Persuasion in Elections when Voting is Costly

Carl Heese   Stephan Lauermann

May 20, 2024

Abstract

We consider the problem of persuasion in the canonical election setting of Feddersen and Pesendorfer (1997) when voting is costly. We show that a sender can provide additional information in such a way that there is an equilibrium that inverts the full information outcome with a probability close to 1 when the voters' exogenously available information is sufficiently imprecise. A similar result holds when the share of partisans is sufficiently small. This note complements Heese and Lauermann (2023), who consider the persuasion problem when voting is costless.

1 Introduction

We consider persuasion of voters in an election in which participation is costly. The basic setup is a variation of Feddersen and Pesendorfer (1997) with a binary state, as in Bhattacharya (2013). As in their setting, voters have exogenous private information ("nature's signal") and heterogeneous preferences. They show that when participation is costless, all equilibria induce outcomes equivalent to those under full information. In Heese and Lauermann (2023), we show that a sender can manipulate the election. Specifically, in the same setting, she can induce any state-dependent outcome by releasing additional information.

Here, we consider the same setting but assume that voting is costly. In addition to their private information and preferences, voters now also draw some voting costs and then decide whether to turn out and, if so, how to vote.

Therefore, an interested party not only has to worry about how its message affects the voting choice but also whether voters turn out at all.

We study this question in a setting with a Poisson distributed voter number (Myerson, 1998). This allows us to use existing approximations of the probabilities to be pivotal. Critically, when there is a turnout margin, the relative probability of being pivotal depends not only on the relative margin of victory—the advantage of the expected winner in percentage of the underdog—but also on the total margin of victory.

Our main result shows that the message structure from Heese and Lauermann (2023) is still effective, provided the voter's private information is not too precise. Specifically, the sender can invert the full information outcome with probability above $1 - \delta$ for any $\delta > 0$ whenever the precision of the voters' private information is below some threshold that depends on δ. Therefore, even with costs, manipulation is still possible and can be very effective. Importantly, we show this for the message structure from Heese and Lauermann (2023), which is not tailored to a setting with costly voting. In this sense, we provide a robustness result. A message that is designed for costly voting may, of course, do even better.[1]

At the end of this note, we state similar results for the case in which the share of voters who are partisans (or almost partisans) is small. For every $\delta > 0$, when the share of voters with extreme preferences is small enough, then the sender can invert the full information outcome with a probability above $1-\delta$. The key idea is that many voters with intermediate types are abstaining, thereby delegating their choice to those who obtain precise information from the sender.

In the appendix, we provide a condensed outline of the proof of the main result that may be helpful to some readers.

[1]An open question left for future research is the full characterization of the set of state-dependent outcomes that are implementable by some appropriately designed message structure when voting is costly.

2 Model

The number of voters is Poisson distributed, with an expected number of n voters; so, the probability that there are k voters is

$$\Pr(k) = \frac{n^k e^{-n}}{k!}.$$

The voters hold a common prior about a binary state $\omega \in \{\alpha, \beta\}$, and the likelihood of α is denoted $\Pr(\alpha)$. A voter type is a pair $t = (y, c)$, where c is the type's cost of voting. We assume that y, c are drawn independently from each other and independently across voters. Specifically, c is drawn from the uniform distribution, $c \sim U[0,1]$, and $y \in [0,1]$ is drawn from a distribution with c.d.f. F, with a continuous density on $(0,1)$, and strictly positive mass points at 0 and 1 (partisans), with

$$m_A = F(0) \text{ and } m_B = 1 - \lim_{y \to 1^-} F(y).$$

The utility of a voter of type y from outcome A in ω is $u(A, \omega)$, given by

$$\begin{aligned} u(A,\alpha) &= (1-y), \text{ and} \\ u(A,\beta) &= -y, \end{aligned}$$

and utilities from outcome B are zero. With this specification, a voter prefers A when she believes α has probability above y (threshold of doubt). We can generalize this specification easily to any distribution for which payoffs are bounded, and there is a mass of "partisans."

We assume that the y distribution is such that the full information outcome is A in α and B in β, i.e., $m_A, m_B < \frac{1}{2}$.

Information: Sender's Message and Nature's Signal For each state ω, there are two "substates." Conditional on $\omega \in \{\alpha, \beta\}$, the likelihood of the substate ω_2 is ε and the likelihood of the substate ω_1 is $1 - \varepsilon$, with $0 < \varepsilon < \frac{1}{2}$. Conditional on the substate, the voters receive signals from two sources independently. First, each voter receives a private message $m \in \{a, b, z\}$, the signals are independent conditional on the substate and the signal probabilities are

$$\begin{aligned} \Pr(a|\alpha_1) &= 1 \text{ and } \Pr(a|\beta_2) = g, \\ \Pr(z|\alpha_2) &= \Pr(z|\beta_2) = (1-g), \end{aligned} \tag{1}$$

for some $0 < g < \frac{1}{2}$, and symmetrically in the other substates,

$$\Pr(b|\alpha_2) = \Pr(a|\beta_2), \tag{2}$$

$$\Pr(b|\beta_1) = \Pr(a|\alpha_1). \tag{3}$$

Second, each voter receives a private signal $s \in \{u, d\}$, the signals are independent conditional on the state and the signal probabilities satisfy

$$\infty > \frac{\Pr(u|\alpha)}{\Pr(u|\beta)} > \frac{\Pr(d|\alpha)}{\Pr(d|\beta)} > 0.$$

We assume also that signals are symmetric and use $r \in (\frac{1}{2}, \frac{2}{3})$ to measure the signal precision,

$$\Pr(u|\alpha) = \Pr(d|\beta) = r.$$

The assumption of symmetric signals can be dispensed with, and there can be more than two signal realizations. What matters is that signals are boundedly informative, and that there is a parameter that uniformly bounds their informativeness (e.g., we could have a continuous signal and assume that the maximal likelihood ratios are bounded by r.)

Note that we consider the sender's message used to "invert" the full information outcome. In a sense, this is the starkest manipulation. We have verified that the same arguments can also be used to show that there are messages for the sender that lead to equilibria in which an essentially constant outcome for both states can be implemented (always A or always B).

Strategies. A (symmetric) pure strategy of the voters is a mapping $\sigma : [0,1]^2 \times \{(s,m) : s \in \{u,d\}, m \in \{a,b,z\}\} \to \{A, B, \emptyset\}$ from types and signal pairs to actions.[2] Here $\sigma(y,c,s,m) = A$ means that the citizen votes A,

[2] One can show that it is without loss of generality only to consider pure strategies. The reason for this is that the type distribution is continuous, which implies that only a measure of zero of types is indifferent under the best response to any non-degenerate strategy.

$\sigma(y,c,s,m) = B$ means that the citizen votes B, and $\sigma(y,c,s,m) = \emptyset$ means that the citizen abstains.

Timing and Equilibrium. First, nature draws the number of voters, the state, and the substate. Second, the private signals are realized. Third, all voters choose their actions simultaneously. Finally, the outcome is given by the simple majority rule, where ties are broken uniformly. We study the symmetric and pure Bayes Nash equilibria of this game.

Notation and pivot probabilities. Take a single citizen and fix the other voters' strategy σ. The "voting rate" for x of type s, m is

$$\Pr(x|s,m;\sigma)$$

the probability that a random type chooses $x \in \{A, B, \emptyset\}$ when having received $(s,m) \in \{u,d\} \times \{a,b,z\}$, where the probability is with respect to the realized y, c. For example,

$$\Pr(A|s,m;\sigma) = \Pr\left(\sigma\left(y,c,s,m\right) = A\right).$$

If σ is a best response, then

$$\Pr(A|s,m;\sigma) = \Pr_{y,c}\left(\mathbb{E}U\left(A|s,m,y\right) \geq \max\left\{\mathbb{E}U\left(\emptyset|s,m,y\right), \mathbb{E}U\left(B|s,m,y\right)\right\}\right).$$

The probability of a vote for x in state ω_i is

$$\sum_{s,m} \Pr(x|s,m;\sigma)\Pr(s,m|\omega_i) \tag{4}$$

the likelihood that a voter chooses the action $x \in \{A, B, \emptyset\}$ in $\omega_i \in \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$. The number of A-votes in ω_i is Poisson distributed with mean

$$\lambda_A = n \sum_{s,m} \Pr(A|s,m;\sigma)\Pr(s,m|\omega_i)$$

The number of B-votes in ω_i is Poisson distributed with mean

$$\lambda_B = n \sum_{s,m} \Pr(B|s,m;\sigma)\Pr(s,m|\omega_i).$$

Denote by T the difference of the A-and B-votes of the other voters (for $T<0$, just swap the expressions). Then,

$$\begin{aligned}
\Pr\left(T=t|\,\omega_i;\sigma\right) &= \sum_{j=0}^{\infty} \mathrm{Poisson}\left(j+t,\lambda_A\right)\mathrm{Poisson}\left(j,\lambda_B\right) \\
&= \sum_{j=0}^{\infty} e^{-\lambda_A}\frac{\lambda_A^{j+t}}{(j+t)!}e^{-\lambda_B}\frac{\lambda_B^j}{j!} \\
&= e^{-(\lambda_A+\lambda_B)}\sum_{j=0}^{\infty}\frac{1}{j!\,(j+t)!}\left(\lambda_A^{j+t}\lambda_B^j\right) \\
&= e^{-(\lambda_A+\lambda_B)}\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{t}{2}}\sum_{j=0}^{\infty}\frac{1}{j!\,(j+t)!}\left(\frac{2\sqrt{\lambda_A\lambda_B}}{2}\right)^{j+t}\left(\frac{2\sqrt{\lambda_A\lambda_B}}{2}\right)^{j} \\
&= e^{-(\lambda_A+\lambda_B)}\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{t}{2}} I_t\left(2\sqrt{\lambda_A\lambda_B}\right), \qquad (5)
\end{aligned}$$

where $I_t(x)$ is the so-called modified Bessel function.

