Persuasion and Information Aggregation in Elections *

Carl Heese † Stephan Lauermann ‡

This paper studies a large majority election with voters who have heterogeneous, private preferences and exogenous private information about an unknown state of the world. We show that a Bayesian persuader can achieve any state-contingent outcome in some equilibrium by providing additional information. In this setting, without the persuader's additional information, a version of the Condorcet jury theorem holds, in the sense that outcomes of large elections satisfy full-information equivalence (Feddersen and Pesendorfer, 1997). Persuasion does not require detailed knowledge of the voters' private information, preferences, or the voting rule. It also requires almost no commitment power on the part of the persuader.

In most elections, a voter's ranking of outcomes depends on her information. For example, a shareholder's view of a proposed merger depends on her beliefs regarding its profitability, and a legislator's support of proposed legislation depends on her beliefs regarding its effectiveness. An interested party that has private information may utilize this dependence by strategically releasing information to affect voters' behavior. For instance, before shareholders vote on a merger, the management may provide strategically chosen information about it through presentations and conversations; similarly, lobbyists may provide selected information to legislators to influence their votes.

We are interested in the scope of such "persuasion" (Kamenica and Gentzkow, 2011) in elections. We study this question in the canonical voting setting of

*May 20, 2024; We are grateful for helpful discussions with Ricardo Alonso, Nageeb Ali, Arjada Bardhi, Dirk Bergemann, Sourav Bhattacharya, Francesc Dilmé, Mehmet Ekmekci, Erik Eyster, Tim Feddersen, Yingni Guo, Matt Jackson, Daniel Krähmer, Elliot Lipnowski, Antonio Penta, Jacopo Perego, Keith Schnakenberg, and Thomas Tröger, as well as comments from audiences at Oxford, Bonn, Yale (lunch), the LSE (lunch), ESWM 2017, SAET 2018, ESEM 2018, the ASSA 2019 meeting in Atlanta, the annual Wallis Institute conference, and other venues. This work was supported by a grant from the European Research Council (ERC 638115), and by the Deutsche Forschungsgemeinschaft (German Research Foundation) under Germany's Excellence Strategy EXC 2126/1-390838866 and the CRC TR 224 (Projects B03 and B04).

† The University of Hong Kong, Faculty of Business and Economics, heese@hku.hk.

‡ University of Bonn, Department of Economics, s.lauermann@uni-bonn.de.

Feddersen and Pesendorfer (1997): There are two possible policies—A and B. Voters' preferences over policies are heterogeneous and depend on an unknown state, α or β. Some voters may prefer A in state α and B in β, with heterogeneous "thresholds of doubt" when they are uncertain about the state, while others are "partisans" who prefer one policy or the other, independently of the state. These preferences are drawn independently across voters and are private information. In addition, all voters privately receive information about the state in the form of a conditionally independent, noisy signal. The election determines the policy by simple majority rule.

For this setting, the Condorcet jury theorem holds: All outcomes of symmetric equilibria of large elections are equivalent to the outcome with a known state ("information aggregation"); see Feddersen and Pesendorfer (1997) and Bhattacharya (2013). We restate this benchmark in Theorem 0, then ask the following question: Can an interested party change that outcome and instead induce a majority for his favorite policy in each state by strategically providing *additional* information to the voters? We study this question using Bayesian persuasion (Kamenica and Gentzkow, 2011), allowing the interested party (sender) to choose any joint distribution over states and messages, and the latter are then privately observed by the voters.

The Condorcet jury theorem may suggest that the scope for persuasion is limited. Indeed, it turns out that the Condorcet jury theorem directly implies that in a large election, any additional information that is either public to all voters or conditionally independent across voters will have no effect on the outcome. So, to have any hope of changing the outcome, the sender must communicate privately with correlated signals. How much can he achieve by doing so?

Our main result (Theorem 1) shows that persuasion can achieve anything: For any given state-dependent target policy, a sender can provide additional information in such a way that a majority will support the target policy with probability close to 1 when the number of voters is large. For example, just by providing additional information, a sender can bring about the opposite of the full-information outcome in every state.

Theorem 1 further highlights the unrestricted scope of persuasion by showing that the same information structure is effective *uniformly* across all parameters of the Condorcet jury setting. In particular, the sender does not need to tailor the information structure to the prior, the private information of individual voters, or the distribution of the voters' private information. The persuasion mechanism is “detail-free” and passes the Wilson doctrine (Wilson, 1987).

The additional information affects the voters' behavior in two ways: directly, by changing their beliefs about the state, and indirectly, by affecting their inference about being "pivotal" for the election outcome. While the direct effect is limited by the well-known Bayesian consistency requirement for beliefs, the pivotal inference turns out to have no such constraint.

Our proofs are constructive and reveal a basic persuasion mechanism. To invert the full-information outcome, the sender can choose an information structure of uncertain quality: With probability $1-\varepsilon$, the information is of "high quality," which means all voters receive a message equal to the state, a in state α and b in state β. Thus, when $\varepsilon = 0$, the election leads to the full-information outcome. However, with probability $\varepsilon > 0$, the information is of "low quality." In this case, in either state, almost all voters receive a message z that is uninformative about the state, while a few voters receive a message that is the opposite of the state; that is, they receive a in β and b in α. Hence, in this situation, a and b carry the opposite of their original meanings.

What makes persuasion effective is that voters interpret their messages in light of the closeness of the election. The closeness of the election tells each voter something about the quality of the other voters' information and thus also reveals something about the quality of her own message. The intuition here is that low-quality information gives rise to more disagreement than high-quality information. Indeed, in the equilibrium that we construct, a close election implies that the information is of low quality, meaning that almost all voters have received message z; therefore, a voter receiving message a will interpret it as being in favor of β rather than α, and vice versa for b.

The fundamental aspect of the information structure is the uncertainty regarding the quality of the others' information. This uncertainty is ubiquitous in real-world settings and may emerge naturally, even without the intentional choice of a particular sender. For instance, voters face uncertainties regarding media bias and the quality of information on social media platforms. Consider encountering a dubious Facebook post. The primary concern is not only to recognize it as misinformation but also to assess its potential influence on the beliefs of other voters and their electoral decisions. Our results illustrate how these strategic considerations can shape electoral outcomes and disrupt information aggregation.[1]

[1]Similar considerations may also be relevant for other political contexts, such as (non-binding) shareholder voting, polls, and protests, which have been studied from the perspective of information aggregation and using variations of the Condorcet model; see Levit and Malenko

Regarding the robustness of our theoretical findings, we noted earlier that the sender's information structure does not need to be tailored to the details of the environment; it is uniformly effective. The proof sketch illustrates the essential idea. The central observation is Lemma 1. It shows that, for every prior and every distribution of the voter's private information, the election will be closer in equilibrium when voters receive very low-quality information than when they have full information about the state. The lemma rests on the voters endogenously adjusting their behavior to environmental changes and is shown to imply that no changes in the sender's information structure itself are required.

Section 3 further discusses the robustness of the persuasion logic by exploring model variations. We start by considering typical features of voting scenarios that are not present in the basic model. In Section 3.1 we consider more general voter preferences, allowing groups of voters to have opposing preferences depending on the state. Such preferences have been used to study distributive politics, in which the state determines which voters will favor each policy.[2] In Section 3.3 we consider participation costs as in Krishna and Morgan (2012). With participation costs, the sender faces an additional challenge since he not only needs to influence what voters vote for but also needs to persuade them to turn out in the first place. In Section 3.2 we consider supermajority rules and in Section 3.4 a state space that is not binary but a continuum, both as in Feddersen and Pesendorfer (1997).

Next we address some common concerns regarding our method, especially about the sender's commitment power and the sender's access to full information about the state. Section 3.5 shows that the sender needs only minimal commitment power. Section 3.6 discusses the case where the sender has only partial information about the state. Two numerical examples in Section 3.7 illustrate the persuasion mechanism further: Persuasion can be quite effective even with relatively small voter numbers, and persuasion remains effective when a large share of voters behaves "sincerely;" thus, we can weaken the sometimes controversial assumption of pivotal voting.

Persuasion of voters has been studied before under the assumption that the sender is "omniscient"—meaning that the sender has complete knowledge of all the voters' types; see Alonso and Câmara (2016), Bardhi and Guo (2018), and Chan, Gupta, Li, and Wang (2019). The classical Condorcet setting of Feddersen and Pesendorfer (1997), in which voters have private preferences and signals, captures scenarios in which the sender has much more limited control over information.

(2011), Morgan and Stocken (2008), and Battaglini (2017) as examples.

[2] See Fernandez and Rodrik (1991), Ali, Mihm, and Siga (2018), and Bhattacharya (2018).

Section 3.8 discusses persuasion with an omniscient sender and how persuasion in the classical setting is distinct and requires a different approach. Finally, Section 4 discusses the paper's relationship to the general literature on persuasion and information aggregation in elections.

We note two broader implications of our analysis. First, making robust predictions about election results may be difficult. If an outside observer knows that voters have access to at least the information assumed in Feddersen and Pesendorfer (1997) but cannot exclude the possibility that they have access to additional information, then no outcome can be excluded as an equilibrium prediction. Second, if one interprets an information structure with a small ε as a small departure from common knowledge, our result adds another observation to the literature on the effects of strategic uncertainty (Weinstein and Yildiz, 2007).

1 Model

There are $2n+1$ voters ("she"), two policies, A and B, and two states of the world, $\omega \in \{\alpha, \beta\}$. The voters share a common prior $p_0 = \Pr(\alpha) \in (0, 1)$.

The preferences of any given voter are characterized by a "threshold of doubt" $y \in [0, 1]$ as follows: A voter of type y prefers policy A if she believes the probability of state α exceeds y.[3] Types are distributed independently across voters (and independently of the state), with full support on $[0, 1]$. The cumulative distribution function is denoted by Φ and assumed to be strictly increasing and continuous on $[0, 1)$, with a strictly positive mass ("atoms") at 0 and 1. The types with thresholds at 0 and 1 are "partisans" who prefer A and B, respectively, independently of their belief; see Figure 1 for an illustration.

Each voter $i \in \{1, \ldots, 2n+1\}$ receives a two-dimensional signal (s_i, m_i) from the finite set $S \times M$, with $\mathbf{s} = (s_1, \ldots, s_{2n+1})$ and $\mathbf{m} = (m_1, \ldots, m_{2n+1})$. Conditional on the state, the signals from the first dimension are independent and identically distributed (i.i.d.) across voters; that is,

$$\Pr(\mathbf{s}|\omega) = \prod_{i=1,\ldots,2n+1} \Pr(s_i|\omega), \tag{1}$$

[3] Here is a simple formulation in terms of payoff types $\hat{y} \in \mathbb{R}$: For a voter with type $\hat{y}$, the payoff from A is $1 - \hat{y}$ in α and $-\hat{y}$ in β and the payoff from B is 0 in both states. With this specification, a voter prefers A whenever she believes the probability of α to be above $\hat{y}$. Types with $\hat{y} \le 0$ and $\hat{y} \ge 1$ are "partisans" who prefer A and B, respectively, independently of their beliefs. An atomless distribution of $\hat{y}$ with $\mathbb{R}$ as its support induces a distribution of thresholds of doubt $y \in [0, 1]$ (with atoms at 0 and 1) via $y = \max\{\min\{\hat{y}, 1\}, 0\}$. Section 3.1 discusses general payoff types.

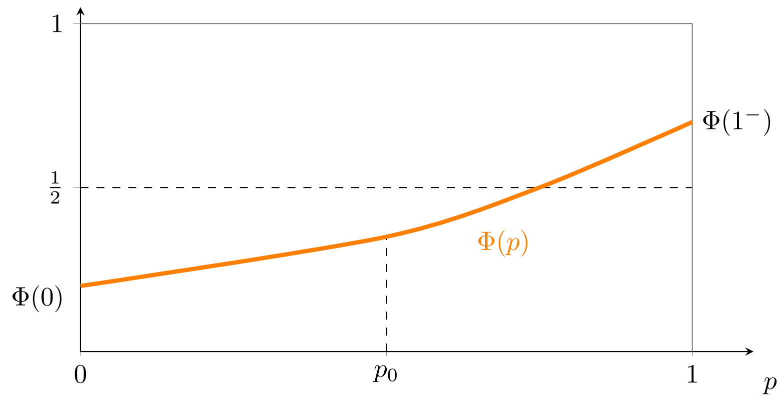

Figure 1: $\Phi(p)$ is the ex-ante probability that a voter prefers A if she believes the probability of α is $p \in (0,1)$. The presence of partisans for A and B implies that $0 < \Phi(0)$ and $\Phi(1^-) < 1$, respectively. At the depicted prior p_0 there is in expectation a majority for B.

and $\Pr(s_i|\omega)$ for $s_i \in S$ does not depend on i. The signals are boundedly informative:

$$0 < \min_{s \in S} \frac{\Pr(s|\alpha)}{\Pr(s|\beta)} < \max_{s \in S} \frac{\Pr(s|\alpha)}{\Pr(s|\beta)} < \infty. \tag{2}$$

There is no restriction on the distribution of the signals from the second dimension, except that it is exchangeable across voters[4] and independent from the first: $\Pr(\mathbf{s}, \mathbf{m}|\omega) = \Pr(\mathbf{s}|\omega) \cdot \Pr(\mathbf{m}|\omega)$.

The first dimension, $\mathbf{s}$, represents the exogenous information that the voters obtain privately as in Feddersen and Pesendorfer (1997); we refer to it informally as "nature's signal." The second dimension, $\mathbf{m}$, represents additional information for the voters—for example, information provided by an interested sender. We refer to it as the "sender's signal," and to its realized signals often as "messages." The independence of $\mathbf{s}$ and $\mathbf{m}$ reflects the inability of the sender to condition on the voters' private information; apart from this, there is no constraint on the sender's signal. Let $\pi^S := \{\Pr(s|\omega)\}_{s \in S, \omega \in \{\alpha,\beta\}}$ and $\pi^M := \{\Pr(\mathbf{m}|\omega)\}_{\mathbf{m} \in M^{2n+1}, \omega \in \{\alpha,\beta\}}$; in the parlance of Bergemann and Morris (2016), $\pi := \pi^S \times \pi^M$ corresponds to an independent expansion of π^S.

The voting game is as follows: First, nature draws the state ω, the preference types according to Φ, and the signal profile $(\mathbf{s}, \mathbf{m})$ according to π. Second, each voter simultaneously submits a vote for A or B after observing her type and signal. Finally, the submitted votes are counted and the majority outcome is selected. This defines a Bayesian game.

[4]This means $\Pr(\mathbf{m}|\omega) = \Pr(\mathbf{m}'|\omega)$ when $\mathbf{m}$ is a permutation of $\mathbf{m}'$.

The game is symmetric, and so we restrict our attention to symmetric strategies $\sigma : S \times M \times [0,1] \to [0,1]$, where $\sigma(s,m,y)$ is the probability that a voter with signal (s,m) and type y votes for A.

We consider only weakly undominated strategies. In particular, voters with types $y \in \{0,1\}$ (partisans) always vote for their preferred policy:

$$\sigma(s,m,0) = 1 \quad \text{and} \quad \sigma(s,m,1) = 0 \quad \forall (s,m). \tag{3}$$

Since there is a strictly positive probability that some A- or B-partisans are present, this rules out "non-responsive" equilibria in which all voters support the same alternative with probability 1. The additional richness assumption

$$\Phi(0) < 1/2 < \Phi(1^-) \tag{4}$$

excludes the trivial scenarios with a majority of A-partisans or B-partisans in expectation.

The best response of the non-partisans is simple: From the viewpoint of a given voter, the pivotal event piv is the one in which the realized types and signals of the other $2n$ voters are such that n of them vote for A and n for B. In this event, her vote determines the outcome; in any other event, the outcome is independent of her vote. Thus, a strategy is optimal if and only if it is optimal conditional on the pivotal event.

Let $\Pr(\alpha|s,m,\text{piv};\sigma')$ denote the posterior probability of α conditional on (s,m) and on the pivotal event, assuming the other voters follow some strategy σ'. The strategy σ is a best response to σ' if

$$\Pr(\alpha|s,m,\text{piv};\sigma') > y \Rightarrow \sigma(s,m,y) = 1 \tag{5}$$

$$\Pr(\alpha|s,m,\text{piv};\sigma') < y \Rightarrow \sigma(s,m,y) = 0. \tag{6}$$

Note that the set of indifferent types has zero measure. It follows that there is no loss of generality in considering pure strategies with $\sigma(s,m,t) \in \{0,1\}$ for all (s,m,t), which we will do from now on.

A symmetric, undominated, and pure Bayesian Nash equilibrium is a strategy $\sigma : S \times M \times [0,1] \to \{0,1\}$ that satisfies (3), (5), and (6) with $\sigma' = \sigma$. Hereafter this is referred to simply as an *equilibrium*.

2 Main Result

2.1 Benchmark: The "Modern" Condorcet Jury Theorem

When the sender's signal π^M is uninformative,[5] our model reduces to a special case of the canonical voting game of Feddersen and Pesendorfer (1997), with a binary state. We restate the classic result as a benchmark: Large elections lead to outcomes that are equivalent to those under full information; that is, for n large, with probability close to 1, a majority votes for A in α and B in β.

Theorem 0 *Let π^M be uninformative and $(\sigma_n^*)_{n\in\mathbb{N}}$ a sequence of equilibria with $2n+1$ voters. Then, the full information outcome is elected as $n \to \infty$,*

$$\lim_{n\to\infty} \Pr\left(A \textit{ wins} \mid \alpha; \sigma_n^*, n\right) = 1 \textit{ and } \lim_{n\to\infty} \Pr\left(B \textit{ wins} \mid \beta; \sigma_n^*, n\right) = 1.$$

The standard proof of Theorem 0 is included in the online appendix, Section F.[6]

In our context, the "modern"[7] Condorcet jury theorem states that when the number of voters is large and their only information is nature's conditionally i.i.d. signal, all equilibrium outcomes are close to the full-information outcome. Thus, to achieve any policy other than the full-information outcome, the sender has to communicate with the voters in some way. How should he do so? What outcomes can he achieve by releasing additional information?

Consider a sender who communicates via *public messages*, meaning that the messages are commonly received by all the voters. When the voters receive a public message m, this shifts their common belief from the prior p_0 to $\Pr(\alpha|m)$. However, given the Condorcet jury theorem, in any subgame following a public message for which $\Pr(\alpha|m)$ is interior, the full-information outcome is attained with probability converging to 1 as $n \to \infty$ (and if the public signal reveals the state, the full-information outcome obtains simply by definition). The Condorcet jury theorem also implies that it is ineffective to communicate via messages that are independent from each other conditional on the state. Therefore, to achieve

[5]For example, $\Pr(\mathbf{m}|\omega;\pi^M) = 1$ for some $\mathbf{m} \in M^{2n+1}$. For the statement of Theorem 0, recall the assumptions we maintain on signals and preferences, especially (1) and (4).

[6]The model of Feddersen and Pesendorfer (1997) is more general than our model. They assume the existence of sub-populations and allow the signal distributions to vary across these; this is not critical. Moreover, they assume a continuum of states ω; see also our discussion in Section 3.4. The binary-state version here is a special case of the model in Bhattacharya (2013).

[7]The "classical" version of the Condorcet jury theorem assumes so-called sincere voting behavior and common interests. It requires the additional condition that signals are sufficiently precise relative to the prior, and it is less robust to the presence of partisans.

an outcome different from the full-information outcome, the sender must communicate privately with the voters and must correlate their messages.

2.2 Main Result: Full Persuasion

We now characterize the policies that a strategic sender can achieve in an equilibrium of a large election by providing additional information to the voters, given the information already provided by nature's signal.

Our main result is that there is no constraint on the set of achievable policies: For any state-dependent policy—even the policy that inverts the full-information outcome—the sender can release messages so that there is an equilibrium σ_n for which in each state the corresponding target policy wins with probability close to 1 when n is large. This holds for all prior beliefs p_0, preference distributions Φ, and nature's signals π^S that are admissible given our assumptions.[8]

Theorem 1 *For any state-dependent policy* $(x(\alpha), x(\beta)) \in \{A, B\}^2$*, there exists a sequence of sender's signals* $(\pi_n^M)_{n\in\mathbb{N}}$ *such that, for any admissible prior belief* p_0*, preference distribution* Φ*, and nature's signal* π^S*, there is an equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *given* $\pi_n = \pi^S \times \pi_n^M$ *that yields the policy when* $n \to \infty$*,*

$$\lim_{n\to\infty} \Pr(x(\omega) \text{ wins} \mid \omega; \sigma_n^*, \pi_n, n) = 1 \qquad \text{for } \omega \in \{\alpha, \beta\}.$$

Note that the *same* sender's signal π_n^M works uniformly across all admissible environments. Thus, the sender needs no detailed information about the prior, Φ, and π^S.[9,10]

The full proof of Theorem 1 is in Appendix B. We sketch the proof in Section 2.3, where we explain in detail how the sender can achieve the outcome that is perhaps the most surprising—namely, the inverted full-information outcome, $x(\alpha) = B$ and $x(\beta) = A$.

Figure 2 depicts the sender's signal that yields this outcome: With probability $1 - \varepsilon$, he sends high-quality information, meaning all of the voters receive message a in state α and message b in state β. With probability ε, he sends low-quality information, and, conditional on the state, the voters' messages are drawn

[8]We admit any $p_0 \in (0, 1)$; any Φ that is strictly increasing and continuous on $[0, 1)$, has atoms at 0 and 1 and satisfies (4); and any nature's signal π^S for which (1) and (2) hold.

[9]Note that every stochastic policy can be achieved by "mixing" over information structures.

[10]Theorem 1 is also a full characterization of the Bayes correlated equilibrium outcomes of the Condorcet setting in Section 2.1. The Bayes correlated equilibria given some exogenous information structure π^S are the Bayes–Nash equilibria that arise from expansions π of π^S; see Bergemann and Morris (2016) for the definitions. Thus, the Condorcet jury theorem fails under this wider equilibrium concept.

