayes Rule; i votes for the alternative $v_i^T \in \{w_{T-1}, a_T\}$ at vote T according to strategy

$$v_i^{T*} \in \arg \max_{\{w_{T-1}, a_T\}} \int_0^1 p(\hat{\theta} | r_i^{T,\infty}, l_i^{T,\infty}) \{ \Pr(w_z = v_i^T | \cdot) \hat{u}_i(v_i^T, \hat{\theta}) + \sum\nolimits_{g=T+1}^{z} [\Pr(w_z = a_g | \cdot) \hat{u}_i(a_g, \hat{\theta})] \} d\hat{\theta},$$

where $\Pr(w_z = v_i^T | \cdot)$ and $\Pr(w_z = a_g | \cdot)$ are conditioned on

$$\theta_i = \hat{\theta}, \{(r_j^{T,\infty}, l_j^{T,\infty})\}_{j \in \mathcal{R}}, \bigcup_{j=1}^{T} D^{i,\infty}, \{\{v_j^{k*}(\cdot)\}_{k=T+1}^{z}\}_{j \in \mathcal{R}}, \{\{\{D_h^{j,k*}(\cdot)\}_{k=1}^{\infty}\}_{j=T+1}^{z}\}_{h \in \mathcal{S}}),$$

and $v_j^{k*}(\cdot)$ is conditioned on

$$w_{k-1}, \{a_h\}_{h=k}^{z}, \{(r_h^{k,\infty}, l_h^{k,\infty})\}_{h \in R}, \bigcup_{h=1}^{k} D^{h,\infty}.$$

The equilibrium speaking strategy at time t in debate T is

$$D_i^{T,t*} \in \arg\max_{2^{\{1,\ldots,n\}}} [\Pr(w_z = w_{T-1}|\cdot) u_i(w_{T-1}, \theta_i) + \sum_{h=T}^{z} \Pr(w_z = a_h|\cdot) u_i(a_h, \theta_i)],$$

where $\Pr(w_z = w_{T-1}|\cdot)$ and $\Pr(w_z = a_h|\cdot)$ are conditioned on

$$\{(r_j^{T,t-1}, l_j^{T,t-1})\}_{j\in\mathcal{R}}, \bigcup_{j=1}^{T-1} D^{j,\infty}, \bigcup_{j=1}^{t-1} D^{T,j}, D_i^{T,t}, D_{-i}^{T,t*}, \{D^{T,j*}\}_{j=t+1}^{\infty}, \{D^{j,\infty*}\}_{j=T+1}^{z}, \{\{v_j^{k*}(\cdot)\}_{k=T}^{z}\}_{j\in\mathcal{R}}).$$

7.2 Proofs

Theorem 4.1

Proof. We proceed by two lemmata, which together establish this theorem.

Lemma 7.1 *Let $\Pi = \{\pi^1, \pi^2\}$ and suppose that not all argument dimensions are active for all receivers. Then, regardless of the debate rule, in every debate equilibrium the debate reveals new arguments if at least one of the following conditions does not hold:*

(a) in the initial information state, the ex ante median voter's most preferred alternative is $\frac{1}{2}$;

(b) $\{E[\theta_i|r_i^0, l_i^0]\}_{i\in\mathcal{R}}$ are symmetric around $\frac{1}{2}$; and

(c) $\{\pi^1, \pi^2\}$ is symmetric around $\frac{1}{2}$.

Otherwise, the speakers are indifferent between making arguments and not.

Proof. Because $|\Pi| = 2$, single-debate and open-debate voting are equivalent, thus we suppress the debate superscript. Let D_i^t be i's behavioral strategy at t in the debate for $i \in O^t$. Let $D^t = \bigcup_{i\in O^t} D_i^t$ and $D^\infty = \bigcup_{t=1}^{\infty} D^t$. Let w represent the winning alternative.