Suppose that a given citizen votes for A. The citizen's vote affects the outcome only in the event when $T=0$ or $T=-1$. Given the uniform tie-breaking, the likelihood that the A-vote shifts the outcome from B to A is[3]

$$\hat{\rho}(A,\omega_i;\sigma) := \frac{1}{2}(\Pr(T=0|\omega_i;\sigma)+\Pr(T=-1|\omega_i;\sigma)).$$

Suppose that the given citizen votes for B. The given citizen's vote affects the outcome only in the event when $T=0$ or $T=1$. Given the uniform tie-breaking, the likelihood that the B-vote shifts the outcome from A to B is

$$\hat{\rho}(B,\omega_i;\sigma) := \frac{1}{2}(\Pr(T=0|\omega_i)+\Pr(T=1|\omega_i)).$$

Let $\rho(x,\omega_i) = \hat{\rho}(x,\omega_i;\sigma)$ be the vector of the pivotal probabilities. The

[3] Sometimes we also write $\Pr(\mathrm{piv}_x|\omega_i)$ instead of $\hat{\rho}(A,\omega_i;\sigma)$; see, e.g., Lemma 2.

relevant expected utility (excluding c and constant terms) is

$$U(A|s,m,y) = \sum_{i\in\{1,2\}} \rho(A,\alpha_i)\frac{\Pr(\alpha_i)\Pr(s,m|\alpha_i)}{\Pr(s,m)}(1-y)-\rho(A,\beta_i)\frac{\Pr(\beta_i)\Pr(s,m|\beta_i)}{\Pr(s,m)}y,$$

and

$$U(B|s,m,y) = -\sum_{i\in\{1,2\}} \rho(B,\alpha_i)\frac{\Pr(\alpha_i)\Pr(s,m|\alpha_i)}{\Pr(s,m)}(1-y)+\rho(B,\beta_i)\frac{\Pr(\beta_i)\Pr(s,m|\beta_i)}{\Pr(s,m)}y.$$

Voting A is a best response if and only if

$$U(A|s,m,y) \geq \max\{c, U(B|s,m,y)\}.$$

3 Main Result and Preliminary Analysis

We show that the sender's message described before can be used to manipulate the election outcome to an arbitrary degree when nature's signal is sufficiently imprecise.

Theorem 1 *Given any preference distribution F and prior $\Pr(\alpha)$ that satisfies our assumptions: For every $\xi > 0$, there exists some $\bar{r} > \frac{1}{2}$ and some parameters (g, ε) for the sender's message such that for all signal precisions $r \in (\frac{1}{2}, \bar{r})$, there is an equilibrium σ when n is large enough for which*

$$\begin{aligned}\Pr(B \textit{ wins majority}|\alpha;\sigma,n) &\geq 1-\xi,\\ \Pr(A \textit{ wins majority}|\beta;\sigma,n) &\geq 1-\xi.\end{aligned}$$

3.1 (Truncated) best responses

The vector of pivotal probabilities $\rho(x,\omega_i)$ for $x \in \{A,B\}$ and $\omega_i \in \{\alpha_1,\beta_1,\alpha_2,\beta_2\}$ is sufficient for the voters' best response. We denote the induced best response given ρ with $\hat{\sigma}(\rho)$, that is, $\sigma = \hat{\sigma}(\rho)$ is the optimal voting behavior if voters believe that a vote for $x \in \{A,B\}$ is pivotal with probability $\rho(x,\omega_i)$ in the respective states.

With this notation, σ^* is an equilibrium if and only if $\sigma^* = \hat{\sigma}(\rho)$ for $\rho = \hat{\rho}(\sigma^*)$, where $\hat{\rho}$ was defined as the "correct" probabilities given the behavior σ.

For any n, the probability of being pivotal is bounded from below

$$\hat{\rho}(x, \omega_i; \sigma) \geq \frac{1}{2} e^{-n} \qquad \text{for } x \in \{A, B\},$$

which is the probability that there is only a single voter.

We now define a "box" for the pivotal probabilities: Given $K, M, W \geq 1$, let $\Pi(K, M, W)$ denote the set of vectors $\rho \in \left[\frac{1}{2K} e^{-n}, 1\right]^8$ such that

$$\begin{aligned}
\frac{\rho(A, \beta_1) + \rho(B, \beta_1)}{\rho(A, \alpha_2)} &\leq \frac{1}{K} \\
\frac{\rho(B, \alpha_1) + \rho(A, \alpha_1)}{\rho(B, \beta_2)} &\leq \frac{1}{K} \\
\frac{1}{M} &\leq \frac{\rho(B, \alpha_2)}{\rho(A, \alpha_2)} \leq M \\
\frac{1}{M} &\leq \frac{\rho(B, \beta_2)}{\rho(A, \beta_2)} \leq M \\
\frac{1}{W} &\leq \frac{\rho(B, \beta_2)}{\rho(A, \alpha_2)} \leq W
\end{aligned}$$

Denote by $\hat{\rho}_T(\sigma)$ denote the **truncated pivotal inference**, defined recursively:

$$\begin{aligned}
\hat{\rho}_T(A, \alpha_2) &= \hat{\rho}(A, \alpha_2) \\
\hat{\rho}_T(A, \beta_1) &= \min\{\hat{\rho}(A, \beta_1), \frac{1}{2K} \rho(A, \alpha_2)\} \\
\hat{\rho}_T(B, \beta_1) &= \min\{\hat{\rho}(B, \beta_1), \frac{1}{2K} \hat{\rho}(A, \alpha_2)\} \\
\hat{\rho}_T(B, \alpha_2) &= \max\{(\frac{1}{M} \hat{\rho}(A, \alpha_2), \min\{(M \hat{\rho}(A, \alpha_2), \hat{\rho}(B, \alpha_2)\} \\
\hat{\rho}_T(B, \beta_2) &= \max(\frac{1}{W} \hat{\rho}(A, \alpha_2), \min\{(W \hat{\rho}(A, \alpha_2), \hat{\rho}(B, \beta_2))\}) \\
\hat{\rho}_T(A, \beta_2) &= \max(\frac{1}{M} \hat{\rho}(B, \beta_2), \min\{(M \hat{\rho}(B, \beta_2), \hat{\rho}(A, \beta_2))\}) \\
\hat{\rho}_T(A, \alpha_1) &= \min\{\hat{\rho}(A, \alpha_1), \frac{1}{2K} \hat{\rho}(B, \beta_2)\} \\
\hat{\rho}_T(B, \alpha_1) &= \min\{\hat{\rho}(B, \alpha_1), \frac{1}{2K} \hat{\rho}(B, \beta_2)\}
\end{aligned}$$

where we dropped the σ argument. In particular,

$$\hat{\rho}_T(\sigma) \in \Pi(K, M, W) \text{ for all } \sigma.$$

Let $R(\rho) = \hat{\rho}_T(\hat{\sigma}(\rho))$, that is, given some ρ, we calculate the voters' best response $\hat{\sigma}(\rho)$ and then $R(\rho)$ is the truncated inference from σ defined above. We call it the **truncated best response**.

Since R is continuous on a compact and convex set, it has a fixed point,

$$\rho^* = R(\rho^*).$$

By definition, if ρ^* is in the interior of $\Pi(K, M, W)$, then ρ^* corresponds to an equilibrium $\sigma^* = \hat{\sigma}(\rho^*)$.

3.2 Normalization and Asymptotics of Best Responses

Because of the Poisson distribution, $\hat{\rho}_T(\sigma) > 0$ (there is always a possibility that there will be no other voter). For the following, we fix some sequence of fixed points ρ of the truncated best response as $n \to \infty$ (meaning, in particular, for fixed K, M, W) and the corresponding sequence of strategies $\sigma = \hat{\sigma}(\rho)$.

To study the outcome of large elections, it will be useful to normalize the participation rates. Specifically, let

$$q(x, s, m) = \frac{\Pr(x|s, m; \sigma)}{\rho(A, \alpha_2)}$$

be the normalized participation rate, and its limit is

$$\bar{q}(x, s, m) = \lim_{n\to\infty} q(x, s, m) = \lim_{n\to\infty} \frac{\Pr(x|s, m; \sigma)}{\rho(A, \alpha_2)}$$

and

$$p(x, \omega_i) = \sum_{s,m} q(x, s, m) \Pr(s, m|\omega_i)$$

with limit

$$\bar{p}(x, \omega_i) = \lim_{n\to\infty} p(x, \omega_i).$$

Without loss of generality, we will consider sequences for which these limits exist (in the extended reals). Of course, the limits are bounded when the sequences are from the box,. Note that

$$\begin{aligned} \lambda_x(\alpha_2) &= n\rho(A,\alpha_2)\,p(x,\omega_i) && (6)\\ &= n\rho(A,\alpha_2)\sum_{s,m} q(x,s,m)\Pr(s,m|\alpha_2)\,. && (7)\end{aligned}$$

4 Proof of the Main Result

In the following, we consider a sequence of vectors of pivotal probabilities for $n \to \infty$ that are fixed points of the truncated best response. In particular, all pivotal probabilities are in the box, and so we consider behavior that is a truncated best response to pivotal probabilities from the box. (The only exceptions are the first two lemmas, Lemma 2 and Lemma 1 where we allow strategies that are best responses to any belief from the box which is not necessarily a fixed point.)[4]

For this and the subsequent analysis until the very end, we fix the parameter $\varepsilon > 0$.

4.1 Bounds on the Participation Rates

Lemma 1 *Fix any K, M, W, g, r. For any sequence of strategies that are best responses to vectors from the box $\Pi(K, M, W)$, for any (s, m),*

$$\begin{aligned} W &\geq \bar{q}(A,s,m) \geq m_A \Pr(\alpha_2|s,m),\\ W &\geq \bar{q}(B,s,m) \geq m_B \frac{1}{W}\Pr(\beta_2|s,m).\end{aligned}$$

To interpret the lemma, note that $\Pr(\omega_2|s,m) > 0$ for all ω_2 and (s,m) except for

$$\Pr(\alpha_2|s,a) = \Pr(\beta_2|s,b) = 0 \text{ for } s \in \{u,d\}\,.$$

Proof. Recall that $m_A > 0, m_B > 0$ are the mass of partisans at $y = 0$ and $y = 1$, respectively.