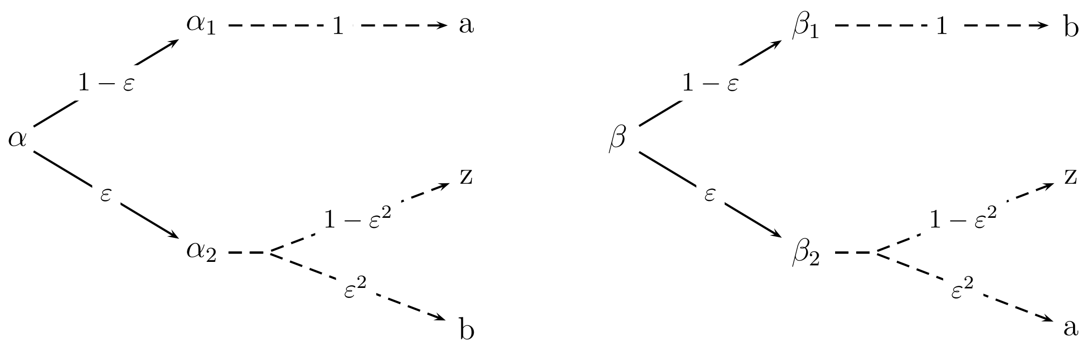

Figure 2: The sender's signal $\pi_n^M(B, A)$ that yields the "inverted" full-information outcome, $x(\alpha) = B$ in α and $x(\beta) = A$ in β, with $\varepsilon = \frac{1}{n}$. With probability $1 - \varepsilon$ the sender's information is of high quality (nodes α_1 and β_1), and with probability ε it is of low quality (nodes α_2 and β_2). Conditional on the quality and the state, the messages are i.i.d. across the voters (dashed lines).

independently and identically as follows: Each voter receives message z with probability $1-\varepsilon^2$ in both states and with the remaining probability her message is b in state α and a in state β. Thus, message z is uninformative about the state, and the meaning of messages a and b depends on the quality of the information.

This signal works because when voters condition on being pivotal, they make an inference about the other voters' message distribution; this is the indirect effect described in the introduction. In the equilibria we construct, this "social" inference leads the voters to believe that the messages are of low quality; i.e., that almost all other voters received message z. Thus, voters with message a or b will believe that their messages are inaccurate, indicating state β or α, respectively. The intuition here is that low-quality information leads to more disagreement among voters than high-quality information, so the election is endogenously closer to being tied; in other words, a tied election is indicative of low-quality messages.[11]

Importantly, this latter observation holds uniformly across environments (prior, preferences distributions, and signal distributions), which is the reason why the sender's signal is uniformly effective. The essential idea is that if the environment changes, voters endogenously adjust their behavior, and therefore no changes in the sender's signal itself are required. This mirrors the logic underlying the Condorcet jury theorem (Theorem 0), which also relies on the voters adjusting their behavior: For example, when the share of partisans for one alternative increases,

[11] An alternative interpretation of the signal is as follows: There is a public signal that reveals the state, with messages a and b, and a strategic sender can manipulate it in a pre-announced way with probability ε. Signal π_n^M could result from such manipulation.

then the voters adjust their behavior to counter this increase.[12] The proof sketch below details how the same endogenous adjustment logic is used for Theorem 1.

2.3 A Sketch of the Proof

This section explains how the sender can achieve the inverted full-information outcome, $x(\alpha) = B$ and $x(\beta) = A$, by using the signal π_n^M depicted in Figure 2. We represent the joint distribution of the messages conditional on the state ω using a two-stage lottery, where nature first draws a common "substate" ω_i and then the voters' individual messages independently conditional on the substate. Thus, the substates capture the correlation of the voters' messages. The substates of the form ω_1 correspond to high-quality information, while those of the form ω_2 correspond to low-quality information.

Given the sender's signal π_n^M from Figure 2 and an equilibrium strategy σ, the probability that a voter supports A in substate ω_j is

$$q(\omega_j;\sigma) = \sum_{s\in S, m\in M} \Pr_\pi(s,m|\omega_j)\,\Phi(\Pr(\alpha|s,m,\text{piv};\sigma)), \tag{7}$$

which is also the expected vote share for A in ω_j. This equation summarizes the effect of the sender's signal: The critical belief $\Pr(\alpha|s,m,\text{piv};\sigma)$ determines the voters' behavior via Φ, and the sender's signal affects the beliefs "directly" via the inference from m and "indirectly" via the inference from the pivotal event.

A key statistic in our analysis is the margin of victory in ω_j,

$$\left|q(\omega_j;\sigma) - \frac{1}{2}\right|,$$

which shapes the voters' inference from the pivotal event. Intuitively, a tie is more likely in a substate in which the election is closer to a tie in expectation, that is, when the margin of victory is smaller. Formally, for any pair ω_i and ω_j and strategy σ,

$$\left|q(\omega_i;\sigma) - \frac{1}{2}\right| < \left|q(\omega_j;\sigma) - \frac{1}{2}\right| \Rightarrow \frac{\Pr(\text{piv}|\omega_i;\sigma)}{\Pr(\text{piv}|\omega_j;\sigma)} > 1. \tag{8}$$

That is, if the margin of victory is smaller in substate ω_i than in ω_j, then a tie is evidence in favor of ω_i; see Claim 1 in Section B.2.1 of the appendix.[13]

[12]Whether voters adjust their behavior to changes in the environment can be tested experimentally and using field data; see Battaglini, Morton, and Palfrey (2010) and Maug and Rydqvist (2009) for some evidence that this happens in some settings.

[13]The pair ω_i and ω_j could be any pair of distinct substates, e.g., $(\omega_i, \omega_j) = (\alpha_1, \beta_2)$, which

The following lemma characterizes the margins of victory in substates α_2 and β_2, in which almost all voters receive message z.

Lemma 1 *Suppose the sender's signal is $\pi_n^M(B, A)$ from Figure 2. Then, for any admissible prior belief p_0, preference distribution Φ, and nature's signal π^S, and for every equilibrium sequence (σ_n^*) given $\pi_n = \pi^S \times \pi_n^M(B, A)$,*

$$\Phi(0) < \lim_{n\to\infty} q(\beta_2; \sigma_n^*) < \frac{1}{2} < \lim_{n\to\infty} q(\alpha_2; \sigma_n^*) < \Phi(1^-). \tag{9}$$

The proof of the lemma is in Section B.2.2 of the appendix. The middle inequalities state that a strict majority votes for B in β_2 and for A in α_2 in expectation. Hence, for these substates, the majority's choice is asymptotically the same as under full information when n is large. The outer inequalities state that the vote shares are bounded away from those under full information. With full information, all non-partisan voters would vote for B in β_2 and for A in α_2, resulting in a vote share of $\Phi(0)$ and $\Phi(1^-)$, respectively.[14]

The lemma can be understood as a consequence of the Condorcet jury theorem, Theorem 0, from the benchmark with an uninformative sender's message. Here is why. In substates α_2 and β_2, voters receive message z with probability close to 1. Suppose that this probability was *equal* to 1. Then message z would contain no information about which of the two substates, α_2 or β_2, holds. Moreover, it would be common knowledge among voters receiving message z that all of the others had received the same message. Therefore, the setting would be exactly as in the benchmark with no additional information from the sender, with α_2 taking the role of α and β_2 taking that of β. Hence Theorem 0 would apply, ensuring a majority vote for A in α_2 and for B in β_2. Now, the construction of π_n^M is such that as $n \to \infty$, the probability of message z goes to 1 rapidly enough in α_2 and β_2 so that the same characterization still applies. Importantly, the characterization of the voters' behavior in Lemma 1 applies uniformly across environments, and, as noted before, this observation is why the sender's signal will be uniformly effective.

Next we consider the behavior of voters with messages a and b. Specifically, we use a fixed-point argument to construct equilibria in which the behavior of these voters implies the inverse of the full-information outcome. Consider any strategy

is a slight abuse of notation because of the identical ω.

[14]The inner inequalities imply the outer ones. Recall that in every undominated strategy profile σ_n, partisan voters with $y = 0$ vote A, and so $\Phi(0) \le q(\beta_2; \sigma_n)$. So, a failure of the leftmost inequality, $\Phi(0) = \lim_{n\to\infty} q(\beta_2; \sigma_n)$, requires that almost all non-partisan voters with $y > 0$ vote for B with probability 1, independently of their signals. Hence, these voters would also vote for B in α_2 with probability 1, contradicting the inner inequality on the right.

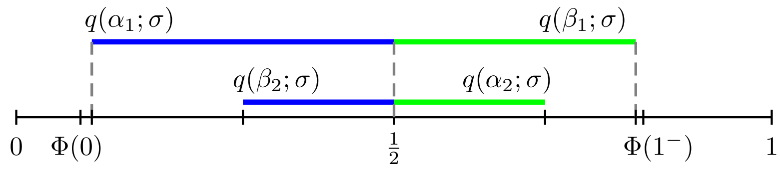

Figure 3: The approximate vote shares in the constructed equilibrium for large n. The blue lines indicate the margins of victory in the substates relevant for voters with message a, i.e., α_1 and β_2. The green lines indicate the margins in the substates relevant for voters with message b, i.e., β_1 and α_2.

profile σ^{cand} with the following properties. A voter who receives message a behaves as if the state is very likely to be β and so votes for A with probability close to $\Phi(0)$. A voter who receives message b behaves as if the state is very likely to be α and so votes for A with probability close to $\Phi(1^-)$. Lastly, a voter who receives message z behaves consistent with Lemma 1, that is, her probability to vote for A satisfies (9).

For large n, the margins of victory implied by the vote shares corresponding to such a strategy σ^{cand} are ordered as follows (we omit the superscript "cand"):

$$q(\alpha_1;\sigma) - \frac{1}{2} < q(\beta_2;\sigma) - \frac{1}{2} < 0 < q(\alpha_2;\sigma) - \frac{1}{2} < q(\beta_1;\sigma) - \frac{1}{2};$$

see Figure 3 for an illustration.

A voter who receives message a is in either substate α_1 or substate β_2. Comparing the margins of victory in the two substates and using Equation (8), substate β_2 is more likely than α_1 conditional on the voter's being pivotal. As shown in the proof, this inference from being pivotal becomes extremely strong as the number of voters becomes large, so that, conditional on the pivotal event, her signal, and her message, the voter believes that the state is almost certainly β; see Claim 2.[15] Therefore, under the best response to σ^{cand}, a voter with message a votes for A with probability close to $\Phi(0)$. By a similar argument, a voter with message b will believe that the state is very likely to be α conditional on her being pivotal, and so she will vote for A with probability close to $\Phi(1^-)$. Thus, given any such σ^{cand}, the best responses for voters with messages a and b will have the same properties.

For the equilibrium construction in Appendix B, following the outline above,

[15]The updating from message a and the updating from the pivotal event move beliefs in opposite directions. The key observation is that the likelihood ratio of message a in α_1 relative to β_2 increases at a polynomial rate, whereas the likelihood ratio of the pivotal event decreases exponentially. Therefore, the latter dominates, resulting in a vanishing posterior likelihood ratio.

we implicitly define a set of strategies that contains such strategies σ^{cand} and then construct a fixed point of the best response within this set when n is large.[16]

The proof in the appendix shows that the other state-dependent target policies can be achieved using similar information structures and with equilibria that have similar properties.

2.4 Other Equilibria

A natural question is whether it is possible to implement a specific target policy in *all* equilibria.

In the extreme case where voters have pure common interests (i.e., share a common preference type $y \in (0,1)$), it is easy to see that the answer is no. A result of McLennan (1998) shows that, in games of pure common interest, the welfare-maximizing symmetric strategy is a symmetric equilibrium. Now, if the voters ignore the sender's signal and follow a sequence of strategies as in the Condorcet jury theorem (Theorem 0), this leads to full-information outcomes. Since the welfare-maximizing strategy can only do better, McLennan's result implies that there also exists an equilibrium sequence that leads to full-information outcomes.

The following theorem shows that this observation holds even when voters do not have pure common interests.

Theorem 2 *Take the sender's signal $\pi_n^M(B,A)$ from Figure 2. There exists an equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$, given $\pi_n = \pi^S \times \pi_n^M$, for which the full-information outcome is elected as $n \to \infty$:*

$$\lim_{n\to\infty} \Pr\left(A \textit{ wins} \mid \alpha; \sigma_n^*, \pi_n, n\right) = 1 \textit{ and } \lim_{n\to\infty} \Pr\left(B \textit{ wins} \mid \beta; \sigma_n^*, \pi_n, n\right) = 1.$$

The theorem is proven in Appendix C. An analogous result holds for the information structures used to achieve the other state-dependent policies, and can be proven using analogous arguments.

One may wonder about equilibrium selection. First, note that it is not straightforward to argue that the full-information outcome is welfare-superior for the voters, especially with the more general class of preferences from Section 3.1.[17]

[16] Specifically, we work with the projection of the best response to the set of candidates and establish that it has a fixed point in the interior of its image when n is large. Methodologically, we follow Bhattacharya (2013) and express equilibrium equivalently in terms of beliefs.

[17] The full-information majority outcome may fail to maximize utilitarian welfare when there is a minority with opposing and "more intense" preferences. Krishna and Morgan (2011) demonstrates that this can have significant consequences with voting costs; see also Section 3.3.

Second, with respect to the level of strategic sophistication required, one may argue that the behavior in the manipulated equilibrium of Theorem 1 is simple, in the sense that voters with message a or b behave "approximately sincerely". They interpret their message conditional on its coming from the low-quality signal structure and simply follow that interpretation. The full-information-equivalent equilibrium from Theorem 2 does not have this property: The behavior of a voter with message a or b will depend very much on the overall preference distribution, not just on the interpretation of the message.

Third, the manipulated equilibrium has a certain "strategic robustness" feature that the full-information-equivalent equilibrium lacks. Namely, for any strategy profile in a neighborhood of the equilibrium profile, the best responses of voters with message a or b will be close to their equilibrium behavior. Similarly, if voters conjecture that the other voters will behave sincerely (i.e., vote A if $\Pr(\alpha|s,m) > y$ and B if $\Pr(\alpha|s,m) < y$), then the best-responses of the voters with message a or b to this conjecture will be close to their equilibrium behavior.[18]

2.5 Bayesian Consistency with Many Voters

Bayesian consistency is understood to constrain persuasion. This constraint is weaker in our setting with multiple receivers. This is because the receivers' behavior depends only on the "critical" posterior beliefs conditional on the strategic event of being pivotal. Bayesian consistency requires that

$$\Pr(\alpha) = \sum_{(s,m)\in S\times M} \left[\Pr(\mathrm{s},\mathrm{m},\mathrm{piv})\Pr(\alpha|s,m,\mathrm{piv}) + \Pr(\mathrm{s},\mathrm{m},\neg\mathrm{piv})\Pr(\alpha|s,m,\neg\mathrm{piv})\right],$$

where $\Pr(\alpha|s,m,\neg\mathrm{piv})$ is the posterior conditional on not being pivotal. With a single voter, $\Pr(\mathrm{piv}) = 1$, and so the expected critical belief $\mathrm{E}(\Pr(\alpha|s,m,\mathrm{piv}))$ is constrained to be the prior. However, when there are many voters, $\Pr(\mathrm{piv})$ becomes small, and consequently, Bayesian consistency imposes only a small constraint on the critical belief $\Pr(\alpha|s,m,\mathrm{piv})$ that pins down voter behavior.

[18]This is not true for voters with message z. However, these voters' behavior is unaffected by equilibrium multiplicity: In every equilibrium, it is uniquely pinned down by an equal-margins condition; see Claim 4 from the proof of Lemma 1.

3 Robustness and Extensions

3.1 General State-Dependent Preferences

In our basic model, non-partisan voters agree unanimously on the best policy when the state is known. However, such unanimity is often unrealistic. For instance, reforms can have distributive consequences and may benefit different groups of voters depending on an uncertain state of the world; see Fernandez and Rodrik (1991) or Ali, Mihm, and Siga (2018). General state-dependent preferences (Bhattacharya, 2013) capture such scenarios: In addition to partisans and voters who prefer policy A in α and policy B in β, we now consider voter types who prefer policy B in α and policy A in β.[19] Bhattacharya (2013) shows that such general preferences have substantial consequences, even in the benchmark setting without a sender. In particular, the outcome of a large election is no longer guaranteed to be full-information equivalent; i.e., information aggregation may fail.

We can study general preference distributions in terms of the implied functions $\hat{\Phi}$ that map a belief p about the state being α to the probability $\hat{\Phi}(p)$ that a randomly drawn voter prefers A to B given p. This is justified by the observation that the set of equilibrium outcomes depends on the preference distribution solely through $\hat{\Phi}$ (Bhattacharya, 2013). We continue to assume that there are partisans, so that $0 < \hat{\Phi}(0)$ and $\hat{\Phi}(1) < 1$, and that, when the state is known, there is in expectation a majority for A in α and for B in β, i.e. $\hat{\Phi}(0) < \frac{1}{2} < \hat{\Phi}(1)$. With these assumptions, our basic model is nested as the case with monotone $\hat{\Phi}(p)$, given by $\hat{\Phi}(p) = \Phi(p^-)$; see also Footnote 3.

In this broader model, because there may be types with opposing preferences, different voters may react in opposite ways to the same piece of information. This potentially complicates persuasion. Nevertheless, in analogy to Theorem 1, we identify a straightforward and easily interpreted condition on $\hat{\Phi}$ under which full persuasion is feasible. The condition requires that the election is closer to being tied in the absence of additional information than when the state is revealed.

To state the condition more precisely, suppose that the sender's signal π^M is uninformative, as in Theorem 0. Then $\hat{\Phi}$ must be such that, for any sequence of

[19] For example, the state of the world may correspond to the sector that benefits from a proposed trade reform, with voters favoring the trade reform in the state in which their own sector stands to benefit.

equilibria $(\sigma_n^*)_{n\in\mathbb{N}}$ given π^M, the equilibrium vote shares $q(\omega;\sigma_n^*)$ satisfy

$$\min\left\{\frac{1}{2}-\hat{\Phi}\left(0\right),\hat{\Phi}(1)-\frac{1}{2}\right\} > \limsup_{n\to\infty}\max_{\omega\in\{\alpha,\beta\}}\left|q(\omega;\sigma_n^*)-\frac{1}{2}\right|; \tag{M}$$

that is, the equilibrium margins of victory when the sender reveals the state (the left side) have to be larger than those when π^M is uninformative (the right side).

In Appendix D we give a formal statement of our result (Theorem 3) and provide the formal analysis. For this we adopt certain regularity conditions on $\hat{\Phi}$ from Bhattacharya (2013),[20] and we restate his characterization of equilibria for general $\hat{\Phi}$. Theorem 3 shows that if $\hat{\Phi}$ satisfies Condition M and the regularity conditions, then any state-dependent policy can be achieved in equilibrium, using the same sender's signal as in Theorem 1.

Condition M lets us extend the arguments used for the equilibrium construction from Theorem 1. Recall the proof sketch in Subsection 2.3 for achieving the inverted full-information outcome: Following the low-quality signal (substates α_2 and β_2), voters behave essentially as in the benchmark with no additional information; following the high-quality signal (substates α_1 and β_1), voters behave as if the sender had revealed the state to them (albeit incorrectly so). Now, consider a general preference distribution $\hat{\Phi}$ that satisfies M and suppose that voters still behave as just described.[21] Then M implies that the margins of victory are smaller in substates α_2 and β_2 than in substates α_1 and β_1. Therefore, conditional on her being pivotal, a voter holding message a is almost certain the state is β, while a voter holding message b is almost certain the state is α. In other words, the central observation about the updating of voters with messages a and b continues to hold, and this allows us to construct an equilibrium with the described voter behavior using similar methods as before.

Condition M is formulated in terms of endogenous equilibrium margins. Lemma 2 in the appendix identifies an equivalent condition on the exogenous fundamentals, namely the preference distribution $\hat{\Phi}$ and nature's signal π^S (given the regularity conditions). In particular, Condition M holds whenever $\hat{\Phi}$ is monotone, as in our basic model. Figure 6 illustrates how M may fail with non-monotone $\hat{\Phi}$.

The monotonicity of $\hat{\Phi}$ is critical for the Condorcet jury theorem as well. Bhattacharya (2013) demonstrates that, whenever $\hat{\Phi}$ is monotone, information aggre-

[20]Basically, $\hat{\Phi}$ should be continuously differentiable, and its derivative should satisfy several genericity conditions, e.g., $\hat{\Phi}'(0)\neq 0$ and $\hat{\Phi}'(1)\neq 0$.

[21]That is, in substates α_2 and β_2, the voters behave as in the benchmark with no additional information, and in substates α_1 and β_1 they behave as if they were sure of one of the states; albeit incorrectly of the wrong one.

gates; however, whenever it is non-monotone, information aggregation fails for some parameters of the model (i.e. for some combination of prior beliefs, signal precision, and supermajority requirements). It is not a coincidence that the monotonicity of $\hat{\Phi}$ is critical in both analyses, since for the full-persuasion result in the basic model with monotone $\hat{\Phi}$ we invoked the Condorcet jury theorem to characterize the voters' behavior.

We do not know whether it is possible to achieve full persuasion for all preference distributions and private signals, potentially using different persuasion strategies if Condition M fails. We leave this as an open question for future research.

3.2 Supermajority Voting Rules

Consider our basic model with monotone preferences, but now governed by a supermajority rule under which A is chosen if it obtains more than a share γ of the votes, where $\gamma \neq \frac{1}{2}$. If there are not too many partisans,

$$\Phi(0) < \gamma < \Phi(1^-), \tag{10}$$

then the full-information outcome is A in α and B in β. If (10) fails strictly, it is either A in both states or B in both states. Bhattacharya (2013) has proven a version of the Condorcet jury theorem in this setting and shown that if preferences are monotone (and there is no additional information from a sender), the outcomes of large elections are equivalent to those under full-information for all $\gamma \in (0, 1)$.

For any γ that satisfies (10), full persuasion can be achieved using the same sender's signal as used to establish Theorem 1. This is because the idea of the original proof remains valid. When the sender's information is of low quality (substates α_2 and β_2), the vote shares are still characterized by the Condorcet jury theorem; that is, Equation (9) from Lemma 1 holds with γ in place of $\frac{1}{2}$. This allows the construction of manipulated equilibria as before, in which the election is closer to a tie in substates α_2 and β_2 than when the sender's signal is of high quality (substates α_1 and β_1), and voters with message a or b behave as if they are sure of one of the states.[22]

Notably, the sender does not need to know the voting rule γ. The original sender's signal is uniformly effective for all γ satisfying (10).