Consider first the case in which conditions (a), (b), and (c) in the statement of the lemma hold. From (c), the median voter is indifferent over $\{\pi^1, \pi^2\}$ and thus, in the absence of debate,

$$\Pr(w = \pi^1) = \Pr(w = \pi^2).$$

From (a), (b), and (c), for any D^∞, in expectation

$$\Pr(w = \pi^1 | D^\infty) = \Pr(w = \pi^2 | D^\infty).$$

Thus, all $i \in \mathcal{S}$ are ex ante indifferent over all possible speech in debate z.

Suppose next that condition (c) does not hold. Then the outcome of the vote over $\{\pi^1, \pi^2\}$ is certain and thus there exists $i \in \{L, R\}$ s.t. any $D_i^\infty \neq \emptyset$ dominates any $D_i^\infty = \emptyset$. If condition (c) holds, but either condition (a) or (b) does not, then

$$\Pr(w = \pi^1 | D^\infty) \neq \Pr(w = \pi^2 | D^\infty).$$

if $D^\infty \neq \emptyset$. Thus, there exists $i \in \{L, R\}$ s.t. any $D_i^\infty \neq \emptyset$ dominates any $D_i^\infty = \emptyset$.

Thus, if at least one of the conditions in the statement of the lemma is violated, then in any equilibrium, new arguments are revealed in debate. ■

Theorem 4.2

We begin with the following instrumental lemma:

Lemma 7.2 *If there is full revelation in* $(j + 1)$*, then regardless of the debate rule, under open debate voting there is full revelation in debate* j *if under* $\bigcup_{k=1}^{j-1} D^{k,\infty}$ *the median voter has strict preference over* $\{w_{j-1}, a_j\}$*.*

Proof. We proceed in a sequence of two claims:

Claim 1. $\exists i \in \mathcal{S}$ s.t. i prefers in expectation $D^{j,\infty} \supseteq \{1,...,n\} \setminus \bigcup_{k=1}^{j-1} D^{k,\infty}$ to any $D^{j,\infty} \subseteq \bigcup_{k=1}^{j-1} D^{k,\infty}$.

Proof of Claim 1. Suppose that the ex ante median voter under $\bigcup_{k=1}^{j-1} D^{k,\infty}$ is not indifferent between a_j and w_{j-1}. We consider an exhaustive set of cases.

Suppose that the ex ante median voter prefers w_{j-1} to a_j. (The case where she prefers a_j is symmetric.)

1. Suppose a_j such that $\min\{w_{j-1}, a_{j+1}, ..., a_z\} < a_j < \max\{w_{j-1}, a_{j+1}, ..., a_z\}$. Let

$$\begin{aligned} x &= \max\{\pi : \pi \in \{w_{j-1}, a_j, a_{j+1}, ..., a_z\} \text{ and } \pi < a_j\}, \text{ and} \\ y &= \min\{\pi : \pi \in \{w_{j-1}, a_j, a_{j+1}, ..., a_z\} \text{ and } \pi > a_j\}. \end{aligned}$$

Let $\pi_m(A)$ be the full-information median voter's most preferred alternative in $A \subseteq \Pi$, and define

$$\begin{aligned} p = &\Pr(\pi_m(\{w_{j-1}, a_{j+1}, a_{j+2}, ..., a_z\}) = x | D^0) \\ &- \Pr(\pi_m(\{w_{j-1}, a_j, a_{j+1}, ..., a_z\}) = x | D^0) \end{aligned}$$

and

$$\begin{aligned} q = &\Pr(\pi_m(\{w_{j-1}, a_{j+1}, a_{j+2}, ..., a_z\}) = y | D^0) \\ &- \Pr(\pi_m(\{w_{j-1}, a_j, a_{j+1}, ..., a_z\}) = y | D^0). \end{aligned}$$

For speaker $i \in \mathcal{S}$, revelation in j is preferred iff

$$p\hat{u}_i(x) + q\hat{u}_i(y) \leq (p+q)\hat{u}_i(a_j).$$

Dividing through by $(p+q)$, we obtain

$$\frac{p}{p+q}\hat{u}_i(x) + \frac{q}{p+q}\hat{u}_i(y) \leq \hat{u}_i(a_j). \tag{2}$$