Lower Bounds: The utility of $y = 0$ (an A-partisan) with $(s,m) \in$

[4]Formally, so far, we have only introduced the participation rates $q(x,s,m)$ and then the voting rate $\lambda_x(\alpha_2)$, etc., for fixed points of the truncated best response. However, the corresponding definitions for the best response σ to any belief in the box are analogous.

$\{(b,u),(b,d),(z,u),(z,d)\}$ from voting A is

$$\begin{aligned} U(A,s,m,0) &= \rho(A,\alpha_2)\Pr(\alpha_2|s,m)+\rho(A,\alpha_1)\Pr(\alpha_1|s,m) \\ &\geq \rho(A,\alpha_2)\Pr(\alpha_2|s,m). \end{aligned}$$

The implied likelihood of voting A under the best response is therefore bounded,

$$q(A,s,m) \geq m_A \Pr(U(A,s,m,0) \geq c) = m_A U(A,s,m,0)$$

where the last equality is from the uniform distribution of c. Thus, taking limits

$$\bar{q}(A,m,s) \geq m_A \Pr(\alpha_2|m,s). \tag{8}$$

Similarly,

$$\bar{q}(B,m,s) \geq m_B \frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)} \Pr(\beta_2|m,s). \tag{9}$$

Since the vector ρ is in $\Pi(K,M,W)$, we have $\frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)} \geq \frac{1}{W}$. The two displayed equations imply the lower bounds.

Upper Bounds: For any signal pair s,m, from $y \leq 1$ we get $U(A,s,m,y) \leq \sum_{i=1,2} \rho(A,\alpha_i)\Pr(\alpha_i|s,m)$. For any $K \geq 1$, $\rho(A,\alpha_1) \leq \rho(A,\alpha_2)\frac{1}{K}W \leq \rho(A,\alpha_2)W$. Thus, $U(A,s,m,y) \leq \rho(A,\alpha_2)W$ for all s,m and y. Integrating with respect to y and using that c is uniform gives $q(A,s,m) \leq W$, and taking limits implies

$$\bar{q}(A,s,m) \leq W.$$

By the analogous argument, the likelihood of a B-vote after any s,m is bounded from above by $\rho(B,\beta_2)$ (since $\rho(B,\beta_1) \leq \frac{1}{K}\rho(A,\alpha_2)$). Further, $\rho(B,\beta_2) \leq W\rho(A,\alpha_2)$ since the fixed point is in $\Pi(K,M,W)$. Together and since $K \geq 1$, this implies

$$\bar{q}(B,s,m) \leq W.$$

The displayed equations imply the upper bounds. ■

4.2 An Approximation Lemma on Pivotal Updating

For the following lemma, we consider a sequence of strategies σ_n that are best responses to some sequence of vectors from the box and for which $\lambda_A(\alpha_2) \to$

∞, the normalized voting rates converge for some (sub-)sequence, and, given $\lambda_x(\alpha_2) = n\rho(x,\alpha_2)\, p(x,\omega_i)$ for $x \in \{A,B\}$, we have

$$\lim \frac{\lambda_A(\omega_i)}{\lambda_B(\omega_i)} = \lim \frac{n\rho(A,\omega_i)\, p(A,\omega_i)}{n\rho(A,\omega_i)\, p(B,\omega_i)} = \frac{\bar{p}(A,\omega_i)}{\bar{p}(B,\omega_i)} \in (0,\infty)\,. \tag{10}$$

Lemma 2 *Take any sequence of strategies σ that are best responses to a sequence of vectors ρ from the "box". Suppose that $\lambda_A(\alpha_2) \to \infty$, that the vector of normalized voting rates $\bar{p}(\cdot,\cdot)$ converges (to finite, possibly zero numbers), and that (10) holds for $\omega_i = \alpha_2$. If $\bar{p}(A,\omega_i') > 0$ and $\bar{p}(B,\omega_i') > 0$ for some substate ω_i, then*

$$\lim_{n\to\infty} \frac{\Pr(PivB|\omega_i')}{\Pr(PivA|\omega_i')} = \frac{1+\sqrt{\frac{\bar{p}(A,\omega_i)}{\bar{p}(B,\omega_i)}}}{1+\sqrt{\frac{\bar{p}(B,\omega_i)}{\bar{p}(A,\omega_i)}}}. \tag{11}$$

Moreover, if for any two substates ω and $\hat{\omega}$ from $\{\alpha_1,\beta_1,\alpha_2,\beta_2\}$,

$$\left(\sqrt{\bar{p}(A,\omega)} - \sqrt{\bar{p}(B,\omega)}\right)^2 > \left(\sqrt{\bar{p}(B,\hat{\omega})} - \sqrt{\bar{p}(A,\hat{\omega})}\right)^2, \tag{12}$$

and $\bar{p}(A,\hat{\omega}) > 0$ and $\bar{p}(B,\hat{\omega}) > 0$, then

$$\lim_{n\to\infty} \frac{\Pr(PivA|\omega) + \Pr(PivB|\omega)}{\Pr(PivA|\hat{\omega})} = \lim_{n\to\infty} \frac{\Pr(PivA|\omega) + \Pr(PivB|\omega)}{\Pr(PivB|\hat{\omega})} = 0. \tag{13}$$

Proof. Preliminaries. We prepare the proof with some preliminary observations. First, we observe some implications of the various assumptions in the lemma: The assumption $\bar{p}(x,\omega_i') > 0$ together with (10) for $\omega_i = \alpha_2$ and $\lambda_A(\alpha_2) \to \infty$ implies that $\lambda_x(\omega_i') \to \infty$. This is because (10) for $\omega_i = \alpha_2$ implies $\bar{p}(A,\alpha_2) > 0$ so that $\lim \frac{\lambda_A(\omega_i)}{\lambda_x(\omega_i)} = \frac{\bar{p}(A,\alpha_2)}{\bar{p}(x,\omega_i)} \in (0,\infty)$.

Note that (10) for $\omega_i = \alpha_2$ also implies $\bar{p}(B,\alpha_2) > 0$ and given $\lambda_A(\alpha_2) \to \infty$ it also implies then $\lambda_B(\alpha_2) \to \infty$.

Second, we restate the previous expressions for the pivotal probabilities, using (5),

$$\Pr(PivB) = \frac{1}{2} e^{-(\lambda_A+\lambda_B)} \left(I_0\left(2\sqrt{\lambda_A\lambda_B}\right) + \left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}} I_1\left(2\sqrt{\lambda_A\lambda_B}\right) \right), \tag{14}$$

and

$$\Pr\left(PivA\right) = \frac{1}{2}e^{-(\lambda_A+\lambda_B)}\left(I_0\left(2\sqrt{\lambda_A\lambda_B}\right) + \left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}} I_1\left(2\sqrt{\lambda_A\lambda_B}\right)\right). \quad (15)$$

The functions I_0 and I_1 are commonly called modified Bessel functions of order 0 and 1. They can be approximated for large numbers x as

$$I_0(x) \approx I_1(x) \approx \frac{e^x}{\sqrt{2\pi x}}, \quad (16)$$

see Abramowitz and Stegun (1968) p.377), where $f(x) \approx g(x)$ means that $\lim_{x\to\infty}\frac{f(x)}{g(x)} = 1$. Given the preliminary observation above at the very start of the proof, $\bar{p}(x,\omega_i') > 0$ for $x \in \{A, B\}$ imply $\lambda_A(\omega_i')\lambda_B(\omega_i') \to \infty$. This way,

$$\begin{aligned}\frac{\Pr\left(PivB|\omega_i'\right)}{\Pr\left(PivA|\omega_i'\right)} &= \frac{I_0\left(2\sqrt{\lambda_A\lambda_B}\right)}{I_0\left(2\sqrt{\lambda_A\lambda_B}\right)}\frac{1+\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}}\frac{I_1\left(2\sqrt{\lambda_A\lambda_B}\right)}{I_0\left(2\sqrt{\lambda_A\lambda_B}\right)}}{1+\left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}}\frac{I_1\left(2\sqrt{\lambda_A\lambda_B}\right)}{I_0\left(2\sqrt{\lambda_A\lambda_B}\right)}} \\ &\approx \frac{1}{1}\frac{1+\sqrt{\frac{\bar{p}_A}{\bar{p}_B}}}{1+\sqrt{\frac{\bar{p}_B}{\bar{p}_A}}}\end{aligned}$$

where the first line simply rewrites (14) and (15), and the approximation uses (16) and (10). Thus, (11) holds. Note that we used the short-hand notation $\bar{p}_x$ for $\bar{p}(x,\omega_i')$ and λ_x for $\lambda(x,\omega_i')$, and we will continue doing so in this proof.