[22] The results from Section 3.1 on the model with general preferences can also be extended to supermajority rules, with Condition M generalized to the condition $\min_{p\in\{0,1\}} \hat{\Phi}(p)^{\gamma}(1 - \hat{\Phi}(p))^{1-\gamma} > \max_{\omega\in\{\alpha,\beta\}} q(\omega;\sigma_n^*)^{\gamma}(1 - q(\omega;\sigma_n^*))^{1-\gamma}$ for all equilibria σ_n^* and n large enough.

3.3 Persuasion with Costly Participation

So far we have assumed that all voters must participate in the election, with no option to abstain. We now ask: If participation is costly and voters can abstain, can the election still be manipulated? Can voters be persuaded not only to elect a policy against their interests—such as the inverse of the full-information outcome—but also to turn out at a cost to do so?

Consider our original setting, and suppose each voter has some random participation cost $c \geq 0$, e.g., uniformly drawn from $[0, 1]$. Krishna and Morgan (2012) have shown that information still aggregates in large elections with participation costs, in a setting in which voters have pure common interests. However, when there is a conflict of interest between voters, the addition of participation costs has substantial consequences, even with a known state. The reason is that the participating voters are generally not representative, as those who "care more" have stronger participation incentives. As shown by Ledyard (1984) and Krishna and Morgan (2015), this often results in the outcome that maximizes utilitarian welfare, rather than the one preferred by the majority.[23]

Thus, participation costs fundamentally change the voting environment, even in the absence of a sender. To illustrate the specific difficulty confronting the sender when the voters face participation costs, let us revisit the equilibrium constructed in Section 2.3. There, the target policy wins by a large margin in substates α_1 and β_1 and by a smaller margin in substates α_2 and β_2. Therefore, votes are much less likely to be pivotal in α_1 and β_1 than in α_2 and β_2. This has implications for the turnout: When n is large and thus ε is small, voters with message a or b have much lower incentives to participate than voters with message z. This is because they believe the substate is more likely to be α_1 or β_1, and so their votes are less likely to be pivotal. However, if voters with message a or b are less likely to participate than those with message z, then the probability of the pivotal event in α_1 and β_1 increases relative to that in α_2 and β_2. This upsets the logic of the original equilibrium construction.

Novel arguments are therefore needed. These arguments require a technically involved analysis that appears in a separate companion note. First, the private voter types now have three dimensions: the preference y, the signal pair (s, m), and the participation cost c. Second, when the number of participating voters is random, the set of pivotal events is much larger and generally hard to characterize.

[23] For an analysis of information aggregation with participation costs and conflicts of interest, see Krishna and Morgan (2011).

To deal with the latter issue, we follow Krishna and Morgan (2012) and consider a Poisson election (Myerson, 1998), for which approximation results are available.

In the companion note we show that in the setting with participation costs, there are broad conditions under which full persuasion as in Theorem 1 is still possible. In particular, the sender can fully manipulate the election outcome when there are no partisans (i.e. when the distribution of threshold types has no mass points at 0 and 1), an assumption that includes the case of pure common values from Krishna and Morgan (2012). What matters in the constructed equilibria is that the swing voter's curse (Feddersen and Pesendorfer, 1996) corrects the problem previously sketched (that the participation incentives may be much lower for voters with message a or b than for those with message z).

Consider the signal, depicted in Figure 2, that inverts the full-information outcome. We construct an equilibrium in which essentially *only* voters with message a or b participate, and almost all of them vote against their message (i.e. they vote B after a and A after b; recall that there are no partisans). Given this behavior, the majority choice in substates α_2 and β_2 is very likely "correct." As a consequence, voters with the less informative message z face a severe swing voter's curse: A vote for A is much more likely to be pivotal in β_2 than in α_2, and vice versa for B. Because of the swing voter's curse, almost all voters with message z indeed strictly prefer to abstain in the constructed equilibrium.

Similar reasoning applies when there are merely few, rather than no, partisans: When the share of partisans is sufficiently small, the sender can still invert the outcome with probability close to 1. The companion note finally provides an extensive analysis of the case in which the share of partisans is fixed and not small. A different argument implies that arbitrarily extreme manipulation is possible in this case if nature's signal is sufficiently imprecise.

3.4 A Continuum of States

To simplify the presentation, in our basic model, we assumed a binary state space, $\{\alpha, \beta\}$. By contrast, the setting in Feddersen and Pesendorfer (1997) features a continuum of states. While, without the sender's signal, the Condorcet jury theorem holds for both settings (equilibrium outcomes satisfy full-information equivalence), the voters' behavior is qualitatively different: With a binary state space, there is a strictly positive margin of victory in equilibrium, and a large share of voters base their voting choices on their private signals. With a continuum of states, the margin of victory vanishes when the election becomes large, and the share of voters who base their voting choices on their private signals also vanishes.

One may conjecture that a continuum of states makes persuasion easier, because of the following observation. In our binary setting, the critical argument was that the margin of victory is larger in substates $\{\alpha_1, \beta_1\}$ than in substates $\{\alpha_2, \beta_2\}$. To make this comparison, we observed that the margin of victory in substates $\{\alpha_2, \beta_2\}$ is identical to the margin of victory when there is no additional information. Since, with a continuum of states, the margin of victory is 0 without additional information, replacing the binary state space with a continuum strengthens this part of the argument.

We have verified that this idea is correct: With a continuum of states, there exists a sender's signal and an equilibrium achieving the inverted full-information outcome; see the brief sketch in the online appendix, Section G..

3.5 Partial Commitment

In this section, we relax the assumption that the sender can perfectly commit to an information structure. To model partial commitment, we follow Lipnowski, Ravid, and Shishkin (2022), Min (2021), and Fréchette, Lizzeri, and Perego (2022). The sender announces an information structure but is committed to it only with probability $\chi \in (0,1)$; otherwise, he can pick any signal profile from its support.

Formally we assume that given some target state-dependent policy, denoted by $(x(\alpha), x(\beta)) \in \{A, B\}^2$, the sender's payoff is 1 if the target policy is achieved and 0 otherwise. An information structure π^M with message set M, a no-commitment sender strategy $\psi^* : \{\alpha, \beta\} \to \Delta(M^{2n+1})$, and a voter strategy σ^* form a *χ-equilibrium* (Lipnowski, Ravid, and Shishkin, 2022) if ψ^* is a best response by the sender given that the voters follow the strategy σ^*, and if σ^* is a voting equilibrium given that the sender commits to π^M with probability χ and otherwise sends messages according to ψ^*.

Perhaps surprisingly, the sender needs almost no commitment power; he can persuade the voters for *any* $\chi > 0$, no matter how small, when n is large.

Theorem 4 *Suppose that the sender is committed with some probability $\chi > 0$. For every state-dependent policy $(x(\alpha), x(\beta)) \in \{A, B\}^2$, there exists a sequence of sender's signals π_n^χ such that for any prior belief $p_0 \in (0,1)$, preference distribution Φ, and nature's signal π^S, there exists a sequence of χ-equilibria $(\pi_n^\chi, \psi_n^*, \sigma_n^*)_{n\in\mathbb{N}}$ that yields the target policy when $n \to \infty$,*

$$\lim_{n\to\infty} \Pr\left(x(\omega) \text{ wins} \mid \omega; \pi_n^\chi, \psi_n^*, \sigma_n^*, \pi^S, n\right) = 1 \quad \text{for } \omega \in \{\alpha, \beta\}.$$

Below, we prove Theorem 4 for the target policy that inverts the full-information outcome, that is, for $(x(\alpha), x(\beta)) = (B, A)$. The proofs for the other cases are along similar lines and are relegated to the appendix.

Take the information structure $\pi_n^M(B, A)$ from Figure 2. As discussed in Section 2.3, with full commitment, there are equilibria σ_n^* in which the target policy is achieved. The voting behavior is such that the vote share of A among voters with message b (resp. a) becomes arbitrarily close to $\Phi(1^-)$ (resp. $\Phi(0)$) in both states. In contrast, the vote share of A among voters with message z is consistent with Lemma 1: It converges to a number strictly between $\frac{1}{2}$ and $\Phi(1^-)$ in state α, and to a number strictly between $\Phi(0)$ and $\frac{1}{2}$ in state β. Given this voting behavior and the target policy, the sender's best response ψ^* is to send message a to all voters when the state is α and message b to all voters when the state is β.

Now note that $\pi_n^M(B, A)$ is already sending the sender's preferred message (a in α and b in β) to voters with probability close to 1. Therefore, for any $\chi > 0$ and n large enough, there is an information structure π_n^χ such that π_n^χ, χ, and ψ^* jointly imply the same message distribution as $\pi_n^M(B, A)$.[24] So σ_n^* remains the voters' best response, and $(\pi_n^\chi, \psi_n^*, \sigma_n^*)_{n\in\mathbb{N}}$ is a χ-equilibrium that yields $(x(\alpha), x(\beta)) = (B, A)$. Since one can find such π_n^χ whenever $\chi > \varepsilon$ and since $\varepsilon = 1/n$, the required commitment power vanishes as $n \to \infty$.

Theorem 4 is related to an alternative interpretation of the original sender's signal from Figure 2: There is already an exogenous public signal in place that reveals the state via messages a and b, and the sender may manipulate this public signal with probability χ in some pre-announced way.

3.6 Partially Informed Sender

Suppose the sender does not know the state $\omega \in \{\alpha, \beta\}$ but only receives a noisy signal θ correlated with it. He can commit to any message structure that is a coarsening of θ. What can persuasion achieve now?

We consider this question in a previous version of this paper, Heese and Lauermann (2019). Specifically, suppose the sender's signal is binary, $\theta \in \{\ell, h\}$, and the voters observe private signals from nature, π^S, as in our basic model. We show that whenever the sender's own information is sufficiently precise relative to π^S, the sender can achieve any policy as a function of his own signal,

[24] The sender's information structure π_n^χ is as follows. In α (resp. β), he sends message a (resp. b) to all voters with probability r, where r solves $\chi r + (1 - \chi) = 1 - \varepsilon$; otherwise, each voter receives message z with probability $1 - \varepsilon^2$ and message b (resp. a) with probability ε^2. This construction is feasible if $\chi > \varepsilon$, which ensures that r is in $(0, 1)$.

$(x(\ell), x(h)) \in \{A, B\}^2$. For example, when the voters' exogenous signals are binary and symmetric across states, it is sufficient for the sender's own information about the state to be at least as informative as the joint signal of two voters (in the Blackwell sense). To prove the result, we show that the sender can use a variant of the information structures π_n^M used to prove Theorem 1, in which the sender's own signal θ assumes the role of the state of the world ω.

3.7 Numerical Results: Small Numbers and Sincere Voters

Two numerical examples in the online appendix (Sections H and I) illustrate our persuasion mechanism further. The first shows that a group of just 11 voters can be persuaded to elect a constant target policy with a probability of 93%. Thus, our result is relevant not just for elections with large numbers of voters, but also for smaller elections and committee decisions.[25]

This paper's pivotal voting model considers the extreme case in which voters react perfectly to the closeness of the election, which may capture the behavior of some but not all voters in practice.[26] The second numerical example includes an exogenous share of 40% "sincere" voters who behave as if their own vote decides the election, based only on the information contained in their signal (s, m), without conditioning on being pivotal. Still, with 199 voters, the election outcome can be manipulated via persuasion with probability close to 1.[27]

3.8 Targeted Persuasion by an Omniscient Sender

Prior work has studied persuasion with an omniscient sender. To discuss this work in our framework, suppose that the realizations of the voters' preferences and exogenous signals are public and, in particular, known to the sender. With such an omniscient sender, it is without loss of generality to consider persuasion via obedient recommendations: Each voter i receives a message m_i, either "vote A" or "vote B," and it is incentive compatible for each voter to follow her recommendation, given the joint distribution of recommendations and the state.

[25] In small committees, communication between voters may be easier, introducing an additional constraint ignored here. However, when voters have heterogeneous preferences, communication between voters remains strategic: In equilibrium, voters condition what information to reveal also on the event that their messages are pivotal for the decisions of others; see Austen-Smith and Feddersen (2006). So, while communication complicates the problem and potentially constrains the sender, the indirect effect of the sender's signal on the pivotal inference remains.

[26] See Guarnaschelli, McKelvey, and Palfrey (2000) for experimental evidence on jury decision-making. Maug and Rydqvist (2009) provide evidence of strategic voting by shareholders.

[27] For persuasion of sincere voters see Gradwohl et al. (2022) & Arieli and Babichenko (2019).

Alonso and Câmara (2016) consider public recommendations. With public recommendations, a voter can correctly predict how others are voting and, hence, whether her vote is pivotal. When they are not pivotal, voters are indifferent. Alonso and Câmara (2016) assume a nonpivotal voter breaks her indifference by voting for the alternative she prefers given her preference and information. This assumption rules out persuasion where the sender exploits the voters' indifference when not pivotal, e.g., by recommending all voters to vote for his target policy. Alonso and Câmara (2016) characterize maximal persuasion subject to this constraint.

Private recommendations allow the sender to achieve full persuasion with strict incentives, as observed by Chan et al. (2019) and Bardhi and Guo (2018). The idea is to "pool incentive constraints."

Here is a sketch of the idea in our setting. Given the realized preferences, let $\Sigma_A \subset \{1, ..., 2n+1\}$ be the set of A-partisans and $\Sigma_B \subset \{1, ..., 2n+1\}$ the set of B-partisans. As in our basic model, the remaining non-partisan voters $\Sigma_N \subset \{1, ..., 2n+1\} \setminus (\Sigma_A \cup \Sigma_B)$ prefer A in α and B in β. Suppose that there are not too many partisans, $m_j := |\Sigma_j| < n$ for $j \in \{A, B\}$; so, neither the A-or the B-partisans constitute a majority alone, and there are at least three non-partisans. Alternative A wins if at least $n+1-m_A$ non-partisan voters support it, and B wins if at least $n+1-m_B$ non-partisans support it.

Here is a recommendation mechanism that yields the inverted full-information outcome—B in α and A in β—with probability at least $1-\varepsilon$, for arbitrarily small $\varepsilon > 0$: With probability $1-\varepsilon$, all non-partisan voters receive the recommendation "vote B" in α and "vote A" in β. With the complementary probability ε, the recommendations are as follows. In state α, a subset of $n+1-m_A$ non-partisan voters is chosen uniformly at random to receive the recommendation "vote A," and the remaining non-partisan voters receive the recommendation "vote B." In state β, a subset of $n+1-m_B$ non-partisan voters is chosen uniformly at random to receive the recommendation "vote B," and the remaining non-partisan voters receive the recommendation "vote A." Partisan voters are always recommended to vote for their preferred alternative. Thus, with probability ε, there is a minimal winning coalition for A in α and for B in β; otherwise, there is a strict majority for the inverted outcome.

Following these recommendations constitutes an equilibrium: If all other voters are obedient, a non-partisan voter is pivotal if and only if she receives the recommendation "vote A" and the state is α or the recommendation "vote B"

and the state is state β. She is never pivotal in any other event.

Obedient behavior remains an equilibrium even with public recommendations. However, a voter has only weak incentives to follow the recommendation when the public recommendation profile implies that she is not pivotal. When the recommendations are private, her incentives are strict because she is uncertain whether the realized profile makes her pivotal; in other words, the information structure pools her incentives across profiles.[28]

Critical to the above construction is that recommendations are finely tuned to details—in particular, to the sender's knowledge of the majority threshold $n+1$ and the voters' ordinal preferences. This allows the sender to pick a minimal winning coalition and thereby control incentives.

Our setting with private preferences and (exogenous) private signals requires a different approach. The sender cannot tailor recommendations to the voters' types, and thus he cannot coordinate their votes to split into precisely $n+1$ and n votes, creating minimal winning coalitions. Instead, the sender achieves full persuasion by using the logic of the Condorcet jury theorem, which implies that equilibrium margins adjust to be relatively close after very low-quality information by the sender; see Lemma 1. This logic also enables the uniform effectiveness of the sender's signal across environments, in stark contrast to the previous examples.

Finally, the messages in our construction must be private, unlike in the construction above. If they were all public, this would essentially reveal the state with probability close to 1, leading to full-information outcomes.

4 Related Literature

Persuasion. Our paper is related to work on persuasion (Kamenica and Gentzkow, 2011) and information design in general (see Bergemann and Morris (2019) for a survey), especially with multiple receivers (see, e.g., Mathevet, Perego, and Taneva (2020)). Previous work with multiple receivers has assumed an omniscient sender. In particular, in prior work in election contexts, the sender knows the voters' preferences, and the voters have no private information. In contrast, we revisit the general Condorcet voting setting of Feddersen and Pesendorfer (1997) with private preferences and exogenous private signals.

[28]The setting and information structure above mirror the example from Chan et al. (2019), p. 193. The preference specification in Bardhi and Guo (2018) is different, but similar ideas apply. In particular, the full support recommendations for their Propositions 6.3 and 6.4 remain incentive compatible when publicly announced.

Alonso and Câmara (2016)'s seminal work on voter persuasion considers public messages and shows that public messages can be used to effectively target different winning coalitions. Kawamura and Vlaseros (2017) demonstrate the effectiveness of public messages experimentally.[29]

Bardhi and Guo (2018) and Chan et al. (2019) study voter persuasion with private messages; see Section 3.8 for a detailed discussion of their full-persuasion findings. However, the main focus of their work is on settings where full persuasion is not possible. Specifically, Bardhi and Guo (2018) consider persuasion when voting takes place under the unanimity rule. Among other results, they reveal heterogeneous welfare implications of certain constraints on evidence presentation ("modes of persuasion"). Chan et al. (2019) study persuasion when it is costly for voters to choose the sender's preferred action. One interpretation is that the sender attempts to convince a quorum of people to turn out.

With a single receiver, there is a considerable body of work on Bayesian persuasion with private receiver information, such as Kamenica and Gentzkow (2011, Section 6) and Dworczak and Pavan (2022) as a more recent contribution; see the survey in Kamenica (2019) for a detailed discussion.

Our setting has multiple receivers and private information; in particular, the sender is not omniscient. We also relax other typical assumptions of the literature, such as the sender's knowledge about the state (Section 3.6) or the sender's commitment power (Section 3.5), following Lipnowski, Ravid, and Shishkin (2022).

There is further work that studies how groups can be influenced through strategic information transmission that is less closely related. For example, Gradwohl, Heller, and Hillman (2022) and Arieli and Babichenko (2019) study persuasion when all voters behave sincerely; Schnakenberg (2015) and Salcedo (2019) consider cheap-talk settings; Titova (2022) studies persuasion with verifiable information; Levy, Barreda, and Razin (2022) study persuasion with correlation neglect; and Schipper and Woo (2019) study persuasion with unawareness.

Information Aggregation. The Condorcet jury theorem provides conditions under which outcomes of large majority elections satisfy full-information equivalence. It relies on the assumption that the voters know the distribution of preferences and the signals conditional on the state. Outcomes do not satisfy full-information equivalence when there is aggregate uncertainty with respect to the preference distribution (Feddersen and Pesendorfer, 1997, Section 6), or when there is aggregate signal uncertainty conditional on the state (Mandler, 2012).[30]

[29] Wang (2013) compares persuasion with public and conditionally i.i.d. private messages.

[30] Acharya and Meirowitz (2017) also considers an uncertain signal distribution.

Aggregate uncertainty reduces the informativeness of a voter being pivotal. In the current paper, there is no aggregate uncertainty about the preferences and exogenous signals. Instead, we show how *additional* information of uncertain quality affects the election outcome and how the same additional information has predictably the same effect uniformly across environments.

Numerous more distantly related models show that elections may poorly aggregate information under a variety of further conditions; see, e.g., Razin (2003), Ali, Mihm, and Siga (2018), Bhattacharya (2013, 2018), Ekmekci and Lauermann (2020), Barelli, Bhattacharya, and Siga (2022), and Kosterina (2023).

5 Conclusion

Persuasion is ubiquitous in politics. We study the power of persuasion (Kamenica and Gentzkow, 2011) in the canonical voting setting of Feddersen and Pesendorfer (1997), in which the majority choice aggregates voters' dispersed private information across all equilibria. This result of Feddersen and Pesendorfer (1997) is a modern version of the classical Condorcet jury theorem, a benchmark that has been called "the first welfare theorem of political economy" (Krishna and Morgan, 2011). How powerful is persuasion relative to the classical force of information aggregation? It turns out to be highly effective: A strategic communicator can fully manipulate the election and achieve any state-dependent outcome just by providing additional information to voters.

Our persuasion result is robust along several dimensions. Notably, it is detail-free (Wilson, 1987)—a single information structure is uniformly effective across environments. Also, it requires only minimal commitment power.

Persuasion works via the voters' uncertainty about the quality of the additional information provided by the sender. In equilibrium, less precise, low-quality information results in less coordinated voting and, consequently, closer elections. Therefore, the closeness of the election tells voters about the quality of the other voters' signals. Given the correlation among signals, it also informs them about the quality of their own signal.

Several questions for future research naturally arise. For the model we have studied, one may ask: What additional information maximizes the probability of a majority for a given target policy when the number of voters is fixed? One may require the additional information to be robustly optimal for a class of environments in some appropriately formalized sense. Similarly, how do our initial results on persuasion with participation costs extend when the signal structure is specifically designed to encourage participation?

The starkness of our main result may be thought-provoking: What factors could limit the power of persuasion? First, our model considers the extreme case in which voters react perfectly to the closeness of the election ("pivotal voting"). One may conjecture that persuasion may be less effective if voters react less sensitively; see the numerical example from Section 3.7 for some initial observations. Second, additional limiting factors may be present in particular applications. For example, shareholder voting has been studied within the Condorcet jury framework; see Maug and Rydqvist (2009), Levit and Malenko (2011), or Bar-Isaac and Shapiro (2020). How effectively can management or other interested parties influence shareholder votes via strategic information release? What institutional and legal factors might shape and potentially constrain persuasion in this context? Another application could be to media markets, which are critical for the aggregation and dissemination of political information; see Prat and Strömberg (2013), Chan and Suen (2008), or Alonso and Miquel (2023). Our results suggest that the most effective manipulation strategy is not necessarily to bias news uniformly in one's direction but rather to seed doubts about the quality of the news and raise uncertainty about its perception by others. If feasible, even a small probability of being able to introduce bias or fake news may be sufficient to affect outcomes.