We know from the concavity of $\hat{u}_i(\pi)$ that

$$\frac{p}{p+q}\hat{u}_i(x) + \frac{q}{p+q}\hat{u}_i(y) < \hat{u}_i(\frac{p}{p+q}x + \frac{q}{p+q}y). \tag{3}$$

Define t s.t.

$$tx + (1-t)y = a_j. \tag{4}$$

There are three exhaustive possibilities:

(a) $t = \frac{p}{p+q}$. Then

$$\hat{u}_i(\frac{p}{p+q}x + \frac{q}{p+q}y) = \hat{u}_i(a_j)$$

and so from (2), it follows that for all $i \in \mathcal{S}$, i prefers full revelation in j to no revelation in j, given full revelation in $j+1$.

(b) $t < \frac{p}{p+q}$. Then $x < y$ implies

$$tx + (1-t)y > \frac{p}{p+q}x + \frac{q}{p+q}y.$$

Given that $\frac{\partial \hat{u}_R}{\partial \pi} > 0$,

$$\hat{u}_R(\frac{p}{p+q}x + \frac{q}{p+q}y) < \hat{u}_R(tx + (1-t)y).$$

Combining this inequality with (3) and (4), we have from (4.1) that $i = R$ prefers full revelation in j to no revelation in j, given full revelation in $j+1$.

(c) $t > \frac{p}{p+q}$. Then by a symmetric argument, $i = L$ prefers full revelation in j to no revelation in j, given full revelation in $j+1$.

2. Suppose $a_j = \max\{w_{j-1}, a_j, a_{j+1}, ..., a_z\}$. Then $R \in \mathcal{S}$ must prefer full revelation in j to delay until $(j+1)$.

3. Suppose $a_j = \min\{w_{j-1}, a_j, a_{j+1}, ..., a_z\}$. Then $L \in \mathcal{S}$ must prefer full revelation in j to delay until $(j+1)$.

Claim 2. Full revelation occurs in j for any O s.t. $\{L, R\} \subseteq \bigcup\limits_{t=1}^{o} O^t$.

Proof of Claim 2. Consider first debates with endogenous termination rules. From Claim 1, $\exists i \in \mathcal{S}$ and $h \in \{1, ..., o\}$ s.t. $i \in O^h$ and $D_i^{j,h}$ s.t. $D_i^{j,h} \backslash \bigcup\limits_{k=1}^{j-1} D^{k,\infty} \neq \emptyset$. Also, from Claim 1, if $\bigcup\limits_{g=1}^{h} D^{i,g} \nsupseteq \{1, ..., n\} \backslash \bigcup\limits_{k=1}^{j-1} D^{k,\infty}$, then under $(\bigcup\limits_{g=1}^{h} D^{i,g}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty})$ $\exists i' \in \mathcal{S}$, $h' \in \{h+1, ..., o\}$ s.t. $i' \in O^{h'}$ and $D_{i'}^{j,h'}$ s.t. $D_{i'}^{j,h'} \backslash (\bigcup\limits_{g=1}^{h} D^{j,g}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty}) \neq \emptyset$. By recursion, new arguments are revelaed until $D^{j,\infty} \supseteq \{1, .., n\} \backslash \bigcup\limits_{k=1}^{j-1} D^{k,\infty}$. Note that, again from Claim 1, anticipation of $D^{j,\infty} \supseteq \{1, ..., n\} \backslash \bigcup\limits_{k=1}^{j-1} D^{k,\infty}$ is consistent with the preference for speech in that the speaker prefers full revelation to no revelation in expectation at the time of her speech.

Suppose next an exogenous termination rule and some speaker order O. From Claim 1, for any $\bigcup_{h=1}^{o-1} D^{j,h}$ s.t. $\bigcup_{h=1}^{o-1} D^{j,h} \not\supseteq \{1,...,n\}\backslash \bigcup_{h=1}^{j-1} D^{h,\infty}$, $\exists i \in \mathcal{S}$ s.t. i prefers in expectation under $(\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty})$, $D^{j,o} \supseteq \{1,...,n\}\backslash((\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty}))$ to any $D^{j,o} \subseteq (\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty})$. Because i realizes benefits from the revelation of new arguments in j iff such revelation changes w_j, and because the probability of such change is increasing in $|D^{j,o}\backslash((\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty}))|$, i also prefers in expectation $D^{j,o} \supseteq \{1,...,n\}\backslash((\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty}))$ to any $D^{j,o} \not\supseteq ((\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty}))$. Thus, if $i \in O^o$, then $D^{j,o} \supseteq \{1,...,n\}\backslash((\bigcup_{h=1}^{o-1} D^{j,h}) \cup (\bigcup_{h=1}^{j-1} D^{h,\infty}))$.