The approximation (16) implies that the pivotal probability satisfies

$$\Pr\left(PivA\right) \approx \frac{1}{2}e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^2}\frac{1+\sqrt{\frac{\lambda_B}{\lambda_A}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}}, \quad (17)$$

provided that $\lim_{n\to\infty}\lambda_A\lambda_B=\infty$. This follows from

$$\begin{aligned}
\Pr(PivA) &= \frac{1}{2}e^{-(\lambda_A+\lambda_B)}\left(I_0\left(2\sqrt{\lambda_A\lambda_B}\right)+\left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}}I_1\left(2\sqrt{\lambda_A\lambda_B}\right)\right)\\
&\approx \frac{1}{2}e^{-(\lambda_A+\lambda_B)}I_0\left(2\sqrt{\lambda_A\lambda_B}\right)\left(1+\left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}}\right)\\
&\approx \frac{1}{2}e^{-(\lambda_A+\lambda_B)}\frac{e^{2\sqrt{\lambda_A\lambda_B}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}}\left(1+\left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}}\right)\\
&= \frac{1}{2}e^{-\left(\sqrt{\lambda_A}+\sqrt{\lambda_B}\right)^2}\frac{1+\left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}}.
\end{aligned}$$

Similarly,

$$\Pr(PivB)\approx\frac{1}{2}e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^2}\frac{1+\sqrt{\frac{\lambda_A}{\lambda_B}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}}. \tag{18}$$

Now, let's turn to the proof of the second part of the lemma. Abbreviate

$$\begin{aligned}
\lambda_x &= \lambda_x(\omega),\\
\tau_x &= \lambda_x(\hat{\omega}).
\end{aligned}$$

Note that (12) implies that

$$\begin{aligned}
& (\sqrt{\lambda_A}-\sqrt{\lambda_B})^2-(\sqrt{\tau_A}-\sqrt{\tau_B})^2 \qquad (19)\\
= \;& n\rho(A,\alpha_2)\left(\sqrt{p(A,\omega)}-\sqrt{p(B,\omega)}\right)^2-n\rho(A,\alpha_2)\left(\sqrt{p(B,\hat{\omega})}-\sqrt{p(A,\hat{\omega})}\right)^2\\
\approx \;& n\rho(A,\alpha_2)\left[\left(\sqrt{\bar{p}(A,\omega)}-\sqrt{\bar{p}(B,\omega)}\right)^2-\left(\sqrt{\bar{p}(B,\hat{\omega})}-\sqrt{\bar{p}(A,\hat{\omega})}\right)^2\right]\\
\to \;& \infty
\end{aligned}$$

Case 1: Suppose $\bar{p}(B,\omega)>0$ **and** $\bar{p}(A,\omega)>0$. Recall that $\bar{p}(x,\hat{\omega})$ for $x=A,B$ are strictly positive by assumption. The previous approximation

(17) implies

$$\frac{\Pr(PivA|\omega)}{\Pr(PivA|\hat{\omega})} \approx \frac{\frac{1}{2}e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^2}}{\frac{1}{2}e^{-(\sqrt{\tau_A}-\sqrt{\tau_B})^2}} \frac{1+\sqrt{\frac{\lambda_B}{\lambda_A}}}{1+\sqrt{\frac{\tau_B}{\tau_A}}} \frac{\sqrt{4\pi\sqrt{\tau_A\tau_B}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}}$$

$$\approx \frac{\frac{1}{2}e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^2}}{\frac{1}{2}e^{-(\sqrt{\tau_A}-\sqrt{\tau_B})^2}} \frac{1+\sqrt{\frac{\bar{p}(B,\omega)}{\bar{p}(A,\omega)}}}{1+\sqrt{\frac{\bar{p}(B,\hat{\omega})}{\bar{p}(A,\hat{\omega})}}} \frac{\sqrt{4\pi\sqrt{\bar{p}(A,\hat{\omega})\,\bar{p}(B,\hat{\omega})}}}{\sqrt{4\pi\sqrt{\bar{p}(A,\omega)\,\bar{p}(B,\omega)}}}$$

and now the claim (13) follows since the last two terms are bounded and bounded away from 0, and the first term vanishes: If (12) holds, then (19) implies (note the sign-change)

$$\frac{\frac{1}{2}e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^2}}{\frac{1}{2}e^{-(\sqrt{\tau_A}-\sqrt{\tau_B})^2}} \to 0.$$

The same argument implies that $\frac{\Pr(PivB|\omega)}{\Pr(PivA|\hat{\omega})} \to 0$, and so (13) holds.

Case 2. Suppose that $\bar{p}(B,\omega)=0$. Since the lemma only considers strategies that are best responses to a belief from the box, it holds $\bar{p}(B,\omega)+\bar{p}(A,\omega)>0$, given Lemma 1, and so $\bar{p}(A,\omega)>0$. [The case of $\bar{p}(A,\omega)=0<\bar{p}(B,\omega)$ is analogous and omitted.] Note that $\bar{p}(A,\omega)>0$ implies $\lambda_A\to\infty$ given the first preliminary observation.

We first show the following auxiliary claim: For every $\mu>0$, the hypothesis of $\bar{p}(A,\omega)>\bar{p}(B,\omega)=0$ implies that

$$\lim_{n\to\infty}\frac{\Pr(PivX|\omega)}{e^{-\lambda_A(1-\mu)}}=0 \text{ for } x\in\{A,B\}. \tag{20}$$

(For an interpretation, recall that $e^{-\lambda_A}$ is the probability there is no A vote. Since there are almost no B votes in expectation, the most likely event in which a vote is pivotal is the one with almost no A vote; see also below for a more precise statement.)

First, consider **Case 2a**, $\lim_{n\to\infty}\lambda_A\lambda_B=\infty$. Then, for large enough n, $\lambda_B\geq\frac{1}{\lambda_A}$ and so

$$\frac{\lambda_A}{\lambda_B}\leq(\lambda_A)^2. \tag{21}$$

Moreover, given $\lim_{n\to\infty} \lambda_A\lambda_B = \infty$ we can use the approximation (18),

$$\begin{aligned}
\Pr(PivB|\omega) &\approx \frac{1}{2}e^{-(\sqrt{\lambda_A}-\sqrt{\lambda_B})^2}\frac{1+\sqrt{\frac{\lambda_A}{\lambda_B}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}} \\
&= \frac{1}{2}e^{-\lambda_A(1-\sqrt{\frac{\lambda_B}{\lambda_A}})^2+\mu\lambda_A}e^{-\mu\lambda_A}\frac{1+\sqrt{\frac{\lambda_A}{\lambda_B}}}{\sqrt{4\pi\sqrt{\lambda_A\lambda_B}}} \\
&\leq \frac{1}{2}e^{-\lambda_A(1-\sqrt{\frac{\lambda_B}{\lambda_A}})^2+\mu\lambda_A}e^{-\mu\lambda_A}(1+\lambda_A) \qquad (22)
\end{aligned}$$

where the equality is from rewriting and, for large n, the inequality from omitting the denominator of the last term and $\sqrt{\frac{\lambda_A}{\lambda_B}} \leq \lambda_A$ from (21). Now, the claim follows for $x = B$ since

$$\lim(1-\sqrt{\frac{\lambda_B}{\lambda_A}})^2 = (1-\sqrt{\frac{\bar{p}_B}{\bar{p}_A}})^2 = 1 \qquad (23)$$

and

$$e^{-\mu\lambda_A}(1+\lambda_A) \to 0. \qquad (24)$$

Using the approximation (17) and the similar rewriting as above, and that

$$\frac{1+\sqrt{\frac{\lambda_B}{\lambda_A}}}{\sqrt{\lambda_A\lambda_B}} \to 0$$

since $\frac{\lambda_B}{\lambda_A} \to 0$ and $\lambda_A\lambda_B \to \infty$, (23) and (24) imply the claim for $x = A$.

Second, consider **Case 2b**, $\lim_{n\to\infty} \lambda_A\lambda_B = k < \infty$. For this case, note that the modified Bessel functions I_0 and I_1 are continuous and strictly positive on $(0,\infty)$, with $I_0(0)=1$ and $I_1(0)=0$. Moreover, $\lim_{x\to 0}\frac{1}{x}I_1(x)=\frac{1}{2}$. Now,

$$\begin{aligned}
\Pr(PivB|\omega) &= \frac{1}{2}e^{-(\lambda_A+\lambda_B)}\left(I_0\left(2\sqrt{\lambda_A\lambda_B}\right)+\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}}I_1\left(2\sqrt{\lambda_A\lambda_B}\right)\right)\\
&= \frac{1}{2}e^{-(\lambda_A+\lambda_B)}\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}}I_1\left(2\sqrt{\lambda_A\lambda_B}\right)\left(\frac{1}{\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}}}\frac{I_0\left(2\sqrt{\lambda_A\lambda_B}\right)}{I_1\left(2\sqrt{\lambda_A\lambda_B}\right)}+1\right)\\
&\approx \frac{1}{2}e^{-\lambda_A(1-\mu)}e^{-\lambda_A\mu}\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}}I_1\left(2\sqrt{\lambda_A\lambda_B}\right)\\
&= \frac{1}{2}e^{-\lambda_A(1-\mu)}e^{-\lambda_A\mu}\frac{2\lambda_A}{2\sqrt{\lambda_B\lambda_A}}I_1\left(2\sqrt{\lambda_A\lambda_B}\right). \qquad (25)
\end{aligned}$$

Where the $\approx$ is from $\frac{\lambda_A}{\lambda_B}\to\infty$, $\lambda_A\lambda_B\to k<\infty$, and continuity of I_0 and I_1 (and their ratio); and the $=$ is simply rewriting $\left(\frac{\lambda_A}{\lambda_B}\right)^{\frac{1}{2}}$. Hence, for $\lim_{n\to\infty}\lambda_A\lambda_B=k\in(0,1)$, continuity of I_1 implies

$$\Pr(PivB|\omega)\approx\frac{1}{2}e^{-\lambda_A(1-\mu)}e^{-\lambda_A\mu}\frac{\lambda_A}{\sqrt{k}}I_1\left(2\sqrt{k}\right)$$

and so the claim follows from $e^{-\lambda_A\mu}\lambda_A\to 0$.

For $k=0$, $\lim_{x\to 0}\frac{1}{x}I_1(x)=\frac{1}{2}$ and (25) imply that

$$\Pr(PivB|\omega)\approx\frac{1}{2}e^{-\lambda_A(1-\mu)}e^{-\lambda_A\mu}2\lambda_A\frac{1}{2}.$$

Therefore, the claim follows for $x=B$. Also, note the intuitive result that for $k=0$, the previously displayed equation simplifies to

$$\Pr(PivB|\omega)\approx\frac{1}{2}e^{-\lambda_A}\lambda_A=\Pr(1\text{ Vote for A}),$$

that is, a B vote is pivotal with a probability equal to roughly one A being present. (The probability of 1 vote for A is much higher than 0 votes for A given $\lambda_A\to\infty$ and since with increasing probability there are 0 other votes for B given $\lambda_B\to 0$, the probabilities of all pivotal events different from a single vote for A are of much lower order.)