Last but not least, majority elections serve as concrete, tractable proxies for other collective decision-making processes and institutions where information aggregation is central.[31] These include informal political processes such as protests or petitions (Battaglini, 2017), information-sharing on social media (Buechel and Mechtenberg, 2019), and polls (Morgan and Stocken, 2008). How would persuasion interact with information aggregation in these environments? For example, could persuasion be used to undermine the formation of protests?

References

Acharya, A., and A. Meirowitz (2017): "Sincere voting in large elections," *Games and Economic Behavior*, 101, 121–131.

Ali, S. N., M. Mihm, and L. Siga (2018): "Adverse Selection in Distributive Politics," Working paper.

[31] This is discussed, for example, by Piketty (1999), who relates the aggregation of dispersed information in political institutions to that in markets (Hayek, 1945).

ALONSO, R., AND O. CÂMARA (2016): "Persuading voters," *American Economic Review*, 106(11), 3590–3605.

ALONSO, R., AND G. P. I. MIQUEL (2023): "Competitive Capture of Public Opinion," Working paper.

ARIELI, I., AND Y. BABICHENKO (2019): "Private bayesian persuasion," *Journal of Economic Theory*, 182, 185–217.

AUSTEN-SMITH, D., AND T. J. FEDDERSEN (2006): "Deliberation, preference uncertainty, and voting rules," *American Political Science Review*, 100(2), 209–217.

BAR-ISAAC, H., AND J. SHAPIRO (2020): "Blockholder voting," *Journal of Financial Economics*, 136(3), 695–717.

BARDHI, A., AND Y. GUO (2018): "Modes of persuasion toward unanimous consent," *Theoretical Economics*, 13(3), 1111–1149.

BARELLI, P., S. BHATTACHARYA, AND L. SIGA (2022): "Full information equivalence in large elections," *Econometrica*, 90(5), 2161–2185.

BATTAGLINI, M. (2017): "Public protests and policy making," *The Quarterly Journal of Economics*, 132(1), 485–549.

BATTAGLINI, M., R. B. MORTON, AND T. R. PALFREY (2010): "The swing voter's curse in the laboratory," *The Review of Economic Studies*, 77(1), 61–89.

BERGEMANN, D., AND S. MORRIS (2016): "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics*, 11(2), 487–522.

——— (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57(1), 44–95.

BHATTACHARYA, S. (2013): "Preference monotonicity and information aggregation in elections," *Econometrica*, 81(3), 1229–1247.

——— (2018): "Condorcet Jury theorem in a spatial model of elections," Working paper.

BUECHEL, B., AND L. MECHTENBERG (2019): "The swing voter's curse in social networks," *Games and Economic Behavior*, 118, 241–268.

CHAN, J., S. GUPTA, F. LI, AND Y. WANG (2019): "Pivotal persuasion," *Journal of Economic Theory*, 180, 178–202.

CHAN, J., AND W. SUEN (2008): "A spatial theory of news consumption and electoral competition," *The Review of Economic Studies*, 75(3), 699–728.

DWORCZAK, P., AND A. PAVAN (2022): "Preparing for the worst but hoping for the best: Robust (Bayesian) persuasion," *Econometrica*, 90(5), 2017–2051.

EKMEKCI, M., AND S. LAUERMANN (2020): "Manipulated electorates and information aggregation," *The Review of Economic Studies*, 87(2), 997–1033.

FEDDERSEN, T., AND W. PESENDORFER (1997): "Voting behavior and information aggregation in elections with private information," *Econometrica*, pp. 1029–1058.

FEDDERSEN, T. J., AND W. PESENDORFER (1996): "The swing voter's curse," *The American Economic Review*, pp. 408–424.

FERNANDEZ, R., AND D. RODRIK (1991): "Resistance to reform: Status quo bias in the presence of individual-specific uncertainty," *The American Economic Review*, pp. 1146–1155.

FRÉCHETTE, G. R., A. LIZZERI, AND J. PEREGO (2022): "Rules and commitment in communication: An experimental analysis," *Econometrica*, 90(5), 2283–2318.

GRADWOHL, R., Y. HELLER, AND A. HILLMAN (2022): "Social media and democracy," Working paper.

GUARNASCHELLI, S., R. D. MCKELVEY, AND T. R. PALFREY (2000): "An experimental study of jury decision rules," *American Political Science Review*, 94(2), 407–423.

HAYEK, F. A. (1945): "The Use of Knowledge in Society," *The American Economic Review*, 35(4), 519–530.

HEESE, C., AND S. LAUERMANN (2019): "Persuasion and Information Aggregation in Large Elections," CRC TR 224 discussion paper series, University of Bonn and University of Mannheim, Germany, https://ideas.repec.org/p/bon/boncrc/crctr224_2019_128.html.

KAMENICA, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

KAMENICA, E., AND M. GENTZKOW (2011): "Bayesian persuasion," *The American Economic Review*, 101(6), 2590–2615.

KAWAMURA, K., AND V. VLASEROS (2017): "Expert information and majority decisions," *Journal of Public Economics*, 147, 77–88.

KOSTERINA, S. (2023): "Information Structures and Information Aggregation in Elections in Threshold Equilibria in Elections," *Economic Theory*, 75, 493–522.

KRISHNA, V., AND J. MORGAN (2011): "Overcoming ideological bias in elections," *Journal of Political Economy*, 119(2), 183–211.

——— (2012): "Voluntary voting: Costs and benefits," *Journal of Economic Theory*, 147(6), 2083–2123.

——— (2015): "Majority rule and utilitarian welfare," *American Economic Journal: Microeconomics*, 7(4), 339–75.

Ledyard, J. O. (1984): "The pure theory of large two-candidate elections," *Public choice*, 44(1), 7–41.

Levit, D., and N. Malenko (2011): "Nonbinding voting for shareholder proposals," *The Journal of Finance*, 66(5), 1579–1614.

Levy, G., I. M. d. Barreda, and R. Razin (2022): "Persuasion with correlation neglect: a full manipulation result," *American Economic Review: Insights*, 4(1), 123–138.

Lipnowski, E., D. Ravid, and D. Shishkin (2022): "Persuasion via weak institutions," *Journal of Political Economy*, 130(10), 2705–2730.

Mandler, M. (2012): "The fragility of information aggregation in large elections," *Games and Economic Behavior*, 74(1), 257–268.

Mathevet, L., J. Perego, and I. Taneva (2020): "On information design in games," *Journal of Political Economy*, 128(4), 1370–1404.

Maug, E., and K. Rydqvist (2009): "Do shareholders vote strategically? Voting behavior, proposal screening, and majority rules," *Review of Finance*, 13(1), 47–79.

McLennan, A. (1998): "Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents," *American political science review*, 92(2), 413–418.

Min, D. (2021): "Bayesian persuasion under partial commitment," *Economic Theory*, 72(3), 743–764.

Morgan, J., and P. C. Stocken (2008): "Information aggregation in polls," *American Economic Review*, 98(3), 864–896.

Myerson, R. B. (1998): "Extended Poisson games and the Condorcet jury theorem," *Games and Economic Behavior*, 25(1), 111–131.

Piketty, T. (1999): "The information-aggregation approach to political institutions," *European Economic Review*, 43(4-6), 791–800.

Prat, A., and D. Strömberg (2013): "The political economy of mass media," *Advances in Economics and Econometrics*, 2, 135.

Razin, R. (2003): "Signaling and election motivations in a voting model with common values and responsive candidates," *Econometrica*, 71(4), 1083–1119.

Salcedo, B. (2019): "Persuading part of an audience," Working paper.

Schipper, B., and H. Y. Woo (2019): "Political awareness and microtargeting of voters in electoral competition," *Quarterly Journal of Political Science*, 14, 41–88.

Schnakenberg, K. E. (2015): "Expert advice to a voting body," *Journal of Economic Theory*, 160, 102–113.

Titova, M. (2022): "Persuasion with verifiable information," Working paper.

Wang, Y. (2013): "Bayesian persuasion with multiple receivers," Working paper.

Weinstein, J., and M. Yildiz (2007): "A structure theorem for rationalizability with application to robust predictions of refinements," *Econometrica*, 75(2), 365–400.

Wilson, R. (1987): "Game-theoretic analyses of trading processes," in *Advances in Economic Theory, Fifth World Congress*, pp. 33–70.

Appendix

A Representation of Equilibrium in Posterior Beliefs

We follow an idea from Bhattacharya (2013) to represent equilibrium as a fixed point in beliefs. From the point of view of a particular voter, given any strategy σ' used by the other voters, the vector of posteriors conditional on the pivotal event and on a message $m \in M$ is denoted by

$$\boldsymbol{\rho}(\sigma') = (\Pr(\alpha|m, \text{piv}; \sigma'))_{m\in M}. \tag{11}$$

It follows from the independence between the sender's signal π^M and nature's signal π^S that $\boldsymbol{\rho}(\sigma') = (\rho_m(\sigma'))_{m\in M}$ pins down the full vector of critical beliefs: For any $m \in M$ and any $s \in S$,

$$\Pr(\alpha|s, m, \text{piv}; \sigma') = \frac{\rho_m(\sigma') \Pr(s|\alpha)}{\rho_m(\sigma') \Pr(s|\alpha) + (1 - \rho_m(\sigma')) \Pr(s|\beta)}. \tag{12}$$

The critical beliefs $(\Pr(\alpha|s, m, \text{piv}; \sigma', \pi))_{(s,m)\in S\times M}$ are a sufficient statistic for the unique best response to σ' for all non-partisan types; see (5) and (6).

For any $\mathbf{p} = (p_m) \in [0, 1]^{|M|}$, we let $\hat{\sigma}^{\mathbf{p}}$ be the unique strategy that is optimal given $\mathbf{p}$, i.e., when a voter with signal (s, m) believes the probability of α is

$$p_{s,m} = \frac{p_m \Pr(s|\alpha)}{p_m \Pr(s|\alpha) + (1 - p_m) \Pr(s|\beta)}. \tag{13}$$

We follow Bhattacharya (2013) and refer to vectors of beliefs $\mathbf{p} = (p_m) \in [0, 1]^{|M|}$ as *induced priors*.

It follows from our discussion that equilibrium is equivalently characterized by a vector of induced priors $\mathbf{p}^* = (p^*_m)_{m\in M}$ such that

$$\mathbf{p}^* = \boldsymbol{\rho}(\hat{\sigma}^{\mathbf{p}^*}). \tag{14}$$

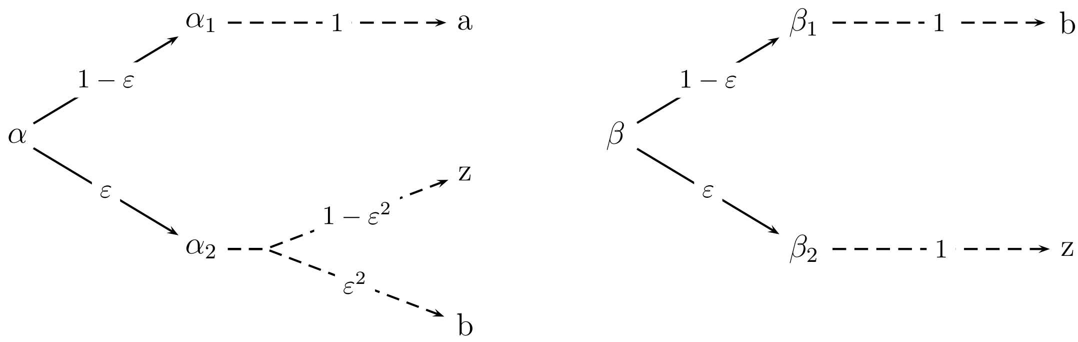

Figure 4: The sender's signal $\pi_n^M(A, A)$ that yields the constant outcome, $x(\alpha) = x(\beta) = A$, with $\varepsilon = \frac{1}{n}$.

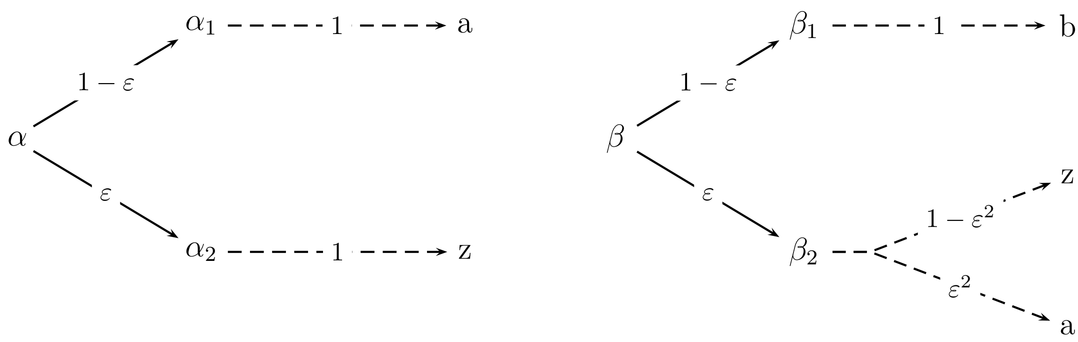

Figure 5: The sender's signal $\pi_n^M(B, B)$ that yields the constant outcome $x(\alpha) = x(\beta) = B$, with $\varepsilon = \frac{1}{n}$.

B Proof of Theorem 1

B.1 Preliminaries

The Sender's Signal. The sender's signals that yield the four possible target policies are as follows. The inverted full-information outcome, $x(\alpha) = B$ and $x(\beta) = A$, is achieved using the sender's signal $\pi_n^M(B, A)$ from Figure 2 in Section 2.3. The constant outcomes $x(\alpha) = x(\beta) = A$ and $x(\alpha) = x(\beta) = B$ are achieved using the sender's signals $\pi_n^M(A, A)$ and $\pi_n^M(B, B)$ from Figures 4 and 5, respectively. Finally, the sender can achieve the full-information outcome either by revealing the state or by sending an uninformative signal, given the Condorcet jury theorem (Theorem 0).

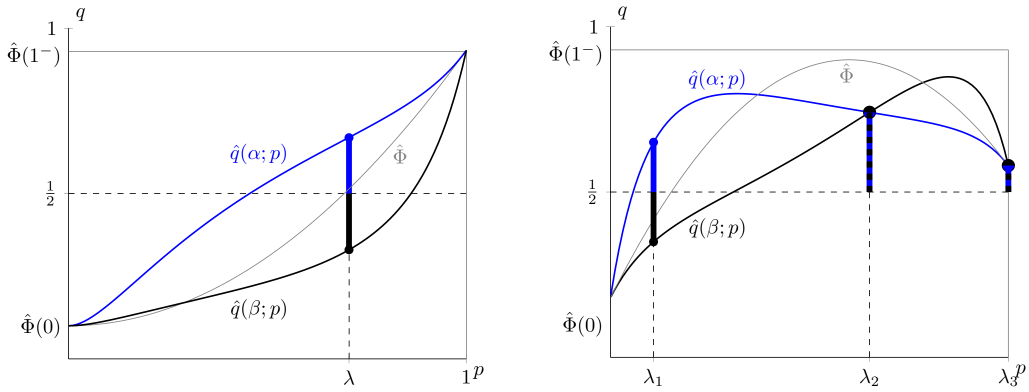

Figure 6: The functions $\hat{q}(\alpha;p)$ and $\hat{q}(\beta;p)$ of the implied vote shares for A in states α and β, respectively, given an induced prior $p \in (0,1)$. The left panel shows the function for an example of a monotone Φ and the right for a non-monotone Φ. The functions also illustrate the equilibria characterized by the equal-margins condition, discussed later. The function Φ is grey, $\hat{q}(\alpha;p)$ blue, and $\hat{q}(\beta;p)$ black.

Notation. We express all relevant equilibrium magnitudes in terms of beliefs (a hat over a symbol signifies functions of beliefs). First, for any $p \in [0,1)$,

$$\hat{q}(\omega;p) = \sum_{s\in S} \Pr(s|\omega)\Phi\left(\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)}\right) \tag{15}$$

is the probability that, in state ω, a voter with induced prior p draws a type y and a signal $s \in S$ such that she then prefers to vote for A. For $p=1$, we have

$$\hat{q}(\omega,1) = \Phi(1^-). \tag{16}$$

Figure 6 illustrates the functions $\hat{q}(\omega;p)$. We will use the facts that $\hat{q}(\omega;p)$ is continuous and strictly increasing in p, and that

$$\hat{q}(\alpha;p) > \hat{q}(\beta;p) \tag{17}$$

for $p \in (0,1)$. These facts follow immediately from the facts that Φ is continuous and strictly increasing on $(0,1)$, and that the distribution of posteriors $\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)}$ in α first-order stochastically dominates the distribution of posteriors $\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)}$ in β.[32]

The sender's signals $\pi_n^M(x(\alpha),x(\beta))$ corresponding to the target outcomes

[32]This is a standard observation for beliefs over binary states. Ordering signals by the induced posteriors, the distributions of signals in α and β are ordered by likelihood ratio dominance, and hence by first-order stochastic dominance.

other than the full-information outcome, shown in Figures 2, 4, and 5, are represented via substates $\alpha_1, \alpha_2, \beta_1, \beta_2$. Given a vector of induced priors $\mathbf{p} = (p_m) \in [0,1]^{|M|}$, the vote share of policy A in substate ω_j is

$$\hat{q}(\omega_j; \mathbf{p}) = \sum_{m \in M} \Pr(m|\omega_j)\hat{q}(\omega; p_m) . \tag{18}$$

The vote share $\hat{q}(\omega_j; \mathbf{p})$ determines the probability of the pivotal event in substate ω_j:

$$\Pr(\text{piv}|\omega_j; \mathbf{p}) = \binom{2n}{n}\Big[q(\omega_j; \mathbf{p})(1 - q(\omega_j; \mathbf{p}))\Big]^n . \tag{19}$$

The pivotal probabilities in (19) and the sender's signal π_n^M determine a voter's posterior beliefs conditional on being pivotal and receiving signal $m \in M$:

$$\hat{\boldsymbol{\rho}}(\mathbf{p}) = (\Pr(\alpha|m, \text{piv}; \mathbf{p}))_{m \in M}. \tag{20}$$

By construction, $\hat{\boldsymbol{\rho}}(\mathbf{p}) = \boldsymbol{\rho}(\hat{\sigma}^{\mathbf{p}})$. With this notation, the proof can be carried out entirely in terms of beliefs.

B.2 Proof

B.2.1 Pivotal Inference and the Margin of Victory

In (8), we noted that a tie is indicative of the substate with the smaller margin of victory. In accordance with our general method of proof (see Appendix A), we express this fact in terms of the vote shares that result from an induced prior $\mathbf{p}$.

Claim 1 *Consider two substates $\omega_i, \omega_j \in \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$. Take any induced prior vector $\mathbf{p}$ for which $\Pr(\text{piv}|\omega_j; \mathbf{p}) \in (0,1)$; if*

$$\left|\hat{q}(\omega_i; \mathbf{p}) - \frac{1}{2}\right| < \left|\hat{q}(\omega_j; \mathbf{p}) - \frac{1}{2}\right|, \tag{21}$$

then

$$\frac{\Pr(\text{piv}|\omega_i; \mathbf{p})}{\Pr(\text{piv}|\omega_j; \mathbf{p})} > 1. \tag{22}$$

Proof. The function $q(1-q)$ has an inverse U-shape on $[0,1]$ and is symmetric around its peak at $q = \frac{1}{2}$. So $|q - \frac{1}{2}| < |q' - \frac{1}{2}|$ implies $q(1-q) > q'(1-q')$. It follows from (19) that (21) implies (22). ■

As the electorate grows, a voter's inference from being pivotal rapidly strengthens. Specifically, the likelihood ratio of being pivotal diverges exponentially in n.

Claim 2 *Consider two substates $\omega_i, \omega_j \in \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$. Take any sequence of induced prior vectors $(\mathbf{p}_n)_{n\in\mathbb{N}}$ for which $\Pr(\text{piv}|\omega_j; \mathbf{p}_n) \in (0,1)$ for all n. If*

$$\lim_{n\to\infty}\left|\hat{q}\left(\omega_i;\mathbf{p}_n\right)-\frac{1}{2}\right| < \lim_{n\to\infty}\left|\hat{q}\left(\omega_j;\mathbf{p}_n\right)-\frac{1}{2}\right|, \tag{23}$$

then, for any $d \geq 0$,

$$\lim_{n\to\infty}\frac{\Pr\left(\text{piv}|\omega_i;\mathbf{p}_n\right)}{\Pr\left(\text{piv}|\omega_j;\mathbf{p}_n\right)}n^{-d} = \infty. \tag{24}$$

Proof. Let

$$k_n = \frac{\hat{q}\left(\omega_i;\mathbf{p}_n\right)\left(1-\hat{q}\left(\omega_j;\mathbf{p}_n\right)\right)}{\hat{q}\left(\omega_j;\mathbf{p}_n\right)\left(1-\hat{q}\left(\omega_i;\mathbf{p}_n\right)\right)}.$$

From (19), the left-hand side of (24) is $\frac{(k_n)^n}{n^d}$. If (23) holds, then $\lim_{n\to\infty} k_n > 1$ because of the properties of $q(1-q)$ (it has an inverse U-shape around 1/2). Therefore, $\lim_{n\to\infty}(k_n)^n = \infty$. Moreover, $(k_n)^n$ diverges exponentially fast and hence dominates the denominator n^d, which is polynomial. ■

B.2.2 Proof of Lemma 1 and Voter Inference after Message z

For readers jumping here directly from the main text, note that the proof utilizes notation and auxiliary methods introduced at the start of the appendix.

We prove Lemma 1 from the main body of the paper: The limit equilibrium vote shares in α_2 and β_2 satisfy the ordering (9).

In fact, Claim 4 shows something stronger: The vote shares in α_2 and β_2 are uniquely pinned down by an equal-margins-of-victory condition across all equilibrium sequences (and equivalently, across all sequences of equilibrium induced prior vectors $\mathbf{p} = (p_a^n, p_z^n, p_b^n))$. Moreover, the conclusion holds even if just p_z^n satisfies the fixed-point condition $p_z^n = \hat{\rho}_z(\mathbf{p})$, even if p_a^n and p_b^n do not. Therefore, the result does not only apply to equilibrium induced prior vectors. We start by showing an auxiliary result: p_z^n cannot converge to an extreme belief.