Suppose $i \notin O^o$. Let $h \equiv \max\{h' : h' \in \{1,...,o\}$ and $i \in O^{h'}\}$. There are two possibilities:

(a) Suppose under $(\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty})$ i prefers, in expectation, any $D^{j,\infty} \supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$ to any $D^{j,\infty} \subseteq ((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$. Then i also prefers in expectation $D^{j,\infty} \supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$ to any $D^{j,\infty} \not\supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$. Thus, $D^{j,h} \supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$.

(b) Suppose under $(\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty})$ i prefers, in expectation, any $D^{j,\infty} \subseteq ((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$ to any $D^{j,\infty} \supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$. Then from Claim 1, it must be that $\exists i' \in \mathcal{S}\backslash\{i\}$ s.t. i prefers, in expectation under $(\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty})$ any $D^{j,\infty} \supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$ to any $D^{j,\infty} \subseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$. Suppose $\exists h' \in \{h, h+1, ..., o\}$ s.t. $i' \in O^{h'}$; then $D_i^{j,h} \subseteq ((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$ dominates any $D_i^{j,h} \not\subseteq ((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$ for i and $D_{i'}^{j,h'} \supseteq \{1,...,n\}\backslash((\bigcup_{k=1}^{h-1} D^{j,k}) \cup (\bigcup_{k=1}^{j-1} D^{k,\infty}))$. Suppose instead $|\{h' : h' \in \{h, h+1, ..., o\}$ and $i' \in O^{h'}\}| \geq 2$. If $D_i^{j,h} \not\supseteq \{1, ..., n\} \backslash ((\bigcup\limits_{k=1}^{h-1} D^{j,k}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty}))$, then $\{D_{i'}^{j,k}\}_{k=h}^{o}$ is such that either $\bigcup\limits_{k=h}^{o} D_{i'}^{j,k} \supseteq \{1, ..., n\} \backslash (D_i^{j,h} \cup (\bigcup\limits_{k=1}^{h-1} D^{j,k}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty}))$ or $\exists h'' \in \{h+1, ..., o\}$ s.t. under $(\bigcup\limits_{k=1}^{h''-1} D^{j,k}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty}))$, in expectation, i prefers any $D^{j,\infty} \subseteq \{1, ..., n\} \backslash (\bigcup\limits_{k=1}^{j-1} D^{k,\infty})$ to $D^{j,\infty} = (\bigcup\limits_{k=1}^{h''-1} D^{j,k})$ and i' prefers $D^{j,\infty} = (\bigcup\limits_{k=1}^{h''-1} D^{j,k})$ to any $D^{j,\infty}$ s.t. $D^{j,\infty} \backslash ((\bigcup\limits_{k=1}^{h''-1} D^{j,k}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty})) \neq \emptyset$. But this implies that i prefers $D_i^{j,h} \supseteq \{1, ..., n\} \backslash ((\bigcup\limits_{k=1}^{h-1} D^{j,k}) \cup (\bigcup\limits_{k=1}^{j-1} D^{k,\infty}))$. ■

We next complete the proof of Theorem 4.2, proving in sequence parts 1, 3, and then 2.