For $x = A$ we have

$$\begin{aligned}\Pr\left(PivA|\omega\right) &= \frac{1}{2}e^{-(\lambda_A+\lambda_B)}\left(I_0\left(2\sqrt{\lambda_A\lambda_B}\right)+\left(\frac{\lambda_B}{\lambda_A}\right)^{\frac{1}{2}}I_1\left(2\sqrt{\lambda_A\lambda_B}\right)\right)\\ &\approx \frac{1}{2}e^{-\lambda_A}I_0\left(k\right),\end{aligned} \tag{26}$$

where the $\approx$ follows from $\frac{\lambda_B}{\lambda_A}\to 0$, $\lambda_A\lambda_B\to k<\infty$, and the continuity of I_0 (including at $k=0$). The claim (20) follows immediately from (26) for all $\mu>0$. (Again, the approximation in (26) is intuitive: The modal event for an A vote to be pivotal is that there are no other voters, which happens with probability $e^{-\lambda_A-\lambda_B}\approx e^{-\lambda_A}$.)

Using the claim (20), we can prove (13) from the Lemma for Case 2: Multiplying (20) by $e^{\lambda_A\mu}e^{-\lambda_A\mu}$ and using (17) for $\hat{\omega}$, which we can do given $\tau_A\tau_B\to\infty$ by $\bar{p}_B\left(\hat{\omega}\right)>0$ and $\bar{p}_A\left(\hat{\omega}\right)>0$):

$$\lim\frac{\Pr\left(PivX|\omega\right)}{\Pr\left(PivA|\hat{\omega}\right)}\leq\lim\frac{e^{-\lambda_A(1-2\mu)}}{\frac{1}{2}e^{-(\sqrt{\tau_A}-\sqrt{\tau_B})^2}}e^{-\lambda_A\mu}\frac{\sqrt{4\pi\sqrt{\tau_A\tau_B}}}{1+\sqrt{\frac{\tau_B}{\tau_A}}}.$$

Now, the claim (13) follows for small enough μ since the terms of the right side vanish: The assumption $0=\bar{p}_B\left(\omega\right)<\bar{p}_A\left(\omega\right)$ implies $\frac{\lambda_A}{\lambda_B}\to\infty$ so that the hypothesis (12) of the lemma implies that $\lambda_A>(\sqrt{\tau_A}-\sqrt{\tau_B})^2$ and so, for sufficiently small $1>\mu>0$,

$$\frac{e^{-\lambda_A(1-2\mu)}}{\frac{1}{2}e^{-(\sqrt{\tau_A}-\sqrt{\tau_B})^2}}\to 0,$$

as in (19). For the remaining terms, note that $\frac{\tau_B}{\tau_A}\to\frac{\bar{p}_B(\hat{\omega})}{\bar{p}_A(\hat{\omega})}$ and given the assumptions of the lemma, namely $\bar{p}_A\left(\hat{\omega}\right)>0$ and $\bar{p}_B\left(\hat{\omega}\right)>0$, one has $\lim_{n\to\infty}\frac{\tau_B}{\tau_A}\in(0,1)$. Then, note that, there is a constant $C>0$ so that for large n, $e^{-\lambda_A\mu}\sqrt{\sqrt{\tau_A\tau_B}}\leq e^{-\lambda_A\mu}\lambda_A\cdot C$. To see why, note that the assumptions $\bar{p}_A\left(\hat{\omega}\right)>0$ and $\bar{p}_B\left(\hat{\omega}\right)>0$ and the implication $\bar{p}_A\left(\omega\right)>0$ of the case assumption $\bar{p}_B\left(\hat{\omega}\right)>0$ (see the discussion at the start of the case) together imply $\lim\frac{\tau_A\tau_B}{\lambda_A\lambda_A}\in(0,1)$. This in turn implies $\sqrt{\sqrt{\tau_A\tau_B}}=\sqrt{\lambda_A\sqrt{\frac{\tau_A\tau_B}{\lambda_A\lambda_A}}}<\lambda_A\cdot C$ for some constant $C>0$ and any n large enough. Now, $e^{-\lambda_A\mu}\lambda_A\to 0$ since $\lambda_A\to\infty$. Recall for this that $\lambda_A\to\infty$ which follows from the first preliminary observation of the proof and the implication $\bar{p}(A,\omega) > 0$ of the case assumption. Finally, we conclude that the term on the right vanishes, too. ■

4.3 Turnout Diverges to Infinity

Lemma 3 *For $n \to \infty$, and any sequence of fixed points of the truncated best responses,*

$$\lim_{n\to\infty} n\rho\left(A,\alpha_2\right) = \infty.$$

Proof. By contradiction. Suppose $\lim_{n\to\infty} n\rho\left(A,\alpha_2\right) = k < \infty$. Then, $\bar{q}\left(x,s,m\right)$ being bounded from Lemma 1 implies that

$$\lambda_x\left(\alpha_2\right) = n\rho\left(A,\alpha_2\right)\sum_{s,m} q\left(x,s,m\right)\Pr\left(s,m|\alpha_2\right)$$

is bounded for large n. However, since

$$\hat{\rho}_T\left(A,\alpha_2\right) \geq \frac{1}{2}e^{-\lambda_A(\alpha_2)}e^{-\lambda_B(\alpha_2)}$$

(where the right side equals the probability that there is no A and no B vote, times $\frac{1}{2}$), $\hat{\rho}_T\left(A,\alpha_2\right)$ is then bounded from below as well and so $\hat{\rho}_T\left(A,\alpha_2\right) = \rho\left(A,\alpha_2\right)$ (since we are taking sequences of fixed points) implies that $n\rho\left(A,\alpha_2\right) \to \infty$. ■

4.4 The M-Bound Does not Bind

In this lemma and throughout this proof, when we say "for all K,g,r" we mean all that are admissible (that is $K \geq 1$, $0 < g < \frac{1}{2}$, and $\frac{1}{2} < r < \frac{2}{3}$). Recall that all fixed points of the truncated best response are in the "box," $\Pi\left(K,M,W\right)$.

Lemma 4 *For every W, there exists some $\bar{M}\left(W\right)$ such that for all K,g,r, if $M \geq \bar{M}\left(W\right)$ then for any sequence of fixed points of the truncated best responses,*

$$\frac{1}{\bar{M}} < \lim_{n\to\infty}\frac{\rho\left(B,\alpha_2\right)}{\rho\left(A,\alpha_2\right)} < \bar{M}.$$

Proof. Note that for all admissible K,g,r, the probabilities $\Pr(\alpha_2|s,z)$ and $\Pr(\beta_2|s,z)$ are bounded from below. Thus, the lower bounds on $\bar{q}\left(x,s,z\right)$ in

Lemma 1 is uniform for all r (and trivially so for K, g). Hence, $\bar{p}(x, \alpha_2)$ are uniformly bounded from below. Together with the upper bounds, we have that

$$\frac{\bar{p}(B, \alpha_2)}{\bar{p}(A, \alpha_2)}$$

are uniformly bounded from below and from above, for all K, g, r (given a fixed W).

Let us verify the conditions for applying the approximation (11) of Lemma 2. Note that Lemma 1 together with Lemma 3 implies $\lambda_A(\alpha_2) \to \infty$, whereas $\frac{\bar{p}(B,\alpha_2)}{\bar{p}(A,\alpha_2)}$ being bounded from above and below implies (10) for $\omega_i = \alpha_2$ and $\lambda_B(\alpha_2) \to \infty$. So, the approximation (11) applies and we obtain ,

$$\frac{\Pr(PivB|\alpha_2)}{\Pr(PivA|\alpha_2)} \approx \frac{1+\sqrt{\frac{\bar{p}_A}{\bar{p}_B}}}{1+\sqrt{\frac{\bar{p}_B}{\bar{p}_A}}}.$$

Therefore, the uniform upper and lower bounds on $\frac{\bar{p}_A}{\bar{p}_B}$ imply uniform upper and lower bounds on $\frac{\Pr(PivB|\alpha_2)}{\Pr(PivA|\alpha_2)}$ for all K, g, r, as required. ■

Note that the discussion in the last proof establishes that the conditions (10) for $\omega_i = \alpha_2$ and $\lambda_A(\alpha_2) \to \infty$ of Lemma 2 are satisfied for any sequence of fixed points of the truncated best response. So, we can in the following generally apply Lemma 2 simply by (implicitly moving over to a sub-sequence for which $\bar{p}(-,-)$ converge) and verifying the additional conditions, e.g., $\bar{p}(A, \omega_i')$ and $\bar{p}(B, \omega_i') > 0$.

4.5 The W-Bound Does not Bind

For the following lemma, recall that we assume $r > \frac{1}{2}$. This lemma is where we use this assumption.

Lemma 5 *Fix any K, M, W, g, r. For any sequence of fixed points of truncated best responses, for $m \in \{a, b, z\}$,*

$$\bar{q}(A, u, m) \geq \bar{q}(A, d, m) \text{ and } \bar{q}(B, d, m) \geq \bar{q}(B, u, m),$$

with strict inequalities for $m = z$.