Claim 3 *For any sequence of induced prior vectors $\mathbf{p}_n = (p_m^n)_{m\in M}$ for which $p_z^n = \hat{\rho}_z(\mathbf{p}_n)$ for all n,*

$$\lim_{n\to\infty}\hat{\rho}_z(\mathbf{p}_n) \notin \{0,1\}. \tag{25}$$

Proof. Take any sequence of induced priors $(\mathbf{p}_n)_{n\in\mathbb{N}}$ with $\mathbf{p}_n = (p_m^n)_{m\in M}$ for which p_z^n has the fixed-point property $p_z^n = \hat{\rho}_z(\mathbf{p}_n)$ for all n.

As a first observation, note that $p_z^n = \hat{\rho}_z(\mathbf{p}_n)$ implies that p_z^n is interior (i.e. $p_z^n \in (0,1)$) for all n. This is because even if $\mathbf{p}_n$ is at the extremes, the existence of partisans implies that the probability of being pivotal is strictly positive in all substates.

As noted in (17), for interior induced priors, the implied vote share of A is higher in α than in β, and so

$$\hat{q}\left(\alpha; p_z^n\right) > \hat{q}\left(\beta; p_z^n\right). \tag{26}$$

Next, we claim that the vote share of A is at most $\frac{1}{n^2}$ smaller in α_2 than in β_2:

$$\hat{q}(\alpha_2; \mathbf{p}_n) - \hat{q}(\beta_2; \mathbf{p}_n) \geq -\frac{1}{n^2}. \tag{27}$$

For messages $m \in \{a, b\}$, the ordering of the implied vote shares $\hat{q}\left(\alpha; p_m^n\right)$ and $\hat{q}\left(\beta; p_m^n\right)$ may be the reverse of (26). However, in α_2 and β_2, the likelihood that a voter does not receive message z is smaller than $\frac{1}{n^2}$. So (27) follows from (18) and (26).

Next, we prove by (25) by contradiction. We have

$$\frac{\hat{\rho}_z(\mathbf{p}_n)}{1-\hat{\rho}_z(\mathbf{p}_n)} = \frac{p_0}{1-p_0} \frac{\Pr(\alpha_2|\alpha)}{\Pr(\beta_2|\beta)} \frac{\Pr(\text{piv}|\alpha_2; \mathbf{p}_n)}{\Pr(\text{piv}|\beta_2; \mathbf{p}_n)}, \tag{28}$$

with the pivotal probability in each substate ω_j given by (19).

Suppose that $\lim_{n\to\infty} \hat{\rho}_z(\mathbf{p}_n) = 0$. We show that this implies

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha_2; \mathbf{p}_n)}{\Pr(\text{piv}|\beta_2; \mathbf{p}_n)} \geq 1, \tag{29}$$

a contradiction. Since almost all voters receive message z in α_2 and β_2, the hypothesis $\lim_{n\to\infty} p_z^n = \lim_{n\to\infty} \hat{\rho}_z(\mathbf{p}_n) = 0$ implies that $\lim_{n\to\infty} \hat{q}(\alpha_2; \mathbf{p}_n) = \lim_{n\to\infty} \hat{q}(\beta_2; \mathbf{p}_n) = \Phi(0)$, given (15) and (18). Since $\Phi(0) < \frac{1}{2}$,

$$\lim_{n\to\infty} \hat{q}(\alpha_2, \mathbf{p}_n) = \lim_{n\to\infty} \hat{q}(\beta_2, \mathbf{p}_n) < \frac{1}{2}. \tag{30}$$

Note that, because of the partisans, $\Phi(0) \leq \hat{q}(\omega_j; \mathbf{p}_n) \leq \Phi(1^-)$ for any induced prior vector and any substate ω_j. Also, the derivative of the function $h(q) = q(1-q)$ is bounded below by some Lipschitz constant $L > 0$ on the compact interval $[\Phi(0), \Phi(1^-)]$. Note that (27) implies

$$h(\hat{q}(\beta_2, \mathbf{p}_n))(\frac{h(\hat{q}(\alpha_2, \mathbf{p}_n))}{h(\hat{q}(\beta_2, \mathbf{p}_n))} - 1) = h(\hat{q}(\alpha_2, \mathbf{p}_n)) - h(\hat{q}(\beta_2, \mathbf{p}_n)) \geq -\frac{L}{n^2}. \tag{31}$$

Recall that the function $h(q)$ is inverse-U-shaped with a peak at $q = \frac{1}{2}$, and by (4), $\Phi(0) < \frac{1}{2} < \Phi(1^-)$. Since $\Phi(0) \leq \hat{q}(\beta_2; \mathbf{p}_n) \leq \Phi(1^-)$, we have $M \leq h(\hat{q}(\beta_2; \mathbf{p}_n)$ for

$M = \min\{(h(\Phi(0)), h(\Phi(1^-)))\}$ and all n. Thus,

$$\frac{h(\hat{q}(\alpha_2, \mathbf{p}_n))}{h(\hat{q}(\beta_2, \mathbf{p}_n))} \geq 1 - \frac{L}{h(\hat{q}(\beta_2\,; \mathbf{p}_n))n^2} \geq 1 - \frac{L}{Mn^2}, \tag{32}$$

where we have rewritten (31) for the first inequality and used $M \leq h(\hat{q}(\beta_2\,; \mathbf{p}_n)$ for the second. It follows from (19) that $\frac{\Pr(\text{piv}|\alpha_2;\mathbf{p}_n)}{\Pr(\text{piv}|\beta_2;\mathbf{p}_n)} \geq (1 - \frac{L}{Mn^2})^n$. Now (29) follows since $\lim_{n\to\infty}(1 - \frac{L}{Mn^2})^n = 1$.[33] A similar argument excludes $\lim_{n\to\infty} \hat{\rho}_z(\mathbf{p}_n) = 1$ (using the bound analogous to (27)). This finishes the proof of (25). ■

Claim 4 *For any π^S, there is some D with*

$$0 < D < \min\left\{\frac{1}{2} - \Phi\,(0)\,, \Phi(1^-) - \frac{1}{2}\right\}$$

such that, for any sequence of induced prior vectors $\mathbf{p}_n = (p_m^n)_{m\in M}$ for which $p_z^n = \hat{\rho}_z(\mathbf{p}_n)$ for all n,

$$\lim_{n\to\infty} \hat{q}(\alpha_2; \mathbf{p}_n) - \frac{1}{2} = \lim_{n\to\infty} \frac{1}{2} - \hat{q}(\beta_2; \mathbf{p}_n) = D. \tag{33}$$

Proof. It follows from (25) that voters do not become certain conditional on their being pivotal and on the substate's being α_2 or β_2; that is,

$$\lim_{n\to\infty} \Pr(\alpha|\{\alpha_2, \beta_2\}, \text{piv}; \mathbf{p}_n) \notin \{0, 1\}.$$

Hence, Claim 2 requires

$$\lim_{n\to\infty} \left|\hat{q}(\alpha_2; \mathbf{p}_n) - \frac{1}{2}\right| = \lim_{n\to\infty} \left|\hat{q}(\beta_2; \mathbf{p}_n) - \frac{1}{2}\right|. \tag{34}$$

An implication of (34) is that, as n becomes large, the voter's induced prior p_z converges to the belief $\lambda \in (0, 1)$ that is the unique interior solution to the equal-margins condition

$$\left|\hat{q}(\alpha; \lambda) - \frac{1}{2}\right| = \left|\hat{q}(\beta; \lambda) - \frac{1}{2}\right| > 0. \tag{35}$$

The existence and uniqueness of such a belief follows directly from the properties of $\hat{q}(\omega; p)$ illustrated in Figure 6. The stated convergence

$$\lim_{n\to\infty} \hat{\rho}_z(\mathbf{p}_n) = \lim_{n\to\infty} p_z^n = \lambda \tag{36}$$

[33] For this, note that $(1 - \frac{L}{Mn^2})^n \leq 1$ for all n. Furthermore, for any $d > 0$, $\frac{L}{Mn} < \frac{1}{d}$ when n is large enough, so that $\lim_{n\to\infty}(1 - \frac{L}{Mn^2})^n \geq \lim_{n\to\infty}(1 - \frac{1}{n}\frac{1}{d})^n$. The claim follows from recalling that $\lim_{n\to\infty}(1 - \frac{1}{n}\frac{1}{d})^n = e^{-\frac{1}{d}}$ and $e^0 = 1$.

follows from (25), which rules out the limits 0 and 1, and from (34), since almost all voters receive message z in α_2 and β_2, so that $\hat{q}(\omega; p_z^n) \to \hat{q}(\omega_2; \mathbf{p}_n)$ for $\omega_2 \in \{\alpha_2, \beta_2\}$. Finally, Claim 4 follows from (35), (36), and (26), with $D = \hat{q}(\alpha; \lambda) - \frac{1}{2}$. ■

As a by-product of the proof of Claim 4, (36) characterizes the unique limit of the induced prior after message z across all equilibria. Next, we turn to the voter's inference after messages a and b.

B.2.3 Voter Inference after Message a or b

Consider a voter who has received a message $m \in \{a, b\}$. The following result shows that if the election is closer to being tied in substates α_2 and β_2 than in substates α_1 and β_1, then the voter's inference from the message is dominated by her inference from the pivotal event.

Claim 5 *Take any sequence of induced prior vectors $(\mathbf{p}_n)_{n\in\mathbb{N}}$ such that*

$$\lim_{n\to\infty} \min_{\omega_1\in\{\alpha_1,\beta_1\}} \left|\hat{q}(\omega_1; \mathbf{p}_n) - \frac{1}{2}\right| > \lim_{n\to\infty} \max_{\omega_2\in\{\alpha_2,\beta_2\}} \left|\hat{q}(\omega_2; \mathbf{p}_n) - \frac{1}{2}\right|. \quad (37)$$

Then, if the sender's signal is $(\pi_n^M(B, A))_{n\in\mathbb{N}}$ or $(\pi_n^M(A, A))_{n\in\mathbb{N}}$, for any $m \in M$,

$$\lim_{n\to\infty} \frac{\Pr(\{\alpha_2, \beta_2\}|s, b, \text{piv}; \mathbf{p}_n)}{\Pr(\{\alpha_1, \beta_1\}|s, b, \text{piv}; \mathbf{p}_n)} = \infty. \quad (38)$$

If the sender's signal is $(\pi_n^M(B, A))_{n\in\mathbb{N}}$ or $(\pi_n^M(B, B))_{n\in\mathbb{N}}$, for any $m \in M$,

$$\lim_{n\to\infty} \frac{\Pr(\{\alpha_2, \beta_2\}|s, a, \text{piv}; \mathbf{p}_n)}{\Pr(\{\alpha_1, \beta_1\}|s, a, \text{piv}; \mathbf{p}_n)} = \infty. \quad (39)$$

Proof. Take the message $m = b$. Then the posterior likelihood ratio is

$$\begin{aligned}\frac{\Pr(\{\alpha_2, \beta_2\}|b, \text{piv}; \mathbf{p}_n)}{\Pr(\{\alpha_1, \beta_1\}|b, \text{piv}; \mathbf{p}_n)} &= \frac{p_0}{1-p_0}\frac{\Pr(\alpha_2|\alpha)}{\Pr(\beta_1|\beta)}\frac{\Pr(b|\alpha_2)}{\Pr(b|\beta_1)}\frac{\Pr(\text{piv}|\alpha_2; \mathbf{p}_n)}{\Pr(\text{piv}|\beta_1; \mathbf{p}_n)} \\ &= \frac{p_0}{1-p_0}\frac{n^{-3}}{(1-\frac{1}{n})}\frac{\Pr(\text{piv}|\alpha_2; \mathbf{p}_n)}{\Pr(\text{piv}|\beta_1; \mathbf{p}_n)}.\end{aligned}$$

Applying Claim 2 for $d = 3$, we see that this quantity diverges to ∞ as $n \to \infty$. So a voter with message b becomes convinced that the state is α for either realization of the private signal $s \in S$:

$$\lim_{n\to\infty} \Pr(\alpha|s, b, \text{piv}; \mathbf{p}_n) = 1. \quad (40)$$

The argument for message a is analogous. ■

We conclude this section by evaluating the limit behavior of the posteriors $\hat{\rho}_m(\mathbf{p}_n) = \Pr(\alpha|m, \text{piv}; \mathbf{p}_n)$ for $m \in \{a, b\}$. Recall that the sender's signal is $(\pi_n^M(x(\alpha), x(\beta)))_{n\in\mathbb{N}}$ for some $(x(\alpha), x(\beta)) \in \{(B, A), (A, A), (B, B)\}$. Take any sequence of induced prior vectors $(\mathbf{p}_n)_{n\in\mathbb{N}}$ that satisfies (37). Note that if $x(\beta) = B$, then a voter who receives message b learns that the state is β. If $x(\alpha) = A$, a voter who receives message a learns that the state is α. Combining this with Claim 5, we conclude that, for any sequence satisfying (37),

$$\lim_{n\to\infty} \hat{\rho}_a(\mathbf{p}_n) = 1 - 1_{x(\alpha)=B}, \tag{41}$$
$$\lim_{n\to\infty} \hat{\rho}_b(\mathbf{p}_n) = 1_{x(\beta)=A}. \tag{42}$$

B.2.4 Fixed-Point Construction

In this section, we construct a sequence of equilibria using the observations from the preceding two sections to prove Theorem 1.

Recall from Section A that equilibrium can equivalently be characterized by a vector of induced priors $\mathbf{p}^* = (p_a^*, p_z^*, p_b^*)$ such that $\mathbf{p}^* = \hat{\boldsymbol{\rho}}(\mathbf{p}^*)$; see (14). Now, take any $\delta > 0$ and let

$$\mathrm{B}_\delta = \left\{ \mathbf{p} \in [0,1]^3 \;\middle|\; p_z \in \left[\frac{\delta}{2}, 1 - \frac{\delta}{2}\right],\ \left|p_a - (1 - 1_{x(\alpha)=B})\right| \le \delta,\ \left|p_b - 1_{x(\beta)=A}\right| \le \delta \right\}.$$

Take any $\mathbf{p} \in \mathrm{B}_\delta$. We define a constrained best response function as its "truncation" to B_δ:

$$\hat{\rho}_z^{tr}(\mathbf{p}) = \begin{cases} \frac{\delta}{2} & \text{if} \quad \hat{\rho}_z(\mathbf{p}) < \frac{\delta}{2}, \\ 1 - \frac{\delta}{2} & \text{if} \quad \hat{\rho}_z(\mathbf{p}) > 1 - \frac{\delta}{2}, \\ \hat{\rho}_z(\mathbf{p}) & \text{otherwise.} \end{cases} \tag{43}$$

The components $\hat{\rho}_a^{tr}$ and $\hat{\rho}_b^{tr}$ are defined analogously. The function $\hat{\boldsymbol{\rho}}^{tr}(\mathbf{p}) = (\hat{\rho}_a^{tr}(\mathbf{p}), \hat{\rho}_z^{tr}(\mathbf{p}), \hat{\rho}_b^{tr}(\mathbf{p}))$ is continuous in $\mathbf{p}$, so that Kakutani's theorem implies that $\hat{\boldsymbol{\rho}}^{tr}(\mathbf{p})$ has a fixed point $\mathbf{p}^* \in \mathrm{B}_\delta$.

The next claim shows that any fixed point $\mathbf{p}^*$ of $\hat{\boldsymbol{\rho}}^{tr}$ is in the interior of B_δ when n is large enough and δ is small enough, i.e., $\hat{\boldsymbol{\rho}}^{tr}(\mathbf{p}^*) = \hat{\boldsymbol{\rho}}(\mathbf{p}^*)$.

Claim 6 *Let the sender's signal be $(\pi_n^M(B, A))_{n\in\mathbb{N}}$, $(\pi_n^M(A, A))_{n\in\mathbb{N}}$, or $(\pi_n^M(B, B))_{n\in\mathbb{N}}$. There is $\delta > 0$ small enough so that for any sequence of fixed points $(\mathbf{p}_n)_{n\in\mathbb{N}}$ of $\hat{\boldsymbol{\rho}}^{tr}$, there is $\bar{n} \in \mathbb{N}$ such that for all $n \ge \bar{n}$, the fixed point $\mathbf{p}_n$ is in the interior of B_δ.*

Proof. Choose some $\tilde{\delta} < p_0$ small enough so that

$$\lim_{n\to\infty} \hat{q}(\alpha_2, \mathbf{p}_n) = \lim_{n\to\infty} \hat{q}(\beta_2, \mathbf{p}_n) < \frac{1}{2}. \tag{44}$$

holds for any $\delta \le \tilde{\delta}$ and any sequence $(\mathbf{p}_n)_{n\in\mathbb{N}}$ with $p_z^n = \frac{\delta}{2}$ for all $n \in \mathbb{N}$. This is possible given $\Phi(0) < \frac{1}{2}$; compare to (30).

Fix some $\delta \le \tilde{\delta}$, and pick any sequence of fixed points $(\mathbf{p}_n)_{n\in\mathbb{N}}$ given δ. Consider $n \in \mathbb{N}$ and suppose that $p_z^n = \frac{\delta}{2}$. Given (44), repeating the same argument as in the proof of (25) (which was an intermediate result in the proof of Claim 4), the assumption $p_z^n = \frac{\delta}{2}$ implies $\frac{\Pr(\text{piv}|\alpha_2;\mathbf{p}_n)}{\Pr(\text{piv}|\beta_2;\mathbf{p}_n)} \ge (1 - \frac{L}{Mn^2})^n$ for two constants $L > 0$ and $M > 0$; compare to (32).[34] Since $\lim_{n\to\infty}(1 - \frac{L}{Mn^2})^n = 1$ and since $p_0 > \delta$, there is $\tilde{n}$ large enough so that $\Pr(\alpha|\text{piv}; z, \mathbf{p}_n) > \delta$ for all $n \ge \tilde{n}$ (note that we can choose $\tilde{n}$ independent of the fixed point sequence and the parameter $\delta \le \tilde{\delta}$ considered.) We arrive at a contradiction to the initial assumption $p_z^n = \frac{\delta}{2}$. In the same way we can find $\hat{\delta} > 0$ small enough and $\hat{n}$ so that for any $\delta \le \hat{\delta}$ and any $n \ge \hat{n}$ no fixed point can have $p_z^n = 1 - \frac{\delta}{2}$.

Thus, fixing $\bar{\delta} = \min(\tilde{\delta}, \hat{\delta})$, for any sequence $(\mathbf{p}_n)_{n\in\mathbb{N}}$ of fixed points, it holds $p_z^n = \hat{\rho}_z(\mathbf{p}_n)$ for $n \ge \max(\tilde{n}, \hat{n})$, and thus Claim 4 implies that

$$\hat{\rho}_z(\mathbf{p}_n) \approx \lambda. \tag{45}$$

Now, the strict monotonicity of Φ implies

$$\left|\Phi(0) - \frac{1}{2}\right| > D, \tag{46}$$

$$\left|\Phi(1^-) - \frac{1}{2}\right| > D, \tag{47}$$

for $D = |\hat{q}(\alpha_2; \lambda) - \frac{1}{2}|$; see Claim 4 for the definition of D. Given the definition of B_δ and the equal-margins property (35) of λ, this implies that there is $\delta \le \bar{\delta}$ so that the expected margins of victory in substates α_2 and β_2 are strictly smaller than those in substates α_1 and β_1 as $n \to \infty$. In other words, $\lim_{n\to\infty} |\hat{q}(\omega_2; \mathbf{p}_n) - \frac{1}{2}| > \lim_{n\to\infty} |\hat{q}(\omega_1; \mathbf{p}_n) - \frac{1}{2}|$ for all $\omega_2 \in \{\alpha_2, \beta_2\}$ and $\omega_1 \in \{\alpha_1, \beta_1\}$; that is, $(\mathbf{p}_n)_{n\in\mathbb{N}}$ satisfies (37). Therefore, recalling the discussion after Claim 5, we deduce that

[34] More precisely, repeat the argument starting from the assumption $p_z < \delta$ immediately before (29) up to Footnote 33.

(41) and (42) hold; that is,

$$\hat{\rho}_a(\mathbf{p}_n) \approx 1 - 1_{x(\alpha)=B}, \tag{48}$$
$$\hat{\rho}_b(\mathbf{p}_n) \approx 1_{x(\beta)=A}. \tag{49}$$

So, fixing such $\delta \leq \bar{\delta}$, for any fixed point sequence $(\mathbf{p}_n)_{n\in\mathbb{N}}$, there is $\bar{n}$ large enough so that $\mathbf{p}_n$ is interior for all $n \geq \bar{n}$.[35] ■

Now we finish the proof of Theorem 1. For the full-information outcome, $x(\alpha) = A$ and $x(\beta) = B$, the statement follows from the Condorcet jury theorem (Theorem 0), and the sender can achieve it by sending an uninformative message, e.g., by letting M be a singleton, $M = \{m\}$. For any other target policy $(x(\alpha), x(\beta)) \in \{(B, A), (A, A), (B, B)\}$, the above fixed-point argument applies: Note that the strategy corresponding to any interior fixed point $\mathbf{p}^*$ of $\hat{\boldsymbol{\rho}}^{tr}$ is an equilibrium. Therefore, Claim 6 implies the existence of a sequence of equilibria $(\sigma_n^*)_{n\in\mathbb{N}}$ for which the corresponding induced priors satisfy (48) and (49); so the vote shares in α_1 and β_1 converge to either $\Phi(0)$ or $\Phi(1^-)$ as $n \to \infty$, depending on the target policy. Since $\Phi(0) < \frac{1}{2}$ and $\Phi(1^-) > \frac{1}{2}$, an application of the law of large numbers implies that the target policy $x(\omega)$ is elected with probability converging to 1 in ω, as $n \to \infty$. This finishes the proof of Theorem 1.