Proof. 1. The information state at the beginning of the z^{th} debate is characterized by $\bigcup\limits_{j=1}^{z-1} D^{j,\infty}$. If $\bigcup\limits_{j=1}^{z-1} D^{j,\infty} = \{1, ..., n\}$, then full revelation has occurred before debate z. Suppose then that $\bigcup\limits_{j=1}^{z-1} D^{j,\infty} \subset \{1, ..., n\}$. So, treating $\bigcup\limits_{j=1}^{z-1} D^{j,\infty}$ as the initial information state and applying Lemma 7.1, we obtain that revelation of new arguments occurs in debate z if conditions (a)-(c) in Lemma 7.1 do not hold. Suppose next that in z, these conditions do hold. Then, similarly to the proof of Lemma 7.1, all $i \in \mathcal{S}$ are ex ante indifferent over all possible speech in debate z. Recall that w_{z-1} is the outcome of the vote over $\{w_{z-2}, a_{z-1}\}$ given information $\bigcup\limits_{j=1}^{z-1} D^{j,\infty}$. Suppose that $w_{z-1} = w_{z-2}$; the case of $w_{z-1} = a_{z-1}$ is symmetric. Given that the median voter under $\bigcup\limits_{j=1}^{z-1} D^{j,\infty}$ is indifferent over $\{w_{z-1}, a_z\}$ and prefers $w_{z-1} = w_{z-2}$ to a_{z-1}, $a_{z-1} \notin [w_{z-2}, a_z]$. If $a_{z-1} < \min\{w_{z-2}, a_z\}$, then, given that all $i \in \mathcal{S}$ are indifferent over speech when the set of alternatives is $\{w_{z-2}, a_z\}$, L must strictly prefer additional revelation in debate $(z-1)$ when the set of effective alternatives is $\{a_{z-1}, w_{z-2}, a_z\}$. If, instead, $a_{z-1} > \min\{w_{z-2}, a_z\}$, then, by symmetry, R prefers additional revelation in debate $(z-1)$. Thus, absence of full revelation prior to debate z is not consistent with the expectation that all three conditions in Lemma 7.1 hold in z. It follows that either full revelation occurs before debate z or, in expectation at $z-1$, full revelation occurs in debate z. From Lemma 7.2, it follows by recursion that the expectation in the initial information state is that full revelation will occur.

3. Part 1 of this theorem implies that, in expectation, there is full revelation in the last debate if not before it. A necessary condition for delay is that there exist $x, y \in \Pi$ s.t. the median voter under D^0 is indifferent over $\{x, y\}$ and, for some $j < z$, $\{w_{j-1}, a_j\} = \{x, y\}$ on the equilibrium path of play under D^0. Because w_{j-1} must be the ex ante median voter's most preferred alternative in $\{a_0, a_1, ..., a_{j-1}\}$, $w_{j-1} \in \{x, y\}$ only if there exists no $a_k \in \{a_0, a_1, ..., a_{j-1}\}$ s.t. the ex ante median voter prefers a_k to x (i.e., $a_k \notin [x, y]$).

2. Suppose A consists of all sets of two distinct elements $x, y \in \Pi$ s.t. the ex ante median voter is indifferent over $\{x, y\}$. Part 3 establishes that any agenda induces full revelation without delay for $A = \emptyset$. For $|A| = 1$.

Suppose $|A| \geq 2$. There exists at most one $\{x, y\} \in A$ s.t. $\nexists \pi \in \Pi$ s.t. $x < \pi < y$ or $y < \pi < x$. If such a pair exists, let $a_z \in \{x, y\}$ and let $a_0 = \{x, y\} \backslash \{a_z\}$. If such a pair does not exist, let a_0 be the ex ante median voter's expected most preferred alternative in Π. Then part 3b establishes that debate is fully revealing without delay. ■

Theorem 4.3

Proof. For every information state, the preference profile is single-peaked. Thus, in every information state, there exists a Condorcet winner. Thus, holding fixed the information state, the voting outcome is invariant with respect to the voting agenda. Because, by definition, in single-debate voting, all debate takes place before voting starts, the information state is fixed during the voting stage. Because all agendas are then strategically equivalent, debate incentives are also invariant with respect to the voting agenda. In contrast, in open-debate voting, the information state is not fixed throughout the voting stage, and debate incentives depend on the presence of particular alternatives in the set of possible outcomes at a time j. Consequently, agendas need not be strategically equivalent. Example 1 establishes that the outcome under open-debate voting is sensitive to the voting agenda.