For any g, W, M, there exists some $\bar{K}(g, W, M)$ such that for all $K \geq$

$\bar{K}\,(g, W, M)$ *and all* r*,*

$$\bar{q}(B, u, a) > \bar{q}(A, u, a) \ \textit{and}\ \bar{q}(A, d, b) > \bar{q}(B, d, b),$$

as well as

$$\bar{q}(B, u, a) > \bar{q}(B, d, b) \ \textit{and}\ \bar{q}(A, d, b) > \bar{q}(A, u, a).$$

and $\bar{K}(g, W, M) \to \infty$ *as* $W \to \infty$ *or* $M \to \infty$*.*

Proof. The first ordering follows from the inspection of

$$U(A|s, m, y) = \sum_{i \in \{1,2\}} \rho(A, \alpha_i) \frac{\Pr(\alpha_i) \Pr(s, m|\alpha_i)}{\Pr(s, m)} (1-y) - \rho(A, \beta_i) \frac{\Pr(\beta_i) \Pr(s, m|\beta_i)}{\Pr(s, m)} y,$$

and

$$U(B|s, m, y) = -\sum_{i \in \{1,2\}} \rho(B, \alpha_i) \frac{\Pr(\alpha_i) \Pr(s, m|\alpha_i)}{\Pr(s, m)} (1-y) + \rho(B, \beta_i) \frac{\Pr(\beta_i) \Pr(s, m|\beta_i)}{\Pr(s, m)} y.$$

Specifically, $U(A|s, m, y)$ is higher for $s = u$ than for $s = d$, since $\frac{\Pr(\alpha_i)\Pr(s,m|\alpha_i)}{\Pr(s,m)}$ is higher and $\frac{\Pr(\beta_i)\Pr(s,m|\beta_i)}{\Pr(s,m)}$ is lower; conversely for $U(B|s, m, y)$. The strict inequalities of the limits for $m = z$ follow from the fact that $\bar{q}(x, s, z)$ are bounded above 0: There are types y that turn out with positive probability after d, and those will turn out with an even higher probability after u

For the second claim, note that for $m = a$

$$U(A|s, a, y) = \rho(A, \alpha_1) \frac{\Pr(\alpha_1) \Pr(s, m|\alpha_1)}{\Pr(s, m)} (1-y) - \rho(A, \beta_2) \frac{\Pr(\beta_2) \Pr(s, m|\beta_2)}{\Pr(s, m)} y,$$

and

$$U(B|s, a, y) = -\rho(B, \alpha_1) \frac{\Pr(\alpha_1) \Pr(s, m|\alpha_1)}{\Pr(s, m)} (1-y) + \rho(B, \beta_2) \frac{\Pr(\beta_2) \Pr(s, m|\beta_2)}{\Pr(s, m)} y.$$

Since $\rho(A, \alpha_1) \leq \frac{1}{K} \rho(B, \beta_2) \leq \frac{W}{K} \rho\,(A, \alpha_2)$, it must be that $\bar{q}\,(A, s, a)$ vanishes when K grows large enough relative to M. Since $\bar{q}(B, s, a)$ is uniformly bounded from below by Lemma 1 for any given g, the claim follows when K is above some $\bar{K}\,(g, M)$. An analogous argument implies the inequality for $m = b$.

The third pair of inequalities follows from the same type of argument.

First, note that $\bar{q}(B,d,b) \leq \frac{\rho(B,\beta_1)}{\rho(A,\alpha_2)} \leq \frac{1}{K}$, while $\bar{q}(B,u,a)$ is bounded from below given W from Lemma 1. This implies that, for K large enough relative to W, the rate $\bar{q}(B,d,b)$ is smaller than $\bar{q}(B,u,a)$; similarly for $\bar{q}(A,u,a)$ and $\bar{q}(A,d,b)$. The bound for K depends on g, since the lower bound in Lemma 1 depends on g. ■

From Lemma 5, for $K \geq \bar{K}(g,W,M)$,

$$\sqrt{\bar{p}(A,\alpha_2)} - \sqrt{\bar{p}(A,\beta_2)} > 0 \text{ and } \sqrt{\bar{p}(B,\beta_2)} - \sqrt{\bar{p}(B,\alpha_2)} > 0; \tag{27}$$

where the first inequality follows since (i) there are more A votes among the voters with z message in α_2 than in β_2 by the first part of the lemma and by (ii) there are more A votes from the voters with b message in α_2 than A votes from voters with a message in β_2 by the second part of the lemma. The second inequality follows analogously.

For the next lemma, note that Lemma 4 implies a bound $\bar{M}(W)$ that is uniform in all other parameters (g,r,K).

Lemma 6 *There exists some $\bar{W}$ such that for $W=\bar{W}$, all (M,g,r), and all $K \geq \bar{K}\left(g,\bar{W},M\right)$, any sequence of fixed points of the truncated best responses from the box $\Pi\left(K,M,\bar{W}\right)$ satisfies the equal-margins-condition*

$$\sqrt{\bar{p}(A,\alpha_2)} - \sqrt{\bar{p}(B,\alpha_2)} = \sqrt{\bar{p}(B,\beta_2)} - \sqrt{\bar{p}(A,\beta_2)} > 0. \tag{28}$$

Moreover, the W bound does not bind,

$$\frac{1}{\bar{W}} < \lim_{n\to\infty} \frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)} < \bar{W}.$$

Proof. From Lemma 1, the voting rates $\bar{p}(x,\omega_2)$ are all strictly positive. Hence, as discussed after Lemma 4, we can use Lemma 2 in the following.

We argue that it must hold that

$$(\sqrt{\bar{p}(A,\alpha_2)} - \sqrt{\bar{p}(B,\alpha_2)})^2 = (\sqrt{\bar{p}(B,\beta_2)} - \sqrt{\bar{p}(A,\beta_2)})^2 \tag{29}$$

Suppose not. In particular, suppose that

$$(\sqrt{\bar{p}(A,\alpha_2)} - \sqrt{\bar{p}(B,\alpha_2)})^2 > (\sqrt{\bar{p}(B,\beta_2)} - \sqrt{\bar{p}(A,\beta_2)})^2. \tag{30}$$

Then, Lemma 2 implies

$$\lim_{n\to\infty}\frac{\Pr\left(PivA|\alpha_2\right)}{\Pr\left(PivB|\beta_2\right)}=0.$$

Therefore, the sequence of truncated fixed points satisfies (note the inversion of the ratio),

$$\lim_{n\to\infty}\frac{\rho\left(B,\beta_2\right)}{\rho\left(A,\alpha_2\right)}=W. \tag{31}$$

For W very large, we claim that this implies that the best response voting rates for B are much higher than for A. That is, $\frac{\bar{p}(B,\alpha_2)}{\bar{p}(A,\alpha_2)}$ and $\frac{\bar{p}(B,\beta_2)}{\bar{p}(A,\beta_2)}$ diverge to infinity as W goes to infinity. This claim requires some arguments:

First, for those with z messages, we have

$$\frac{\bar{q}\left(B,s,z\right)}{\bar{q}\left(A,s,z\right)}\to\infty \tag{32}$$

for $W\to\infty$, since $\bar{q}\left(A,s,z\right)$ is bounded from above,

$$\bar{q}\left(A,s,z\right)\leq\Pr(\alpha_2|u,z) \tag{33}$$

while

$$\bar{q}\left(B,s,z\right)\geq m_B\frac{\rho\left(B,\beta_2\right)}{\rho\left(A,\alpha_2\right)}\Pr\left(\beta_2|s,z\right)\to\infty. \tag{34}$$

Second, for those with a messages, since $K\geq\bar{K}\left(g,W,M\right)$, Lemma 5 yields

$$\frac{\bar{q}\left(B,s,a\right)}{\bar{q}\left(A,s,a\right)}\geq 1. \tag{35}$$

Third, observe that

$$\bar{q}(A,s,a)\leq\frac{\rho(A,\alpha_1)}{\rho(A,\alpha_2)}\leq\frac{\rho(A,\alpha_1)}{\rho(B,\beta_2)}\frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)}\leq\frac{1}{K}\frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)} \tag{36}$$

Then,

$$\frac{\bar{p}(A,\beta_2)}{\bar{p}(B,\beta_2)} = \frac{\sum_s \Pr(s|\beta)\Big[g(\bar{q}(A,s,z)+(1-g)\bar{q}(A,s,a)\Big]}{\sum_s \Pr(s|\beta)\Big[g\bar{q}(B,s,z)+(1-g)\bar{q}(B,s,a)\Big]} \tag{37}$$

$$\leq \frac{\sum_s \Pr(s|\beta)\Big[g(\bar{q}(A,s,z)+(1-g)\bar{q}(A,s,a)\Big]}{\sum_s \Pr(s|\beta)g\bar{q}(B,s,z)} \tag{38}$$

$$\leq \frac{\Pr(\alpha_2|u,z)}{m_B\frac{\rho(B,\beta_2)}{\rho(A,\beta_2)}\Pr(\beta_2|d,z)} + \frac{(1-g)}{gm_B\Pr(\beta_2|d,z)K}, \tag{39}$$

where, for the last inequality, we used (33) and (34) for the first summand and (34) and (36) for the second summand.

Now, the first summand of the left-hand side of (37) converges to zero as $W\to\infty$, given (31). Further, $W\to\infty$ implies $K\to\infty$, and thus that the second summand of the left-hand side goes to zero as well. Thus, $\frac{\bar{p}(B,\beta_2)}{\bar{p}(A,\beta_2)}$ goes to infinity when W grows large. An analogous argument shows that also $\frac{\bar{p}(B,\alpha_2)}{\bar{p}(A,\alpha_2)}$ diverges as W grows large.

As a consequence of $\frac{\bar{p}(B,\alpha_2)}{\bar{p}(A,\alpha_2)}$ and $\frac{\bar{p}(B,\beta_2)}{\bar{p}(A,\beta_2)}$ becoming arbitrarily large for W large (and K large),

$$\sqrt{\bar{p}\,(B,\alpha_2)} > \sqrt{\bar{p}\,(A,\alpha_2)}. \tag{40}$$

From the ordering in (27), the difference $\sqrt{\bar{p}\,(B,\omega_2)}-\sqrt{\bar{p}\,(A,\omega_2)}$ is even higher in β_2,

$$\sqrt{\bar{p}\,(B,\beta_2)}-\sqrt{\bar{p}\,(A,\beta_2)} > \sqrt{\bar{p}\,(B,\alpha_2)}-\sqrt{\bar{p}\,(A,\alpha_2)}.$$

Since the right side is also strictly positive by (40), we get that

$$(\sqrt{\bar{p}\,(B,\beta_2)}-\sqrt{\bar{p}\,(A,\beta_2)})^2 > (\sqrt{\bar{p}\,(A,\alpha_2)}-\sqrt{\bar{p}\,(B,\alpha_2)})^2,$$

which contradicts our starting hypothesis (30). An analogous argument excludes the converse case. Thus, (29) must hold.

Given the ordering from Lemma 5, this implies (28) whenever W is above some $\bar{W}$.