C Proof of Theorem 2: Other Equilibria

For any vector $\mathbf{p} = (p_m)_{m\in M}$, consider the vote share vector $(\hat{q}(\omega_i; \mathbf{p}))_{\omega_i\in\{\alpha_1,\beta_1,\alpha_2,\beta_2\}}$. Truncate this vector so that the vote share in α_1 is weakly larger than $\frac{1}{2}$ and the vote share in β_1 is weakly smaller than $\frac{1}{2}$:

$$\hat{q}^{\text{tr}}(\omega_i; \mathbf{p}) = \begin{cases} \frac{1}{2} \text{ if } \omega_i = \alpha_1 \text{ and } \hat{q}(\alpha_1; \mathbf{p}) < \frac{1}{2}, \\ \frac{1}{2} \text{ if } \omega_i = \beta_1 \text{ and } \hat{q}(\beta_1; \mathbf{p}) > \frac{1}{2}, \\ \hat{q}(\omega_i; \mathbf{p}) \text{ otherwise.} \end{cases} \tag{50}$$

Finally, define $\hat{\rho}^{\text{q-tr}}(\mathbf{p})$ as the vector of posteriors conditional on being pivotal and on receiving message $m \in M$ that arises given the pivotal probabilities calculated via (19) and when $\hat{q}(\omega_j; \mathbf{p})$ is replaced by $\hat{q}^{\text{tr}}(\omega_j; \mathbf{p})$.

First, by continuity, the self-map on $[0, 1]^{|M|}$ from $\mathbf{p}$ to the "truncation"

[35] In principle, $\bar{n}$ may depend on the fixed-point sequence. However, one can show that there is a universal $\bar{n}$ that applies to all fixed point sequences. This requires more work and is not needed for our purposes.

$\hat{\rho}^{\text{q}-\text{tr}}(\mathbf{p})$ has a fixed point for every n. Second, for any sequence of fixed points $(\mathbf{p}_n)_{n\in\mathbb{N}}$, we show below that

$$\lim_{n\to\infty} \hat{q}(\alpha_1; \mathbf{p}_n) \;>\; \frac{1}{2} \text{ and} \tag{51}$$

$$\lim_{n\to\infty} \hat{q}(\beta_1; \mathbf{p}_n) \;<\; \frac{1}{2}. \tag{52}$$

This implies that the fixed points are interior for n sufficiently large. Hence, the corresponding strategies $\mathbf{p}_n$ are equilibria.

We prove (51) and (52) by contradiction. For (51), suppose that $\lim_{n\to\infty} \hat{q}(\alpha_1; \mathbf{p}_n) = \frac{1}{2}$. Then Lemma 1 (which established the equal-margins condition for the substates ω_2) together with Claim 2 implies that $\Pr(\alpha|s, a, \text{piv}; \mathbf{p}_n) \to 1$ for $s \in S$ as $n \to \infty$. Thus, $\hat{q}(\alpha_1; \mathbf{p}_n) \to \Phi(1^-)$. However, this contradicts the initial assumption, since $\Phi(1^-) > \frac{1}{2}$. A similar argument applied to the assumption $\lim_{n\to\infty} \hat{q}(\beta_1; \mathbf{p}_n) = \frac{1}{2}$ also leads to a contradiction, establishing (52).

Finally, for these equilibrium sequences, an application of the law of large numbers implies that A is elected with probability converging to 1 in α_1, and B is elected with probability converging to 1 in β_1, as $n \to \infty$. Since the substates α_1 and β_1 are realized with probability converging to 1 as $n \to \infty$, the claim of the theorem follows.

D General Preferences: Theorem 3

We state an analog of Theorem 1 for a model in which the voter's payoffs can depend on the state in a general way. Except for these more general preferences, the model from Section 1 is unchanged. In particular, we maintain our assumptions on the prior p_0, nature's signal π^S, and the sender's signal.

We define a voter's preference type as a vector $t = (u(x, \omega))_{x\in\{A,B\},\omega\in\{\alpha,\beta\}}$, where $u(x, \omega) \in \mathbb{R}$ denotes the voter's payoff when x is elected in ω (so $t \in \mathbb{R}^4$). For example, a voter is a partisan for A if $u(A, \omega) > u(B, \omega)$ for $\omega \in \{\alpha, \beta\}$ and a partisan for B if the reverse inequalities hold. As in the basic model, we consider symmetric equilibria and we assume that partisans vote for their preferred policy no matter their private signal.

The distribution of preference types $t \in \mathbb{R}^4$ is denoted by G. Given this distribution, the probability that a voter draws some t for which she prefers A to B given a belief p that the state is α is

$$\hat{\Phi}(p) := \Pr(\{t : p(u(A, \alpha) - u(B, \alpha)) \geq (1 - p)(u(B, \beta) - u(A, \beta))\}).$$

The function $\hat{\Phi}$ takes essentially the role of the cumulative distribution function Φ from the basic model. Note that $\hat{\Phi}$ is not a cumulative distribution function, and we can have $\hat{\Phi}(1) < 1$. More importantly, $\hat{\Phi}$ may be non-monotone.

Equilibrium can again be characterized via vectors of induced prior beliefs. Recall that for the basic model, we showed that an equilibrium corresponds to a fixed point $\mathbf{p}^*$ of the mapping $\hat{\boldsymbol{\rho}}(\cdot)$ defined by Equations (15), (16), and (18)–(20). Now, with a general preference distribution G, the relevant observation is that the vote share implied by an induced prior vector only depends on the preference distribution through $\hat{\Phi}$. So, to represent general preferences, we can simply replace $\Phi(p^-)$ by $\hat{\Phi}(p)$ in Equations (15) and (16); an equilibrium then corresponds to a fixed point of the resulting mapping $\hat{\boldsymbol{\rho}}(\cdot)$.

We assume that the distribution G of preference types t is atomless, and that, as in the basic model, the following conditions hold:

G1. The function $\hat{\Phi}$ satisfies the richness condition (4): $\hat{\Phi}(0) < 1/2 < \hat{\Phi}(1)$.

G2. The expected share of partisans for both A and B is positive.

Condition G1 implies that the full-information outcome is A in α and B in β. If the full-information outcome were the same in both states, there would be no meaningful information aggregation problem, nor would it be possible to generate a majority for either alternative using extreme beliefs, as in the proof of Theorem 1. Condition G2 implies that the vote shares under full information are interior, $\hat{\Phi}(0) > 0$ and $\hat{\Phi}(1) < 1$, thus trivial equilibria are excluded.

We now re-state the central condition for full persuasion from the main text: When the sender provides full information, the margins of victory must be larger than when he provides no information. Formulated in terms of induced priors:

M. The prior, preference distribution G, and nature's signal π^S must be such that

$$\min\left\{\frac{1}{2} - \hat{\Phi}(0), \hat{\Phi}(1) - \frac{1}{2}\right\} > \limsup_{n\to\infty} \left|\hat{q}(\omega; \mathbf{p}_n^*) - \frac{1}{2}\right|$$

for any $\omega \in \{\alpha, \beta\}$ and any sequence of equilibrium induced priors $(\mathbf{p}_n^*)_{n\in\mathbb{N}}$ given an uninformative sender's signal (that is, each $\mathbf{p}_n^*$ corresponds to an equilibrium in the original Condorcet jury theorem setting).

We adopt two regularity conditions from Bhattacharya (2013) on the derivative of $\hat{\Phi}$:[36]

[36] Condition R1 is implied by the conditions A1 and A2 in Bhattacharya (2013). Condition R2 requires that any belief $p \in (0, 1)$ that satisfies (35) is "regular" in his terminology.

R1. $\hat{\Phi}$ is continuously differentiable, $\hat{\Phi}'(0) \neq 0$ and $\hat{\Phi}'(1) \neq 0$, and there is no open interval over which $\hat{\Phi}'(p) = 0$.

R2. For any $p \in (0,1)$ satisfying (35) (with $p = \lambda$), the following hold for the implied derivatives of $\hat{q}(\omega, p)$:

- If $\hat{q}(\alpha, p) = \hat{q}(\beta, p)$, then $\hat{q}'(\alpha, p) \neq \hat{q}'(\beta, p)$.
- If $\hat{q}(\alpha, p) \neq \hat{q}(\beta, p)$, then $\theta'^*(p) \neq 0$, where $\theta^*(p)$ is the implicit solution to $x^\theta(1-x)^{1-\theta} = y^\theta(1-y)^{1-\theta}$ with $x = \hat{q}(\alpha, p)$ and $y = \hat{q}(\beta, p)$.

We are now ready to state the analog of Theorem 1 for general preferences.

Theorem 3 *For any state-dependent policy $(x(\alpha), x(\beta)) \in \{A, B\}^2$, there exists a sequence of sender's signals $(\pi_n^M)_{n\in\mathbb{N}}$ such that for any environment (p_0, G, π^S) satisfying Conditions G1, G2, R1, R2, and M, there is an equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ given $\pi_n = \pi^S \times \pi_n^M$ that yields the target policy:*

$$\lim_{n\to\infty} \Pr(x(\omega) \textit{ wins} \mid \omega; \sigma_n^*, \pi_n, n) = 1 \qquad \textit{for } \omega \in \{\alpha, \beta\}. \tag{53}$$

To prepare for the proof, we provide two conditions, stated in terms of primitives (the preference function $\hat{\Phi}$ and nature's signal π^S) that are jointly equivalent to Condition M, which was stated in terms of equilibrium objects. The first condition is that $\hat{\Phi}$ is locally monotone at the extreme beliefs:

$$\exists\ \varepsilon > 0 : \hat{\Phi} \text{ is strictly increasing on } \ [0, \varepsilon] \text{ and on } [1-\varepsilon, 1]. \tag{54}$$

Given the monotonicity condition (54), the continuity of $\hat{\Phi}$, and Condition G1, it follows from the intermediate value theorem that there exists at least one induced prior $\lambda \in (0,1)$ such that the implied vote shares have equal margins of victory,[37]

$$\left|\hat{q}(\alpha; \lambda) - \frac{1}{2}\right| = \left|\hat{q}(\beta; \lambda) - \frac{1}{2}\right|; \tag{55}$$

[37] Recall the definition of the implied vote share for the general-preferences model, that is, $\hat{q}(\omega; p) = \mathrm{E}(\hat{\Phi}(\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)}))$ for all $p \in (0,1)$. Given (2), for any $p \in (0,1)$, the distribution of the posteriors $\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)}$ in α (as implied by the distribution of s) strictly first-order stochastically dominates the distribution of the posteriors $\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)}$ in β. Also, given (2), there is ε' such that $\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)} \leq \varepsilon$ for all $p \leq \varepsilon'$ and $s \in S$ and $\frac{p\Pr(s|\alpha)}{p\Pr(s|\alpha)+(1-p)\Pr(s|\beta)} \geq 1-\varepsilon$ for all $p \geq 1-\varepsilon'$. Since $\hat{\Phi}$ is strictly increasing on $[0,\varepsilon]$ and $[1-\varepsilon, 1]$ by (54), we obtain that $\hat{q}(\alpha; p) > \hat{q}(\beta; p)$ for all $p \leq \varepsilon'$ and all $p \geq 1-\varepsilon'$. Then, G1 implies $|\hat{q}(\alpha; p) - \frac{1}{2}| < |\hat{q}(\beta; p) - \frac{1}{2}|$ when $p \approx 0$ and $|\hat{q}(\alpha; p) - \frac{1}{2}| > |\hat{q}(\beta; p) - \frac{1}{2}|$ when $p \approx 1$. Thus, the claim follows from the intermediate value theorem and the continuity of $\hat{\Phi}$ (which implies the continuity of $\hat{q}(\omega, -)$).

see Figure 6. The second condition is that the margin of victory under full information is larger than the margin implied by any induced prior $\lambda \in (0,1)$ satisfying (55):

$$\min\left\{\frac{1}{2}-\hat{\Phi}(0), \hat{\Phi}(1)-\frac{1}{2}\right\} > \left|\hat{q}(\alpha;\lambda)-\frac{1}{2}\right| \text{ for all } \lambda \in (0,1) \text{ satisfying (55).} \tag{56}$$

Lemma 2 *Given G1, G2, R1 and R2, the two conditions* (54) *and* (56) *are jointly equivalent to Condition M.*

Remark. When $\hat{\Phi}$ is increasing, (54) and (56) hold. This is immediate for (54). For (56), it follows because there is a unique interior λ satisfying (55), which gives rise to the margins of size $D < \min\left\{\frac{1}{2}-\hat{\Phi}(0), \hat{\Phi}(1)-\frac{1}{2}\right\}$ from Claim 4. When $\hat{\Phi}$ is not monotone, (56) can fail; see the right panel of Figure 6. In particular, (55) may have multiple solutions.

Proof. Condition M is a condition on sequences of equilibria given an uninformative sender's signal. For this case, our setting specializes to the one in Bhattacharya (2013).[38] Thus, we can directly utilize his equilibrium characterizations (in particular, his Lemma 3). Specifically, his results imply that given G1, G2, R1, and R2, there is a sequence of equilibrium induced prior vectors $(\mathbf{p}_n^*)_{n\in\mathbb{N}}$ converging to $\lambda \in [0,1]$ if and only if

(i) $\lambda = 0$ and $\hat{\Phi}'(0) < 0$;

(ii) $\lambda = 1$ and $\hat{\Phi}'(1) < 0$;

(iii) $\lambda \in (0,1)$ and it satisfies (55).

Having made these observations, we now show that M is equivalent to (54) and (56) given G1, G2, R1, and R2.

Condition M implies (54): Suppose not; suppose (54) fails. Recalling R1, we see that $\hat{\Phi}$ is continuously differentiable and $\hat{\Phi}'(0) \neq 0$, $\hat{\Phi}'(1) \neq 0$. So, if (54) fails, this implies that either $\hat{\Phi}'(0) < 0$, $\hat{\Phi}'(1) < 0$, or both. Therefore, the observations (i) and (ii) above imply that there exists a sequence of equilibria with induced priors converging to 0 and/or 1. But then M cannot hold, a contradiction.

Condition M implies (56): This follows from the observation (iii) because, if $\lambda \in (0,1)$ satisfies (55), then there is a sequence of equilibria with induced priors converging to it.

[38] The only difference is that Bhattacharya (2013) assumes that nature's signal has binary realizations, but one can easily check that this difference is inconsequential.

Condition M is implied by (54) and (56): By (54) and the observations (i) and (ii), all limits of sequences of equilibrium induced priors must be interior and, therefore, satisfy (55). Hence, (56) implies M. ■

Proof of Theorem 3. We show that there exist a sender's signal and an equilibrium sequence that yield a given target outcome whenever the conditions (54) and (56) hold. Given Lemma 2 and the regularity conditions, these properties are implied by Condition M of Theorem 3.

The proof closely follows the original proof of Theorem 1. The previous setting differs from the current one only in the assumption of the (global) monotonicity of $\hat{\Phi}$, which was used only in the statement and proof of Claim 4.

We therefore start by stating and proving a modified version of Claim 4. Claim 4 established that the margins of victory equalize in substates α_2 and β_2:

$$\lim_{n\to\infty}\left|\hat{q}(\alpha_2;\mathbf{p}_n)-\frac{1}{2}\right| = \lim_{n\to\infty}\left|\hat{q}(\beta_2;\mathbf{p}_n)-\frac{1}{2}\right|. \tag{57}$$

With monotone $\hat{\Phi}$, this pins down the limit of the induced priors p_z^n as the unique solution to (55). With non-monotone $\hat{\Phi}$, there may be multiple solutions to (55). Thus, we have to drop the uniqueness statement.

Moreover, the monotonicity of $\hat{\Phi}$ was used to rule out equilibrium sequences in which the induced priors p_z converge to 0 or 1. We show that local monotonicity at the extremes (the condition (54)) is sufficient for this purpose, given the regularity conditions.

Claim 7 *Let the sender's signal be* $(\pi_n^M(B,A))_{n\in\mathbb{N}}$, $(\pi_n^M(A,A))_{n\in\mathbb{N}}$, *or* $(\pi_n^M(B,B))_{n\in\mathbb{N}}$. *If* (54), *G1, and G2 hold, and* $\hat{\Phi}$ *is continuous, then, for any sequence of induced prior vectors* $\mathbf{p}_n = (p_m^n)_{m\in M}$ *for which* $p_z^n = \hat{\rho}_z(\mathbf{p}_n)$ *for all* n,

$$\lim_{n\to\infty}\left|q(\alpha_2;\mathbf{p}_n)-\frac{1}{2}\right| = \lim_{n\to\infty}\left|q(\beta_2;\mathbf{p}_n)-\frac{1}{2}\right|. \tag{58}$$

and

$$\lim_{n\to\infty} p_z^n \notin \{0,1\}. \tag{59}$$

Proof. In the proof of Claim 4, the strict monotonicity of $\hat{\Phi}$ was used to show that the vote shares in α and β given any interior induced prior p are ordered as $\hat{q}(\alpha,p) > \hat{q}(\beta,p)$; see (26). With general (non-monotone) $\hat{\Phi}$, this ordering may fail for some p. However, when (54) holds, the vote shares are still ordered for interior

p close enough to 0 and 1. More precisely, there is some $\varepsilon' > 0$ such that

$$\hat{q}(\alpha; p) > \hat{q}(\beta; p) \text{ for all } 0 < p \leq \varepsilon' \text{ and } 1 > p \geq 1 - \varepsilon'; \tag{60}$$

see Footnote 37 for a detailed argument. This observation replaces (26). Given (60), the argument following (26) can be used verbatim to establish (59). Finally, given (59), (58) follows from Claim 2. ■

In the equilibrium construction used to prove Theorem 1, we applied Claim 4 to argue that, for $n \to \infty$, the limits of the margins of victory in states α_2 and β_2 are necessarily smaller than those that would have obtained under full information. In the present setting, Claim 7 and (56) and $\lim_{n\to\infty} q(\alpha_2; \mathbf{p}_n) = q(\alpha, \lambda)$ for $\lambda = \lim_{n\to\infty} p_z^n$ imply the same thing. Therefore, the equilibrium construction in Section B.2.4 can be replicated exactly here, finishing the proof of Theorem 3. ■

E Proof of Theorem 4 (Remaining Cases): Partial Commitment

Consider the constant target policy $x(\alpha) = x(\beta) = A$. (The proof for the case $x(\alpha) = x(\beta) = B$ is completely analogous.)

Take the sender's signal $\pi_n^M(A, A)$ from Figure 4. As shown in the proof of Theorem 1, with full commitment, there are equilibria σ_n^* that yield the target policy. The voting behavior is such that the vote share of A among voters with message a or b becomes arbitrarily close to $\Phi(1^-)$ in both states. In contrast, the vote share of A after z is consistent with the Condorcet jury theorem: In α, it converges to a number strictly between $\frac{1}{2}$ and $\Phi(1^-)$, and in β to one strictly between $\Phi(0)$ and $\frac{1}{2}$ (cf. Lemma 1 and (45), (48) and (49)). Given a slight modification of $\pi_n^M(A, A)$, denoted by $\tilde{\pi}_n^M(A, A)$, in which the messages a and b are identified with each other, there is an equilibrium $\tilde{\sigma}_n^*$ in which the voting behavior exhibits the same properties. Given this voting behavior and the target policy $x(\alpha) = x(\beta) = A$, the sender's best response ψ^* is to send message a to all voters in both states.

Now, for any $\chi > 0$ and n large enough, we construct a modified sender's signal $\tilde{\pi}_n^\chi$ such that $\tilde{\pi}_n^\chi$, χ, and ψ^* jointly imply the same signal distribution as $\tilde{\pi}_n^M(A, A)$. Then $\tilde{\sigma}_n^*$ is a best response of the voters to this signal, and $(\tilde{\pi}_n^\chi, \psi_n^*, \tilde{\sigma}_n^*)_{n\in\mathbb{N}}$ is a χ-equilibrium that yields the constant policy A.

The sender's signal $\tilde{\pi}_n^\chi$ is constructed as follows: In both states, he sends message a to all voters with probability r, where r solves $\chi r + (1-\chi) = 1-\varepsilon$. In α, otherwise, each voter receives message z with probability $1-\varepsilon^2$ and message a with probability ε^2, independently of all the other voters. In β, otherwise, each voter receives message z. This construction is feasible if $\chi > \varepsilon$, which ensures that r is in $(0,1)$. One can check that $\tilde{\pi}_n^\chi$, χ, and ψ^* indeed jointly imply the same signal distribution as $\tilde{\pi}_n^M(A, A)$.

Online Appendix

F Proof of the Condorcet Jury Theorem

We use the representation of equilibrium in the belief space that was introduced in Appendix A: Equilibrium is characterized by the fixed-point equation (14).

We adopt the expressions for the equilibrium quantities in terms of induced priors from Section B.1. Since the sender's signal π^M is uninformative, the posterior conditional on being pivotal $\Pr(\alpha|\text{piv}, m)$ is the same for all messages m. Therefore, it is without loss to consider only induced prior vectors $\mathbf{p} = (p_1, \ldots, p_m)$ where all components are equal, $p = p_i$ for all $i = 1, \ldots, m$, and we identify such induced prior vectors $\mathbf{p}$ with p. In particular, given any p, the probability that a voter with induced prior p has a realized type and signal such that she prefers A is

$$\begin{aligned}
\hat{q}(\omega; p) &= \sum_{s \in S} \Pr(s|\omega) \Phi\left(\frac{p \Pr(s|\alpha)}{p \Pr(s|\alpha) + (1-p) \Pr(s|\beta)}\right) \quad \text{for } p \in [0, 1), \\
\hat{q}(\omega, 1) &= \Phi(1^-),
\end{aligned}$$

as in (15) and (16).

The vote share $\hat{q}(\omega; p)$ determines the probability of being pivotal in ω, denoted by $\Pr(\text{piv}|\omega; p)$, and the posterior conditional on being pivotal, denoted by $\hat{\rho}(p)$. These are defined by analogs of (19) and (20) respectively, where we replace the substate ω_j with the state ω in all instances.

To prepare for the proof, we restate our previous observations regarding a voter's inference from being pivotal. Then, in Step 2, we show that voters cannot become certain of the state when pivotal; specifically, the equilibrium induced prior does not converge to 0 or 1. From the fact that the equilibrium induced prior must be interior, two key points follow: (i) The limit vote share of A is strictly higher in α than in β (Step 3), and (ii) the limits of the margins of victory must be equal (Step 4). Combining these observations, we conclude that the margin of victory for A must be positive in α and negative in β, implying the Condorcet jury theorem (Theorem 0) for large n by the law of large numbers.