■

Theorem 4.4

Proof. From Theorem 4.2, there exists an agenda that induces fully revealing debate without delay under open-debate voting regardless of debate rule. Choose such an agenda. From Theorem 4.3, with respect to single-debate voting, the agenda is irrelevant and the equilibrium informativeness is greatest if the right debate rule is chosen. Fix that debate rule. This pair of agenda and debate rule induces the most informative equilibrium debate in both the open-debate case and the single-debate case. For the open-debate case, we have just established that equilibrium debate is fully informative without delay. To see that even with the best debate rule (and so, for any debate rule), single-debate voting is sometimes not fully informative, suppose that in Example 1 there is a single debate before the first vote. If the agenda is, once again, $a_0 = \frac{1}{7}, a_1 = \frac{1}{2}, a_2 = \frac{6}{7}$, then restriction to single-debate voting eliminates all informative argumentation. Each speaker prefers $\frac{1}{2}$ with certainty to the lottery of $\frac{1}{7}$ with probability p, $\frac{6}{7}$ with probability p, and $\frac{1}{2}$ with probability $1 - 2p$. Thus, regardless of the debate rule, $\forall p \in (0, \frac{1}{2})$, there is no revelation under single-debate voting. ■

References

[1] Austen-Smith, David. 1993. "Interested Experts and Policy Advice: Multiple Referrals Under Open Rule." *Games and Economic Behavior* 5, 3-43.

[2] Austen-Smith, David, and Timothy Feddersen. 2006. "Deliberation, Preference Uncertainty, and Voting Rules." *American Political Science Review* 100 (2), 209-217.

[3] Battaglini, Marco. 2002. "Multiple Referrals and Multidimensional Cheap-talk." *Econometrica* 70 (4), 1379-1401.

[4] Crawford, Vincent and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50, 1431-51.

[5] Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini. 2001. "Standard State Space Models Preclude Unawareness." *Econometrica* 66 (1), 159-173.

[6] Dickson, Eric S., Catherine Hafer, and Dimitri Landa. 2008. "Cognition and Strategy: a Deliberation Experiment." *Journal of Politics* 70 (4), 2008.

[7] Gerardi, Dino and Leeat Yariv. 2006. "Deliberative Voting." *Journal of Economic Theory* 134, 317-338.

[8] Glazer, Jacob and Ariel Rubinstein. 2001. "Debates and Decisions: on a Rationale for Argumentation Rules." *Games and Economic Behavior* 36, 158-73.

[9] Glazer, Jacob and Ariel Rubinstein. 2006. "On the Pragmatics of Persuasion: a Game Theoretical Approach." *Theoretical Economics* 1 (4).

[10] Hafer, Catherine and Dimitri Landa. 2007. "Deliberation as Self-Discovery and Institutions for Political Speech." *Journal of Theoretical Politics* 19 (3).

[11] Hafer, Catherine and Dimitri Landa. 2006. "Deliberation and Social Polarization." New York University Mimeo.

[12] Heifetz, Aviad, Martin Meier and Burkhard C. Schipper. 2006. "Interactive Unawareness." *Journal of Economic Theory* 130, 78-94.

[13] Krishna, Vijay and John Morgan. 2001. "A Model of Expertise." *Quarterly Journal of Economics* 116, 747-775.

[14] Lanzi, Thomas and Jerome Mathis. 2004. "Argumentation in Sender-Receiver Games." Mimeographed.

[15] Li, Jing. 2006. "Information Structures with Unawareness." Duke University Mimeo.

[16] Lipman, Barton and Duane J. Seppi. 1995. "Robust Inference in Communication Games with Partial Provability." *Journal of Economic Theory* 66 (2), 370-405.

[17] McKelvey, Richard and Richard Niemi. 1978. *Journal of Economic Theory* 18, 1-22.

[18] Meirowitz, Adam. 2007. "In Defense of Exclusionary Deliberation: Communication and Voting with Private Beliefs and Values." *Journal of Theoretical Politics* 19 (3).

[19] Patty, John. forthcoming. "Arguments-Based Collective Choice." *Journal of Theoretical Politics*.