The same argument also implies that W must not be binding when W is above $\bar{W}$: If it were binding, then the implied inequality (40) implies a contradiction to (28). ■

4.6 The K-Bound Does not Bind

Lemma 7 *For every $\delta \in (0,1)$, there exists some $\bar{g}(\delta) < \frac{1}{2}$ such that for $g \in (0, \bar{g}(\delta))$, $W = \bar{W}$, and all other parameters (K, M, r), for any sequence of fixed points of the truncated best responses from the box $\Pi(K, M, W)$,*

$$\begin{aligned} g\bar{q}(B,d,a) &< \delta\bar{q}(A,d,b). \\ g\bar{q}(A,u,b) &< \delta\bar{q}(B,u,a) \end{aligned}$$

Proof. From $y \leq 1$,

$$U(B,d,a,y) \leq \rho(B,\beta_2)\Pr(\beta_2|d,a)$$

and so

$$\begin{aligned} g\bar{q}(B,d,a) &\leq g\frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)}\Pr(\beta_2|d,a) \\ &\leq g\bar{W}\frac{g\varepsilon\Pr(\beta)r}{g\varepsilon\Pr(\beta)r + (1-g)(1-r)\Pr(\alpha)}. \end{aligned}$$

Also, as in Lemma 1,

$$\bar{q}(A,d,b) \geq m_A \Pr(\alpha_2|d,b) = m_A\frac{g\varepsilon\Pr(\alpha)r}{g\varepsilon\Pr(\alpha)r + (1-g)(1-r)\Pr(\alpha)}.$$

Cancelling g, given $\bar{W}$, there is some $\bar{g}(\delta) < \frac{1}{2}$ such that for all $g \leq \bar{g}(\delta)$ and all r,

$$g\bar{W}\frac{\varepsilon\Pr(\beta)r}{g\varepsilon\Pr(\beta)r + (1-g)(1-r)\Pr(\alpha)} \leq \delta m_A\frac{\varepsilon\Pr(\alpha)r}{g\varepsilon\Pr(\alpha)r + (1-g)(1-r)\Pr(\alpha)}.$$

This implies the first inequality from the lemma. The second follows from the same idea. ■

Including r, K as arguments, note that $\bar{q}(x,s,m,r,K)$ is uniformly bounded from above (given some fixed W). Hence, we can pick some converging subsequences with finite limits, denoting them

$$\bar{q}^*(x,s,m) = \lim_{r\to\frac{1}{2}^+}\lim_{K\to\infty}\bar{q}(x,s,m,r,K) \text{ for all } x,s,m.$$

Lemma 8 *For fixed g, W, M and all $m \in \{a, b, z\}$,*

$$\bar{q}^*(x, d, m) = \bar{q}^*(x, u, m)$$

and

$$\bar{q}^*(A, s, a) = \bar{q}^*(B, s, b) = 0 \text{ for } s \in \{u, d\}.$$

The lemma is immediate from the fact that $s \in \{u, d\}$ contains vanishing information for $r \to \frac{1}{2}$, and its proof therefore omitted. In the following, we drop s from the argument and write $\bar{q}^*(x, m)$ in the following. The lemma implies that, with the notation

$$\bar{p}^*(x, \omega_i) := \lim_{r \to \frac{1}{2}^+} \lim_{K \to \infty} \bar{p}(x, \omega_i),$$

we have

$$\begin{aligned}
\bar{p}^*(A, \alpha_2) &= g\bar{q}^*(A, b) + (1-g)\bar{q}^*(A, z) \\
\bar{p}^*(B, \alpha_2) &= (1-g)\bar{q}^*(B, z) \\
\bar{p}^*(B, \beta_2) &= g\bar{q}^*(B, a) + (1-g)\bar{q}^*(B, z) \\
\bar{p}^*(A, \beta_2) &= (1-g)\bar{q}^*(A, z) \\
\bar{p}^*(A, \beta_1) &= \bar{q}^*(A, b) \\
\bar{p}^*(B, \beta_1) &= 0.
\end{aligned}$$

Combining this with Lemma 6 (which applies since $K \geq K(g, W, M)$ by $K \to \infty$), for $W = \bar{W}$ and fixed M, we have

$$\begin{aligned}
EM^* \quad : &= \sqrt{g\bar{q}^*(A, b) + (1-g)\bar{q}^*(A, z)} - \sqrt{(1-g)\bar{q}^*(B, z)} \\
&= \sqrt{g\bar{q}^*(B, a) + (1-g)\bar{q}^*(B, z)} - \sqrt{(1-g)\bar{q}^*(A, z)} \geq 0.
\end{aligned}$$

Lemma 9 *Given $W = \bar{W}$ and $M = \bar{M}(\bar{W})$, there exists some $\bar{g}_x < \frac{1}{2}$ such that for all $g \leq \bar{g}_x$,*

$$\begin{aligned}
\sqrt{\bar{q}^*(A, b)} &> EM^*, \\
\sqrt{\bar{q}^*(B, a)} &> EM^*.
\end{aligned}$$

Proof. Suppose

$$\bar{q}^*(B,z) \leq \bar{q}^*(A,z).$$

Then,

$$(1-g)\,\bar{q}^*(B,z) \leq (1-g)\,\bar{q}^*(A,z) < (1-g)\,\bar{q}^*(A,z)+g\bar{q}^*(A,b) \leq (1-g)\,\bar{q}^*(B,z)+g\bar{q}^*(B,a),$$

where the first inequality is by assumption, the strict inequality is by $g\bar{q}^*(A,b) > 0$, and the last inequality is from the equal-margin condition as follows: Let $a' = (1-q)\bar{q}^*(A,z)$, $b' = (1-g)\bar{q}^*(B,z)$, $c' = g\bar{q}^*(B,a)$, and $d' = g\bar{q}^*(A,b)$. Then, the equal-margin condition implies that $\sqrt{a'+d'}-\sqrt{b'} = \sqrt{b'+c'}-\sqrt{a'}$; compare to the expression for the margin EM^* right before Lemma 9. Thus, $\sqrt{a'+d'} = \sqrt{b'} - \sqrt{a'} + \sqrt{b'+c'}$, which implies $\sqrt{a'+d'} \leq \sqrt{b'+c'}$, given the assumption that $b' \leq a'$. Thus, $a'+d' \leq b'+c'$, which is what the last inequality above states. Therefore, subtracting $(1-g)\,\bar{q}^*(B,z)$

$$0 \leq (1-g)\,\bar{q}^*(A,z) + g\bar{q}^*(A,b) - (1-g)\,\bar{q}^*(B,z) \leq g\bar{q}^*(B,a). \tag{41}$$

Note that, for any positive numbers $a', b', c', d' \geq 0$ (overwriting the notation from above now), the inequality $0 \leq a'-b' < c'$ implies $\sqrt{a'}-\sqrt{b'} \leq \sqrt{b'+c'} - \sqrt{b'} \leq \sqrt{c'}$, where the first inequality is simply rewriting the condition to $a' \leq b'+c'$, using monotonicity of $\sqrt{\cdot}$, and finally subtracting $\sqrt{b'}$ and the second inequality is from $\sqrt{b'+c'}-\sqrt{b'}$ decreasing in b' by strict concavity of $\sqrt{\cdot}$ and by $b' > 0$ which we showed in Lemma 1. Therefore, (41) implies

$$EM^* = \sqrt{(1-g)\,\bar{q}^*(A,z) + g\bar{q}^*(A,b)} - \sqrt{(1-g)\,\bar{q}^*(B,z)} < \sqrt{g\bar{q}^*(B,a)}.$$

For $g \leq \bar{g}(1/2)$, Lemma 7 implies that $g\bar{q}^*(B,a) < \bar{q}^*(A,b)$. Hence, $\sqrt{g\bar{q}^*(B,a)} < \sqrt{\bar{q}^*(A,b)}$ and so, the claim follows,

$$EM^* < \sqrt{\bar{q}^*(A,b)}.$$

Now, suppose

$$\bar{q}^*(B,z) \geq \bar{q}^*(A,z).$$

Then,

$$(1-g)\,\bar{q}^*(A,z) \leq (1-g)\,\bar{q}^*(B,z) < (1-g)\,\bar{q}^*(B,z)+g\bar{q}^*(B,a) \leq (1-g)\,\bar{q}^*(A,z)+g\bar{q}^*(A,b)\,. \tag{42}$$

Therefore, comparing the last inequality and subtracting $(1-g)\,\bar{q}^*(A,z)$ from both sides

$$0 \leq (1-g)\,\bar{q}^*(B,z) + g\bar{q}^*(B,a) - (1-g)\,\bar{q}^*(A,z) < g\bar{q}^*(A,b)\,.$$

By the analogous argument as above, this implies $EM^* \leq \sqrt{g\bar{q}^*(A,b)}$. Hence, again, the claim of the lemma follows for $\bar{q}^*(A,b)$.

The second inequality in the lemma involving $\bar{q}^*(B,a)$ can be established analogously. ■

Lemma 10 *For all g with $g \leq \bar{g}_x < \frac{1}{2}$, there exists some $\bar{r}(g)$ such that for all r with $\frac{1}{2} < r \leq \bar{r}(g)$, there is some $\bar{K}(r,g) \geq \bar{K}(g,\bar{W},\bar{M}(\bar{W}))$ such that when $W=\bar{W}$, $M=\bar{M}\left(\bar{W}\right)$, and $K \geq \bar{K}(r,g)$:*

$$\begin{aligned} \sqrt{\bar{p}(A,\alpha_1)} - \sqrt{\bar{p}(B,\alpha_1)} &> \sqrt{\bar{p}(B,\beta_2)} - \sqrt{\bar{p}(A,\beta_2)} > 0\,, \\ \sqrt{\bar{p}(A,\beta_1)} - \sqrt{\bar{p}(B,\beta_1)} &> \sqrt{\bar{p}(A,\alpha_2)} - \sqrt{\bar{p}(B,\alpha_2)} > 0. \end{aligned}$$

Proof. Let $W=\bar{W}$, $M=\bar{M}\left(\bar{W}\right)$ and fix some $g \leq \bar{g}_x$. Note that

$$\sqrt{\bar{p}^*(A,\alpha_1)} - \sqrt{\bar{p}^*(B,\alpha_1)} = \sqrt{\bar{q}^*(A,a)} - \sqrt{\bar{q}^*(B,a)} = \sqrt{\bar{q}^*(B,a)}$$

where the first equality is from by definition of $\bar{p}^*$ and the second from Lemma 8. Hence, from Lemma 9, for any converging subsequence, $g \leq \bar{g}_x$ implies

$$\lim_{r\to\frac{1}{2}^+}\lim_{K\to\infty}\sqrt{\bar{p}(A,\alpha_1,r,K)}-\sqrt{\bar{p}(B,\alpha_1,r,K)} > \lim_{r\to\frac{1}{2}^+}\lim_{K\to\infty}\sqrt{\bar{p}(B,\beta_2,r,K)}-\sqrt{\bar{p}(A,\beta_2,r,K)}.$$