F.1 Inference from Being Pivotal

We restate Claims 1 and 2, which extend immediately to the case of an uninformative sender's signal if we replace the substates ω_j, ω_i by the states α and β.

Claim 8 *Take any induced prior p for which* $\Pr(\text{piv}|\beta;p) \in (0,1)$*; if*

$$\left|\hat{q}(\alpha;p) - \frac{1}{2}\right| < (>) \left|\hat{q}(\beta;p) - \frac{1}{2}\right|, \tag{61}$$

then

$$\frac{\Pr(\text{piv}|\alpha;p)}{\Pr(\text{piv}|\beta;p)} > (<)\, 1. \tag{62}$$

Claim 9 *Consider any sequence of induced priors* $(p_n)_{n\in\mathbb{N}}$ *for which* $\Pr(\text{piv}|\beta;p_n) \in (0,1)$ *for all n. If*

$$\lim_{n\to\infty} \left|\hat{q}(\alpha;p_n) - \frac{1}{2}\right| < \lim_{n\to\infty} \left|\hat{q}(\beta;p_n) - \frac{1}{2}\right|, \tag{63}$$

then, for any $d \geq 0$,

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;p_n)}{\Pr(\text{piv}|\beta;p_n)} n^{-d} = \infty. \tag{64}$$

F.2 Proof

Step 1 *For all n and every equilibrium induced prior p_n^*, the vote share of A is larger in α than in β:*

$$0 < \hat{q}(\beta;p_n^*) < \hat{q}(\alpha;p_n^*) < 1. \tag{65}$$

The ordering $\hat{q}(\beta;p_n^*) < \hat{q}(\alpha;p_n^*)$ follows from (15) because Φ is strictly increasing and the distribution of the posteriors $\Pr(\alpha|s,\text{piv};p_n^*)$ in α (as implied by the distribution of s) strictly first-order stochastically dominates the distribution of the posteriors in β. The fact that $0 < \hat{q}(\omega;p_n^*) < 1$ for $\omega \in \{\alpha,\beta\}$ holds because $0 < \Phi(0)$ and $\Phi(1^-) < 1$.

Step 2 *Voters cannot become certain of the state conditional on being pivotal; that is, the inference from the pivotal event must remain bounded:*

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;p_n^*)}{\Pr(\text{piv}|\beta;p_n^*)} \in (0,\infty). \tag{66}$$

Suppose not; suppose instead, for example, that conditional on being pivotal, voters become convinced that the state is β, i.e., $\eta := \lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;p_n^*)}{\Pr(\text{piv}|\beta;p_n^*)} = 0$. This would imply $\lim_{n\to\infty} \Pr(\alpha|s, \text{piv}; p_n^*) = 0$ for $s \in S$. Then, given $\Phi(0) < \frac{1}{2}$, a strict majority would support B in both states. However, given the ordering of the vote shares in (65), the election is then closer to being tied in state α. Therefore, Claim 8 implies that voters would update toward state α conditional on being pivotal, in contradiction to the hypothesis that the induced prior converges to $\eta = 0$.

Step 3 *The limit of the vote share of A is larger in α than in β:*

$$\lim_{n\to\infty} \hat{q}(\alpha; p_n^*) > \lim_{n\to\infty} \hat{q}(\beta; p_n^*). \tag{67}$$

By (2), (12), and (66), the limits of the posteriors conditional on the pivotal event and on $s \in S$ are interior. So (67), like (65), follows from Φ being strictly increasing.

Step 4 *In both states, the election is equally close to being tied in expectation as $n \to \infty$; that is,*

$$\lim_{n\to\infty} \hat{q}(\alpha; p_n^*) - \frac{1}{2} = \lim_{n\to\infty} \frac{1}{2} - \hat{q}(\beta; p_n^*). \tag{68}$$

Since, by (66), voters must not become certain conditional on being pivotal, Claim 9 requires that

$$\lim_{n\to\infty} \left| \hat{q}(\alpha; p_n^*) - \frac{1}{2} \right| = \lim_{n\to\infty} \left| \hat{q}(\beta; p_n^*) - \frac{1}{2} \right|. \tag{69}$$

Given the ordering of the limits of the vote shares from (67), (69) implies (68).

It follows from (67) and (68) that

$$\lim_{n\to\infty} \hat{q}(\alpha; p_n^*) > \frac{1}{2} > \lim_{n\to\infty} \hat{q}(\beta; p_n^*).$$

Therefore, given the conditional independence of the signals s, the weak law of large numbers implies that A wins in state α with probability converging to 1 as $n \to \infty$, and B wins in state β with probability converging to 1 as $n \to \infty$. This proves Theorem 0.

Remark. By the same reasoning, Theorem 0 also holds when we allow the information structure π^S to vary with n (keeping the signal set S fixed), as long as the limit information structure is not completely uninformative—that is, as long as it satisfies (2).

G Sketch of the Sender's Signal for a Continuum of States

Consider the following setting, which is based on Feddersen and Pesendorfer (1997): A state $\omega \in [0,1]$ is drawn from a distribution with a continuous, strictly positive density. The voters' preference types $y \in [0,1]$ are drawn independently across voters and independent of the state, with the distribution of y also admitting a continuous, strictly positive density. If a voter changes the outcome from B to A, her payoff is $v(y,\omega) = u(y,A,\omega) - u(y,B,\omega)$, which is continuous and strictly increasing in the state ω and the voter's preference type y. The extreme types are partisans, with $v(1,0) > 0$ and $v(0,1) < 0$. The distribution of y is such that under full information (i.e., if the state ω were known), in expectation a strict majority of types would prefer B if $\omega = 0$ and A if $\omega = 1$. By continuity, this means that there is some interior cutoff state such that there is in expectation a majority for A if ω is above the cutoff and for B if it is below. Without loss of generality, we can assume the cutoff state is $\omega = 0.5$. Voters receive private signals $s \in \{u, d\}$ that are i.i.d. across voters conditional on the state ω, with $\frac{\Pr(u|\omega)}{\Pr(d|\omega)} \in (0,\infty)$ strictly increasing in ω. The voters simultaneously decide whether to vote A or B based on their private preference types y and their private signals s; the outcome is then chosen by simple majority rule. Feddersen and Pesendorfer (1997) show that for any sequence of symmetric Bayesian Nash equilibria indexed by the electorate size n, for $n \to \infty$ the full-information outcome is chosen with probability converging to 1; that is, A wins a majority in state ω if and only if $\omega > 0.5$.

We show that persuasion is still effective in this setting by constructing a sender's signal and a sequence of equilibria for which the majority outcome is the opposite of the full-information outcome. The sender's signal is as follows. Take any intervals $[b_1, b_2] \subset (0, 0.5)$ and $[a_1, a_2] \subset (0.5, 1)$ of length $\varepsilon > 0$ each. With probability $1 - \varepsilon$, all voters receive message a when $\omega > 0.5$ and message b when $\omega \leq 0.5$. With probability ε, the message distribution is as follows: In all states, except on the intervals $[b_1, b_2]$ and $[a_1, a_2]$, all voters receive message z. For states in $[b_1, b_2]$, voters receive message a with some probability ε' and z otherwise, and for states in $[a_1, a_2]$, voters receive message b with probability ε' and z otherwise. In summary, on the intervals $[b_1, b_2]$ and $[a_1, a_2]$, there is a chance ε that most voters receive an essentially uninformative message z while the rest receive message a or b. Outside these intervals, all of the voters always receive the same message, which is either z or a (when $\omega > 0.5$) or b (when $\omega \leq 0.5$).

For ε' small enough, the margin of victory implied by the behavior of the voters with message z is close to 0. The reasoning is analogous to that in our basic model. Voters who receive message z have almost no additional information about the state, and they know that almost all of the other voters have also received message z. If ε' were equal to 0, they would know that *all* of the voters received message z and z would provide *no* information about the state. Then, from their perspective, the setting would be exactly the same as the setting of Feddersen and Pesendorfer (1997) with no additional information. For that setting, Feddersen and Pesendorfer (1997) show that the margin of victory vanishes in all equilibria. We construct our sender's signal so that ε' goes to 0 rapidly enough for this observation to remain valid.

For the voters with message a or b, one can construct an equilibrium in which a strict majority of them vote the inverse of their message (B after a and A after b), provided n is sufficiently large and $\varepsilon, \varepsilon'$ are sufficiently small, with b_2 and a_1 bounded away from 0.5. This behavior is rationalized by the fact that conditional on being pivotal, a voter with message a believes that the state is almost surely in $[b_1, b_2]$ and a voter with message b believes it is almost surely in $[a_1, a_2]$. Since b_2 and a_1 are bounded away from 0.5, this implies strictly positive margins of victory for B among voters with message a and for A among voters with message b. These beliefs conditional on being pivotal result from the relative margins of victory: For states outside $[b_1, b_2]$ and $[a_1, a_2]$, in the case where all voters receive message a or message b, their voting behavior implies a strict majority for B or A, respectively, and thus a strictly positive margin of victory. By contrast, in the states $[b_1, b_2]$ and $[a_1, a_2]$, there is a chance that most other voters have received message z, in which case the margin of victory is close to 0.

The behavior of the voters with message a or b implies that for this sender's signal, the inverse of the full-information outcome receives a majority with probability close to 1 when n is large and ε is small.

The logic above can also be used to achieve more complicated state-dependent outcomes. For example, suppose that we keep the structure otherwise identical but now, with probability $1 - \varepsilon$, all voters receive message a when $\omega \in [0, 0.1) \cup [0.2, 0.3) \cup \ldots \cup [0.8, 0.9)$ and message b when $\omega \in [0.1, 0.2) \cup [0.3, 0.4) \cup \ldots \cup [0.9, 1.0]$. Then, in equilibrium, the majority will alternate between A and B as the state moves from 0 to 1. We conjecture that a similar information structure can yield any outcome that does not change too irregularly as the state changes.

H Numerical Example: A Jury

We give an example showing that our persuasion mechanism can be effective when there are as few as 11 voters, as in a standard jury. For this example, we use the sender's signal $\pi_n^M(A,A)$ from Figure 4, with $\varepsilon = \frac{1}{3n}$. We assume there are $2n+1$ voters. Furthermore, we let $\Phi(p) = p$ for all $p \in [0,1]$, $p_0 = \frac{1}{3}$, and the exogenous private signal is binary, $s \in S = \{u, d\}$, with $r = \Pr(u|\alpha) = \Pr(d|\beta) = \frac{2}{3}$. There are no partisans in this example.

We show that under these primitives, when $2n+1 \geq 11$, there is an equilibrium σ_n^* in which A is elected with probability greater than 99.99% in substates α_1 and β_1. Therefore, the overall probability of A being elected exceeds $0.9999\Big[p_0(1 - \frac{1}{3n}) + (1-p_0)(1-\frac{1}{3n})\Big]$, which is larger than 93% for $n \geq 5$.

Proof. We will construct an equilibrium in terms of voter beliefs as described by vectors $(p_{s,m})_{(s,m)\in S\times M}$. (This is similar to our construction in terms of beliefs $(p_m)_{m\in M}$ in Appendix A.)

Specifically, we consider the set of beliefs $\mathbf{p} \in (p_{s,m})_{(s,m)\in S\times M}$ for which

$$p_{s,m} \geq \frac{99}{100} \text{ for } (s,m) \in S \times \{a,b\}, \tag{70}$$

$$\frac{1}{2} \leq p_{u,z} \leq \frac{5}{6}, \tag{71}$$

$$\frac{1}{6} \leq p_{d,z} \leq \frac{1}{2}, \tag{72}$$

and look for some $\mathbf{p}$ for which $\hat{\sigma}^{\mathbf{p}}$ is an equilibrium. Here $\hat{\sigma}^{\mathbf{p}}$ is again the voter strategy that is optimal given $\mathbf{p}$, that is, when a voter with signal (s,m) believes the probability of α is given by $p_{s,m}$ (cf. (13)).

It is useful to write down the bounds above in terms of likelihood ratios:

$$\frac{p_{s,m}}{1-p_{s,m}} \geq 99 \text{ for } (s,m) \in S \times \{a,b\}, \tag{73}$$

$$1 \leq \frac{p_{u,z}}{1-p_{u,z}} \leq 5, \tag{74}$$

$$\frac{1}{5} \leq \frac{p_{d,z}}{1-p_{d,z}} \leq 1. \tag{75}$$

For any $\mathbf{p}$, we consider the posteriors conditional on being pivotal and on the pairs (s,m), provided that all other voters behave according to $\hat{\sigma}^{\mathbf{p}}$; that is, we consider $\hat{\rho}(\mathbf{p}) = (\Pr(\alpha|s,m,\text{piv};\mathbf{p}))_{(s,m)\in S\times M}$. In analogy to (43), we denote by $\hat{\boldsymbol{\rho}}^{\text{tr}}$ the truncation of $\hat{\boldsymbol{\rho}}$ to the set of beliefs satisfying (70)–(72). We call the mapping $\mathbf{p} \mapsto \hat{\boldsymbol{\rho}}^{\text{tr}}(\mathbf{p})$ the "truncated best reponse."

Below, we show that any fixed point of the truncated best response is interior when $2n+1 \geq 11$. Hence, the corresponding strategy $\hat{\sigma}^{\mathbf{p}}$—the uniquely optimal one when a voter with signal (s,m) believes that the probability of α is $p_{s,m}$—is an equilibrium strategy.

Since such an equilibrium satisfies the bounds above, the likelihood that A gets elected in substates α_1 and β_1 can be calculated to be at least 99.99 percent when there are at least $2n+1=11$ voters, which proves the claim for this example.

Before we do this, we state an auxiliary result which guarantees that when $n \geq 5$, the probability of receiving message z is high enough so that a key observation from the proof of the Condorcet jury theorem applies. Let $\hat{q}(\omega_i;\mathbf{p})$ be the vote share of A in ω_i given $\hat{\sigma}^{\mathbf{p}}$.

Claim 10 *Consider any fixed point* $\mathbf{p}$ *of the truncated best response* $\hat{\boldsymbol{\rho}}^{\mathrm{tr}}(\cdot)$. *Then*

$$\hat{q}(\beta_2;\mathbf{p}) < \hat{q}(\alpha_2;\mathbf{p}) \tag{76}$$

for any $n \geq 5$.

Proof. First, suppose that $\frac{\Pr(\alpha|d,z,\mathrm{piv};\mathbf{p})}{\Pr(\beta|d,z,\mathrm{piv};\mathbf{p})} \leq \frac{1}{5}$. Then $p_{d,z}=\frac{1}{6}$ since $\mathbf{p}$ is a fixed point of $\hat{\boldsymbol{\rho}}^{\mathrm{tr}}(\cdot)$. Furthermore, $\frac{\Pr(\alpha|u,z,\mathrm{piv};\mathbf{p})}{\Pr(\beta|u,z,\mathrm{piv};\mathbf{p})} \leq \frac{1}{5}(\frac{r}{1-r})^2=\frac{4}{5}$. Hence, $p_{u,z}=\frac{1}{2}$ since $\mathbf{p}$ is a fixed point of $\hat{\boldsymbol{\rho}}^{\mathrm{tr}}(\cdot)$. Thus,

$$p_{u,z}-p_{d,z}=\frac{1}{2}-\frac{1}{6}=\frac{1}{3}.$$

Second, suppose that $\frac{1}{5} \leq \frac{\Pr(\alpha|d,z,\mathrm{piv};\mathbf{p})}{\Pr(\beta|d,z,\mathrm{piv};\mathbf{p})} \leq \frac{1}{2}$. Then $p_{d,z}=\Pr(\alpha|d,z,\mathrm{piv};\mathbf{p})$. Moreover, $\frac{\Pr(\alpha|u,z,\mathrm{piv};\mathbf{p})}{\Pr(\beta|u,z,\mathrm{piv};\mathbf{p})} \leq 2$. Hence, $\Pr(\alpha|u,z,\mathrm{piv};\mathbf{p}) \leq p_{u,z}$. Now let $x=\frac{p_{d,z}}{1-p_{d,z}}$. Then

$$\frac{\partial}{\partial x}\left(\frac{4x}{1+4x}-\frac{x}{1+x}\right) = \frac{4}{(1+4x)^2}-\frac{1}{(1+x)^2}\cdot\frac{x}{1+x}. \tag{77}$$

Furthermore, $\frac{4}{(1+4x)^2}-\frac{1}{(1+x)^2}\frac{x}{1+x}>0$ for $x<\frac{1}{2}$. Since $p_{d,z}=\frac{x}{1+x}$ and $p_{u,z} \geq \frac{4x}{1+4x}$, this shows that

$$p_{u,z}-p_{d,z} \geq \frac{\frac{4}{5}}{1+\frac{4}{5}}-\frac{\frac{1}{5}}{1+\frac{1}{5}}=\frac{4}{9}-\frac{1}{6}=\frac{5}{18}.$$

Third, suppose that $\frac{1}{2} < \frac{\Pr(\alpha|d,z,\mathrm{piv};\mathbf{p})}{\Pr(\beta|d,z,\mathrm{piv};\mathbf{p})} \leq 1$. Then $p_{d,z}=\Pr(\alpha|d,z,\mathrm{piv};\mathbf{p})$. Moreover, $2 < \frac{\Pr(\alpha|u,z,\mathrm{piv};\mathbf{p})}{\Pr(\beta|u,z,\mathrm{piv};\mathbf{p})} \leq 4$. Hence, $\Pr(\alpha|u,z,\mathrm{piv};\mathbf{p})=p_{u,z}$. From (77), we see that $\frac{\partial}{\partial x}(\frac{4x}{1+4x}-\frac{x}{1+x})<0$ for $x>\frac{1}{2}$. Since $p_{d,z}=\frac{x}{1+x}$ and $p_{u,z}=\frac{4x}{1+4x}$, this shows that

$$p_{u,z} - p_{d,z} \geq \frac{2}{3} - \frac{\frac{1}{2}}{1+\frac{1}{2}} = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

Fourth, suppose that $\frac{\Pr(\alpha|d,z,\text{piv};\mathbf{p})}{\Pr(\beta|d,z,\text{piv};\mathbf{p})} \geq 1$. Then $p_{d,z} = \frac{1}{2}$. Furthermore, $\frac{\Pr(\alpha|u,z,\text{piv};\mathbf{p})}{\Pr(\beta|u,z,\text{piv};\mathbf{p})} \geq 4$. Hence $p_{u,z} \geq \frac{4}{5}$. Therefore,

$$p_{u,z} - p_{d,z} \geq \frac{4}{5} - \frac{1}{2} = \frac{3}{10}.$$

Note that $\min\left(\frac{1}{3}, \frac{5}{18}, \frac{1}{3}, \frac{3}{10}\right) = \frac{5}{18} \approx 0.27$. Using this, we can evaluate the difference $\hat{q}(\alpha_2;\mathbf{p}) - \hat{q}(\beta_2;\mathbf{p})$. Note that

$$\begin{aligned}
\hat{q}(\alpha_2;\mathbf{p}) &\geq (1-\varepsilon^2)(r\Pr(A|u,z) + (1-r)\Pr(A|d,z)), \\
\hat{q}(\beta_2;\mathbf{p}) &\leq (1-\varepsilon^2)((1-r)\Pr(A|u,z) + r\Pr(A|d,z)) + \varepsilon^2.
\end{aligned}$$

Thus,

$$\hat{q}(\alpha_2;\mathbf{p}) - \hat{q}(\beta_2;\mathbf{p}) \geq (1-\varepsilon^2)(2r-1)\frac{5}{18} - \varepsilon^2,$$

where we used that $\Pr(A|s,m) = \Phi(p_{s,m}) = p_{s,m}$ for $s \in \{u,d\}$ and that the difference $p_{u,z} - p_{d,z}$ is at least $\frac{5}{18}$. A calculation shows that

$$(2r-1)\frac{5}{18} = \frac{1}{3} \cdot \frac{5}{18} = \frac{5}{54}.$$

Furthermore, for any $n \geq 5$,

$$\frac{5}{54}(1-\varepsilon^2) \geq \frac{5}{54} \cdot \frac{224}{(3n)^2} > \varepsilon^2 = \frac{1}{(3n)^2}.$$

Thus, (76) holds for $n \geq 5$. ■

Claim 11 *Any fixed point of $\hat{\boldsymbol{\rho}}^{\text{tr}}(\cdot)$ is interior when $n \geq 5$.*

The proof is by contradiction. We suppose that there is a non-interior fixed point $\mathbf{p} = (p_{s,m})_{(s,m)\in S\times M}$ and derive a contradiction in each possible case.

Proof of Claim 11.

Case 1 *$p(s,m) = \frac{99}{100}$ for some $(s,m) \in S \times \{a,b\}$.*

Note that a voter with message a learns that the state is α, so $p_{s,a} = 1$ for all $s \in S$. Thus, $p_{s,a}$ cannot be equal to $\frac{99}{100}$ for any fixed point of the truncated best response. In the following, let $s = b$ and suppose that $p(s, b) = \frac{99}{100}$. Then, for any $\mathbf{p}$ satisfying (70)–(72),

$$q(\omega_1; \mathbf{p}) - \frac{1}{2} \geq \frac{49}{100}$$

for $\omega_1 \in \{\alpha_1, \beta_1\}$. Furthermore,

$$\begin{aligned}\left|\hat{q}(\omega_2; \mathbf{p}) - \frac{1}{2}\right| &\leq \max\left\{\frac{2}{3}\left(\frac{1}{2} - \frac{1}{6}\right)(1-\varepsilon^2) + \varepsilon^2, \frac{2}{3}\left(\frac{5}{6} - \frac{1}{2}\right)(1-\varepsilon^2) + \varepsilon^2)\right\} \\ &= \frac{2}{9}(1-\varepsilon^2) + \varepsilon^2\end{aligned}$$

for $\omega_2 \in \{\alpha_2, \beta_2\}$. For $n \geq 5$, we have $\frac{2}{9}(1-\varepsilon^2) + \varepsilon^2 \leq \frac{2}{9} \cdot \frac{224}{225} + \frac{1}{225}$. Therefore,

$$\begin{aligned}\frac{\Pr(\text{piv}|\omega_2; \mathbf{p})}{\Pr(\text{piv}|\omega_1'; \mathbf{p})} &\geq \left(1 + \frac{(\frac{49}{100})^2 - (\frac{2}{9} \cdot \frac{224}{225} + \frac{1}{225})^2}{\frac{1}{4} - (\frac{49}{100})^2}\right)^n \\ &\geq (19)^n\end{aligned}$$

for $n \geq 5$, $\omega_1' \in \{\alpha_1, \beta_1\}$, and $\omega_2 \in \{\alpha_2, \beta_2\}$. As a consequence,

$$\begin{aligned}\frac{\Pr(\alpha|s, m, \text{piv}; \mathbf{p})}{\Pr(\beta|s, m, \text{piv}; \mathbf{p})} &\geq \frac{1}{\varepsilon^3} \cdot \frac{p_0}{1-p_0} \cdot \frac{\Pr(d|\alpha)}{\Pr(d|\beta)} \cdot 19^n \\ &= \frac{1}{(3n)^3} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 19^n.\end{aligned}$$

For $n = 5$, this implies that $\frac{\Pr(\alpha|s,m,\text{piv};\mathbf{p})}{\Pr(\beta|s,m,\text{piv};\mathbf{p})} \geq 183$. Since $\frac{1}{(3n^3)} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 19^n$ is increasing in n, this implies that for $n \geq 5$, $\Pr(\alpha|s, m, \text{piv}; \mathbf{p}) > \frac{99}{100}$. However, this contradicts the initial assumption that $p(s, m) = \frac{99}{100}$.

Case 2 $p_{u,z} = \frac{1}{2}$.

Since $\mathbf{p}$ is a fixed point of $\hat{\rho}(\cdot)$, this implies

$$\frac{\Pr(\alpha|u, z, \text{piv}; \mathbf{p})}{\Pr(\alpha|u, z, \text{piv}; \mathbf{p})} \leq 1.$$

Hence, $\frac{p_{d,z}}{1-p_{d,z}} \leq (\frac{1-r}{r})^2 = \frac{1}{4}$. So $p_{d,z} \leq \frac{1}{5}$. This implies

$$\begin{aligned}\hat{q}(\alpha_2; \mathbf{p}) &\leq \left(r \cdot \frac{1}{2} + (1-r) \cdot \frac{1}{5}\right)(1-\varepsilon^2) + \varepsilon^2 \\ &= \left(\frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{5}\right)\left(1 - \frac{1}{(3n)^2}\right) + \frac{1}{(3n)^2} \\ &= \left(\frac{2}{6} + \frac{1}{15}\right)\left(1 - \frac{1}{(3n)^2}\right) + \frac{1}{(3n)^2}.\end{aligned}$$

Note that $(\frac{2}{6}+\frac{1}{15})(1-\frac{1}{(3n)^2})+\frac{1}{(3n)^2} < \frac{1}{2}$ for $n=3$, since $(\frac{2}{6}+\frac{1}{15})\cdot\frac{80}{81}+\frac{1}{81}\approx 0.407$. Since $(\frac{2}{6}+\frac{1}{15})(1-\frac{1}{(3n)^2})+\frac{1}{(3n)^2}$ is strictly decreasing in n, this implies that $(\frac{2}{6}+\frac{1}{15})(1-\frac{1}{(3n)^2})+\frac{1}{(3n)^2} < \frac{1}{2}$ for $n \geq 3$. Hence, $\hat{q}(\alpha_2;\mathbf{p}) \leq \frac{1}{2}$. Together with (76) from Claim 10, this implies that $\hat{q}(\beta_2;\mathbf{p}) < \hat{q}(\alpha_2;\mathbf{p}) \leq \frac{1}{2}$. Thus, by Claim 8,

$$\Pr(\alpha|z, \text{piv}; \mathbf{p}) > p_0 = \frac{1}{3}.$$

But this implies that $\Pr(\alpha|u, z, \text{piv}; \mathbf{p}) > \frac{1}{2}$, which contradicts the initial assumption that $p_{u,z} = \frac{1}{2}$.

Case 3 $p_{u,z} = \frac{5}{6}$.

We have $\frac{\Pr(\alpha|u,z,\text{piv};\mathbf{p})}{\Pr(\alpha|u,z,\text{piv};\mathbf{p})} \geq 5$ and so $\frac{\Pr(\alpha|d,z,\text{piv};\mathbf{p})}{\Pr(\alpha|d,z,\text{piv};\sigma^{\mathbf{P}})} \geq 5(\frac{(1-r)}{r})^2 = \frac{5}{4}$. Therefore, $p_{d,z} = \frac{1}{2}$ and $\hat{q}(\beta_2;\mathbf{p}) \geq \frac{1}{2}$. Together with Lemma 10, this implies $\hat{q}(\alpha_2;\mathbf{p}) > \hat{q}(\beta_2;\mathbf{p}) \geq \frac{1}{2}$ for $n \geq 5$. As a consequence, $\Pr(\alpha|z, \text{piv}; \mathbf{p}) < p_0 = \frac{1}{3}$, so $\Pr(\alpha|u, z, \text{piv}; \mathbf{p}) < \frac{1}{2}$, which contradicts the initial assumption that $p_{u,z} = \frac{5}{6} > \frac{1}{2}$.

Case 4 $p_{d,z} = \frac{1}{2}$.

The assumption implies that $q(\beta_2;\mathbf{p}) \geq \frac{1}{2}$. Together with Claim 10, this implies $\hat{q}(\alpha_2;\mathbf{p}) > \hat{q}(\beta_2;\mathbf{p}) \geq \frac{1}{2}$ for $n \geq 5$. As a consequence, $\Pr(\alpha|z, \text{piv}; \mathbf{p}) < p_0 = \frac{1}{3}$, so $\Pr(\alpha|d, z, \text{piv}; \mathbf{p}) < \frac{1}{2}$, which contradicts the initial assumption that $p_{d,z} = \frac{1}{2}$.

Case 5 $p_{d,z} = \frac{1}{6}$.

The assumption implies that $\frac{p_{d,z}}{1-p_{d,z}} = \frac{1}{5}$, so $\frac{\Pr(\alpha|d,z,\text{piv};\mathbf{p})}{\Pr(\alpha|d,z,\text{piv};\mathbf{p})} \leq \frac{1}{5}$. Hence, $\frac{\Pr(\alpha|u,z,\text{piv};\mathbf{p})}{\Pr(\alpha|u,z,\text{piv};\mathbf{p})} \leq \frac{4}{5}$, so that $\Pr(\alpha|u, z, \text{piv}; \mathbf{p}) \leq \frac{1}{2}$. Therefore, $p_{u,z} = \frac{1}{2}$ and

$$\hat{q}(\alpha_2;\mathbf{p}) \leq \left(\frac{1}{3}\cdot\frac{1}{6}+\frac{2}{3}\cdot\frac{1}{2}\right)(1-\varepsilon^2)+\varepsilon^2.$$

For $n \geq 4$, we have $\varepsilon \leq \frac{1}{12}$ and $(\frac{1}{3}\cdot\frac{1}{6}+\frac{2}{3}\cdot\frac{1}{2})(1-\varepsilon^2)+\varepsilon^2 = (\frac{1}{18}+\frac{6}{18})(1-\varepsilon^2)+\varepsilon^2 < \frac{1}{2}$. Thus, $\hat{q}(\alpha_2;\mathbf{p}) < \frac{1}{2}$. Together with Claim 10, this implies $\hat{q}(\beta_2;\mathbf{p}) < \hat{q}(\alpha_2;\mathbf{p}) < \frac{1}{2}$ for $n \geq 5$. As a consequence, $\Pr(\alpha|\text{piv}, z; \mathbf{p}) > p_0 = \frac{1}{3}$, and therefore $\frac{\Pr(\alpha|d,z,\text{piv};\mathbf{p})}{\Pr(\alpha|d,z,\text{piv};\mathbf{p})} > \frac{1}{2}\cdot\frac{1-r}{r} = \frac{1}{4}$, which contradicts $\frac{\Pr(\alpha|d,z,\text{piv};\mathbf{p})}{\Pr(\alpha|d,z,\text{piv};\mathbf{p})} \leq \frac{1}{5}$.■ ■

I Numerical Example: Behavioral Types

We give an example showing that our persuasion mechanism can be effective even when a significant proportion of the voters are "sincere," rather than "pivotal" as in the basic model. Specifically, we suppose that each voter is sincere with probability $\kappa = 40\%$ and pivotal with the remaining probability. A sincere voter's behavior depends only on her private information (s, m): she always votes as if her own vote will decide the election. Formally, a sincere voter with threshold of doubt y and signal pair (s, m) votes as follows:

$$\begin{aligned} \Pr(\alpha|s,m) > y \quad &\Rightarrow \text{vote for } A, \\ \Pr(\alpha|s,m) < y \quad &\Rightarrow \text{vote for } B. \end{aligned}$$

This contrasts with the behavior of pivotal voters as described in (5) and (6).

For this example, we use a slight modification of the sender's signal $\pi_n^M(A, A)$ from Figure 4. The modification is denoted by $\tilde{\pi}_n^M(A, A)$, and it arises simply via identifying message b with message a (so, in α_1 and β_1, all voters receive message a, in α_2 they receive message a with probability ε^2 and message z otherwise, and in β_2 they all receive message z). As usual, we assume there are $2n + 1$ voters. Furthermore, we fix the prior $p_0 = 0.4$, let $S = \{u, d\}$ with $r = \Pr(u|\alpha) = \Pr(d|\beta) = 0.6$, and let $\Phi(p) = p$ for all $p \in [0, 1]$. There are no partisans in this example.

We show that under these primitives, when there are at least 199 voters, there is an equilibrium σ_n^* for which A is elected with probability greater than 99%.[39]

Here, it is convenient to construct the equilibria by working in the space of vectors of expected vote shares. For any strategy σ' of the pivotal voters,

$$\hat{q}(\omega_j; \sigma', \kappa) = \kappa \mathrm{E}(\Phi(\Pr(\alpha|m, s)^-)) + (1 - \kappa) \Pr(\{y : \sigma'(s, m, y) = 1\}) \tag{78}$$

is the expected vote share in substate ω_j when each voter is sincere with probability κ, and $\hat{\mathbf{q}}(\sigma') = (\hat{q}(\omega_j; \sigma', \kappa))_{\omega_j \in \{\alpha_1, \alpha_2, \beta_1, \beta_2\}}$ is the vector of the expected vote shares. Given (5) and (6), for the best response σ of the pivotal voters to the strategy profile where the pivotal voters follow σ', we have

$$\hat{q}(\omega_j; \sigma, \kappa) = \kappa \mathrm{E}(\Phi(\Pr(\alpha|m, s)^-)) + (1 - \kappa) \mathrm{E}(\Phi(\Pr(\alpha|m, s; \mathrm{piv}; \sigma', \kappa)^-)) \tag{79}$$

where the posterior $\Pr(\alpha|m, s; \mathrm{piv}; \sigma', \kappa)$ arises by updating from Bayes' rule given the pivotal likelihoods in the substates induced by $q = \hat{q}(\omega_j; \sigma', \kappa)$, as in (19).

[39] Without additional information from the sender, information would aggregate (meaning A wins in α and B in β). However, given the asymmetric prior and the signal precision $r = 0.6$, this would not be the case if the share of sincere voters were too close to 1. The latter follows from two of the observations we make in the paragraphs after Equation (84). The first observation is that the ex-ante expected vote share is 0.4. This implies that if $r = 1$, the vote share in β_2 is smaller than 0.4 as $n \to \infty$. The second observation is that $\psi(p) < 0.05$. This implies that if $r = 1$, the vote share in α_2 is smaller than 0.45 as $n \to \infty$.

Importantly, the vector of vote shares $\hat{\boldsymbol{q}}(\sigma')$ pins down the best response σ entirely; therefore, we can characterize equilibrium in terms of vote share vectors, as follows. Given a vote share vector $\mathbf{q}$, denote by $\sigma^{\mathbf{q}}$ the best response to $\mathbf{q}$. A strategy σ' of the pivotal voters is an equilibrium strategy if

$$\sigma^{\hat{\boldsymbol{q}}(\sigma')} = \sigma',$$

and an equilibrium strategy of the pivotal voters is equivalently characterized by a vector of vote shares $\mathbf{q}^*$ for which

$$\hat{\boldsymbol{q}}(\sigma^{\mathbf{q}^*}) = \mathbf{q}^*. \tag{80}$$

Now, take any $\mathbf{q} = (q(\omega_j))_{\omega_j \in \{\alpha_1,\alpha_2,\beta_1,\beta_2\}}$ with

$$q(\omega_1) \geq 0.75 \quad \text{for} \quad \omega_1 \in \{\alpha_1, \beta_1\}, \tag{81}$$

$$q(\alpha_2) \geq 0.39, \tag{82}$$

$$q(\beta_2) \leq 0.61. \tag{83}$$

We define a constrained version of the best response function $\hat{\mathbf{q}}(\sigma^{(\cdot)})$ as its "truncation" to the domain D defined by (81)–(83): For $\omega_1 \in \{\alpha_1, \beta_1\}$,

$$\hat{q}^{tr}(\omega_1; \kappa, \sigma^{\mathbf{q}}) = \begin{cases} 0.75 & \text{if } q(\omega_1; \kappa, \sigma^{\mathbf{q}}) < 0.75, \\ q(\omega_1; \kappa, \sigma^{\mathbf{q}}) & \text{otherwise.} \end{cases}$$

The components $\hat{q}^{tr}(\alpha_2; \kappa, \sigma^{\mathbf{q}})$ and $\hat{q}^{tr}(\beta_2; \kappa, \sigma^{\mathbf{q}})$ are defined analogously. The function $\hat{q}^{tr}(\cdot)$ is continuous in $\mathbf{q}$, so that Kakutani's fixed point theorem implies that it has a fixed point $\mathbf{q}^* \in D$.

In the following, we show that any fixed point $\mathbf{q}^*$ is interior when $n \geq 99$. So the corresponding strategy $\sigma = \sigma^{\mathbf{q}^*}$ is an equilibrium strategy of the pivotal voters with vote shares $\hat{\mathbf{q}}(\sigma)$ satisfying the constraints (81)–(83).

Since the vote shares in α_1 and β_1 exceed 0.75, an evaluation of the binomial distribution shows that the likelihood that A is elected in these states exceeds 0.999. Since α_1 and β_1 occur with a likelihood close to 1, A is elected with probability close to 1, as claimed in the main text.

Claim 12 *Any fixed point of $\hat{q}^{tr}(\cdot)$ is interior when $n \geq 99$.*

We proceed in five steps.

Step 1 *Take any vote share vector from D. When $n \geq 99$, the difference between the vote shares in α_2 and β_2 under the best response is at least* 0.015 *but at most* 0.045.

Fix $n \geq 99$. Since fewer than 0.1 percent of the voters receive a signal other than z in substates α_2 and β_2, it suffices to show that the vote shares after message z differ by at most 0.04 (since $0.04 + 0.001 < 0.045$), and by at least $0.015 \cdot \frac{100}{99}$.

For this, consider the function

$$\psi(p) = (2q-1)\Big[\frac{pr}{pr + (1-r)(1-r)} - \frac{p(1-r)}{p(1-r)+(1-p)r}\Big], \tag{84}$$

for $p \in (0,1)$. Here $\frac{pr}{pr+(1-p)(1-r)}$ is a voter's posterior belief about the likelihood of α after receiving signals z and u, if the posterior after z alone is p. Similarly, $\frac{p(1-r)}{p(1-r)+(1-p)r}$ is the posterior about the likelihood of α after signals z and d, if the posterior after z alone is p. Given $\Phi(p) = p$, the function (84) describes the difference between the likelihood in substate α_2 and the likelihood in substate β_2 that a random voter with message z will vote A if her belief about the likelihood of α after message z is p. (Sincere and pivotal voters may act upon different p in equilibrium.)

We make three observations about the function (84): First, it takes the maximum value 0.04 at $p = \frac{1}{2}$, given the signal precision $r = 0.6$. This already implies that under the best response to any strategy, the vote shares in α_2 and β_2 will differ by at most 0.045. Second, it is positive. Third, the belief $p = \Pr(\alpha|z)$ that pins down the behavior of sincere voters after receiving z is in $[0.39, 0.41]$ when $n \geq 99$, since

$$\Pr(\alpha|z) = \frac{0.4(1-\frac{1}{n^2})}{0.4(1-\frac{1}{n^2}) + 0.6}. \tag{85}$$

A calculation shows that the function (84) is bounded below by 0.038 for $p \in [0.39, 0.41]$.

The last two observations imply that the vote share of a random voter (who is sincere with likelihood $\kappa = 0.4$) after message z is at least $0.4 \cdot 0.038 > 0.0152$ higher in α_2 than in β_2. Since the share of voters receiving message a is at most $\frac{1}{99^2} < 0.0001$ in α_2 and β_2, the vote share is at least 0.015 larger in α_2 than in β_2 under the best response to any $\mathbf{q} \in D$.

Step 2 *There is no fixed point $\mathbf{q}^*$ with $q^*(\beta_2) = 0.61$ when $n \geq 99$.*

Suppose otherwise. Then, by Step 1, the vote shares under the best response satisfy $q(\alpha_2; \kappa, \sigma^{\mathbf{q}^*}) > q(\beta_2; \kappa, \sigma^{\mathbf{q}^*}) > 0.5$ when $n \geq 99$. Thus, the election is more likely to be tied in β_2 than in α_2. But then $\Pr(\alpha|z) \leq 0.4$ by (85), and $\sum_{s\in\{u,d\}} \Pr(s|\beta_2)\Pr(\alpha|s,z,\text{piv}) < 0.4$, where the $\Pr(\alpha|s,z,\text{piv})$ are the posteriors implied by $\mathbf{q}^*$. Given $\Phi(p) = p$, this means the vote share in β_2 is smaller than 0.41 when $n \geq 99$ (since, for $n \geq 99$, the share of voters receiving message a is smaller than 0.01 in β_2). But this contradicts the initial assumption that the vote share in β_2 would be equal to 0.61.

Step 3 *There is no fixed point* $\mathbf{q}^*$ *with* $q^*(\alpha_2) = 0.39$ *when* $n \geq 99$.

Suppose otherwise. Then, by Step 1, the vote shares under the best response satisfy $0.5 > q(\alpha_2;\kappa;\sigma^{\mathbf{q}^*}) > q(\beta_2;\kappa;\sigma^{\mathbf{q}^*})$ when $n \geq 99$. Thus, the election is more likely to be tied in α_2 than in β_2. But then $\Pr(\alpha|z) \geq 0.39$ by (85), and $\sum_{s\in\{u,d\}} \Pr(s|\beta_2)\Pr(\alpha|s,z,\text{piv}) > 0.4$, where the $\Pr(\alpha|s,z,\text{piv})$ are the posteriors implied by $\mathbf{q}^*$. Given $\Phi(p) = p$, this means the vote share in α_2 is larger than 0.39 when $n \geq 99$ (since the voters receiving message a vote with an ever higher likelihood for A). But this contradicts the initial assumption.

Step 4 *There is no fixed point* $\mathbf{q}^*$ *with* $q^*(\beta_1) = 0.75$ *when* $n \geq 99$.

Suppose otherwise. Given the vote share vector $\mathbf{q}^*$, consider the implied pivotal likelihood ratio

$$\begin{aligned}
&\frac{\Pr(\text{piv}|\omega_2')}{\Pr(\text{piv}|\omega_1)} \\
\geq\ & \left[1 + \min_{\omega_1\in\{\alpha_1,\beta_1\},\omega_2'\in\{\alpha_2,\beta_2\}} \frac{(q(\omega_1)-\frac{1}{2})^2 - (q(\omega_2')-\frac{1}{2})^2}{\frac{1}{4} - (q(\omega_1)-\frac{1}{2})^2}\right]^n \\
\geq\ & \left(1 + \left(\frac{(\frac{25}{100})^2 - (\frac{11}{100})^2}{\frac{1}{4} - (\frac{25}{100})^2}\right)\right) \\
\geq\ & \left(1 + \frac{25^2 - 121}{2500 - 25^2}\right)^n \\
\geq\ & (1.26)^n.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\Pr(\alpha|a,\text{piv})}{\Pr(\beta|a,\text{piv})} &\geq \frac{p_0}{1-p_0} \cdot \frac{\Pr(\alpha_2|\alpha)\Pr(a|\alpha_2)\Pr(\text{piv}|\alpha_2)}{\Pr(\beta_1|\beta)\Pr(a|\beta_1)\Pr(\text{piv}|\beta_1)} \\
&\geq \frac{2}{3}\frac{\frac{1}{n^3}}{(1-\frac{1}{n})}(1.26)^n \\
&\geq 5900 \quad \text{for} \quad n \geq 99.
\end{aligned}$$

This implies

$$\frac{\Pr(\alpha|d,a,\text{piv})}{\Pr(\beta|d,a,\text{piv})} \geq 100.$$

Since all voters receive message a in β_1, this implies that the likelihood that a pivotal voter votes A in β_1 exceeds $\Pr(\alpha|d, a, \text{piv}) \geq \frac{100}{101} > 0.99$. Since a is more likely to be received in α, the likelihood that a sincere voter votes A in β_1 exceeds

$$\begin{aligned} & \sum_{s\in\{u,d\}} \Pr(s|\beta) \frac{p_0 \Pr(s|\alpha)}{p_0 \Pr(s|\alpha) + (1-p_0)\Pr(s|\beta)} \\ = \; & 0.4 \cdot \frac{0.4 \cdot 0.6}{0.4 \cdot 0.6 + 0.6 \cdot 0.4} + 0.6 \cdot \frac{0.4 \cdot 0.4}{0.4 \cdot 0.4 + 0.6 \cdot 0.6} \\ > \; & 0.384. \end{aligned}$$

Thus, we obtain the lower bound $q(\beta_1; \kappa; \sigma^{\mathbf{q}^*}) > 0.4 \cdot 0.384 + 0.6 \cdot 0.999 = 0.753$. However, this contradicts the initial assumption that $q(\beta_1; \sigma^{\mathbf{q}^*}, \kappa) \leq 0.75$.

Step 5 *There is no fixed point* $\mathbf{q}^*$ *with* $q^*(\alpha_1) = 0.75$ *when* $n \geq 99$.

This follows from Step 4, since in both substates α_1 and β_1 all voters receive message a, and therefore $q^*(\alpha_1) = 0.75$ implies $q^*(\beta_1) = 0.75$. ■