Hence, there exists some $\bar{r}(g)$ such that for all r with $\frac{1}{2} < r \leq \bar{r}(g)$ there is some $\bar{K}(r,g)$ such that for all $K \geq \bar{K}(r,g)$,

$$\sqrt{\bar{p}(A,\alpha_1,r,K)} - \sqrt{\bar{p}(B,\alpha_1,r,K)} > \sqrt{\bar{p}(B,\beta_2,r,K)} - \sqrt{\bar{p}(A,\beta_2,r,K)}.$$

The second inequality can be shown analogously. ■

Lemma 11 *For all g with $g \leq \bar{g}_x < \frac{1}{2}$, there exists some $\bar{r}(g)$ such that for all r with $\frac{1}{2} < r \leq \bar{r}(g)$, there is some $\bar{K}(r,g)$ such that when $W = \bar{W}$, $M = \bar{M}\left(\bar{W}\right)$, and $K \geq \bar{K}(r,g)$: For every sequence of fixed points of the truncated best responses ρ and their induced behavior σ^{ρ},*

$$\lim_{n\to\infty} \frac{\Pr\left(PivB|\beta_1;\sigma^{\rho}\right) + \Pr\left(PivA|\beta_1;\sigma^{\rho}\right)}{\Pr\left(PivA|\alpha_2;\sigma^{\rho}\right)} = 0$$

Proof. Recall the discussion after Lemma 4, which explains that the conditions of Lemma 2 are satisfied. By Lemma 1, it holds $\bar{p}(A,\alpha_2)$ and $\bar{p}(B,\alpha_2) > 0$. Therefore, the claim follows from Lemma 2 and Lemma 10. ■

4.7 Proof of Theorem 1

From Lemmas 4, 6, and 11, when $g \leq \bar{g}_x$ and $r \leq \bar{r}(g)$, then, for $W = \bar{W}$, $M = \bar{M}\left(\bar{W}\right)$, and $K = \bar{K}(g,r)$, for n large enough, none of the constraints from the truncation are binding, implying that the fixed points of the truncated best responses are interior and, hence, equilibria. This proves the existence of a sequence of manipulated equilibria for g small enough and r close enough to $\frac{1}{2}$.

This observation holds for all $\varepsilon > 0$, where ε was the parameter of the sender's message. Thus, for any given ξ, we can choose some $\varepsilon < \frac{1}{2}\xi$ and then some $g \leq \bar{g}_x$ (with $\bar{g}_x$ defined in Lemma 9) such that for all $r \leq \bar{r}(g)$ (with $\bar{r}(g)$ defined in Lemma 10) there exists a sequence of equilibria with the property that, with probability $1-\varepsilon$, the substates α_1 and β_1 realize and, for n large enough, with a probability larger than $1-\varepsilon$, in these equilibria a strict majority votes for B in α and A in β, respectively. Now, note that Lemma 1 implies that $\bar{p}(A,\omega_i) + \bar{p}(B,\omega_i) > 0$ for $\omega_i \in \{\alpha_1, \beta_1, \alpha_2, \beta_2\}$ and that $\bar{p}(A,\alpha_2) \leq W$. Therefore $\lim_{n\to\infty} \frac{\lambda_A(\omega_i)+\lambda_B(\omega_i)}{\lambda_A(\alpha_2)} = \frac{\bar{p}(A,\omega_i)+\bar{p}(B,\omega_i)}{\bar{p}(A,\alpha_2)} \in \mathbb{R}^{>0}$ Therefore, Lemma 3 implies that $\lambda(A,\omega_i) + \lambda(B,\omega_i) \to \infty$ as $n \to \infty$. This means that the expected number of participants grows large in every substate. Finally, an application of the law of large numbers proves the theorem.

5 Full Manipulation with Few Partisans

One can derive a similar result when the share of partisans is sufficiently small.

Theorem 2 *Given any precision* $r \in \left(\frac{1}{2}, 1\right)$ *and prior* $\Pr(\alpha) \in (0,1)$*: For every* $\xi > 0$ *and* $\bar{y}$ *with* $0 < \bar{y} < \frac{1}{2}$*, there exists some* $\Delta > 0$ *such that for all type distributions* F *for which* $F(\bar{y}) + F(1-\bar{y}) < \Delta$*, there are some parameters* (g, ε) *for the sender's signal and an equilibrium* σ *such that when* n *is large enough,*

$$\begin{aligned} \Pr(B \textit{ wins majority} | \alpha; \sigma, n) &\geq 1 - \xi, \\ \Pr(A \textit{ wins majority} | \beta; \sigma, n) &\geq 1 - \xi. \end{aligned}$$

When there are no partisans, $F(0) = 1 - F(1^-) = 0$*, then there are some parameters* (g, ε) *for the sender's signal and an equilibrium* σ *such that when* n *is large enough,*

$$\begin{aligned} \Pr(B \textit{ wins majority} | \alpha; \sigma, n) &= 1, \\ \Pr(A \textit{ wins majority} | \beta; \sigma, n) &= 1. \end{aligned}$$

The proof of the theorem is in the appendix. It uses many of the basic techniques from Theorem 1 to characterize equilibrium with costly voting. To keep the proof short, we only prove the theorem under the simplifying assumption that the setting is symmetric across states and outcomes.

With few partisans, the proof relies on ideas from the swing voter's curse. The basic idea is that voters who obtain the uninformative message z and have intermediate types $y \in [\bar{y}, 1-\bar{y}]$ abstain and thereby "delegate" the decision to those with messages a/b. This is rationalized by the fact that in state α_2, only voters with message b participate and almost all of them vote "correctly" for A. Similarly, in state β_2, only voters with message a participate and almost all of them vote "correctly" for B. Hence, voters with an uninformative signal z can only add noise. Formally, there is 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 a vote for B is much more likely to be pivotal in α_2 than in β_2. Given this swing voters' curse, voters with z signals and interior y strictly prefer abstaining.

As a consequence of the abstention of the voters with $y \in [\bar{y}, 1-\bar{y}]$, the outcome is almost completely decided by the voters with an a or b message. Importantly, in states α_2 and β_2, total turnout is smaller than in α_1 and β_1. This is because only a share $g < 1$ obtains an a or b message, respectively. In states α_1 and β_1, however, all voters obtain these messages, and so turnout is larger. The critical observation is that, therefore, a vote is much more likely to be pivotal in α_2 and in β_2 (with small turnout) than in states α_1 and β_1. Hence, conditional on being pivotal, voters with an a or b signal become convinced that the state is β_2 and α_2, respectively.

Note that this is similar to the effect in Ekmekci and Lauermann (2022), who consider elections in which the expected number of potential voters is state-dependent. As here, the critical effect is that a vote is more likely to be pivotal in the state with lower expected voter numbers, keeping everything else constant.

6 Conclusion

We asked a narrow question: Under what conditions does a message with a structure as in Heese and Lauermann (2023) allow to invert the full information outcome with probability 1.

We expect that one can use the previous arguments to show that, for all environments (preference distributions and private signal precisions), there is a message structure that allows *some* manipulation, in the sense that the additional information upsets the full information equivalence result in Feddersen and Pesendorfer (1997).

An interesting open question is the effectiveness of other message structures. What is the set of (stochastic) state-dependent outcomes that is implemented in some equilibrium for some message structure? It may well be that when the share of partisans is sufficiently large and nature's signal sufficiently precise, there is a constraint on the ability of persuasion. We leave this conjecture for future research.

References

Abramowitz, M., and I. A. Stegun (1968): *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55. US Government printing office.

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

Ekmekci, M., and S. Lauermann (2022): "Information aggregation in Poisson elections," *Theoretical Economics*, 17(1), 1–23.

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

Heese, C., and S. Lauermann (2023): "Persuasion and Information Aggregation in Large Elections," Discussion paper.

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

A Appendix: Proof of Theorem on Few Partisans

To improve readability and focus on the main ideas, we simplify the model by considering a state-symmetric setting (symmetry of the prior, the distributions of the types and signals, and the strategies) and by assuming that all voters with types outside the interval $[\bar{y}, 1-\bar{y}]$ are partisans with types $y=0$ and $y=1$. Their combined mass is denoted by Δ, and we aim to show that for sufficiently small Δ, a sender can manipulate the election outcome with high probability.

The theorem for the general case follows from the same lines of reasoning, albeit with more notation and case distinctions.

Proof. The proof requires familiarity with the proof of Theorem 1 because it uses its notation and many of its arguments.

Let $\bar{M}$ be sufficiently large such that if the swing voter's curse (SVC) is larger than $\bar{M}$, then voters with intermediate types $y \in [\bar{y}, 1-\bar{y}]$ and receiving a z message have a strict preference to abstain. Formally, we require that for any $y \geq \bar{y}$,

$$U(A|s,z,y) < 0,$$

where $U(A|s,z,y)$ is the expected utility of voting for A given signal s, message z, and type y. Let $\bar{M}_0$ denote the threshold above which this condition holds.

We seek equilibria where voters with z messages participate only if their types are at the extremes ($y=0$ or $y=1$). The participation rate of these voters relative to the pivotal probability ρ is given by

$$q(A,u,z) = r\Delta \quad \text{and} \quad q(A,d,z) = (1-r)\Delta,$$

and by symmetry,

$$q(B,d,z) = r\Delta \quad \text{and} \quad q(B,u,z) = (1-r)\Delta.$$

For voters with b messages, we have

$$U(A|s,b,y) = \rho(A,\alpha_2)\frac{\Pr(\alpha_2|s,b)}{\Pr(s,b)}(1-y) - \rho(A,\beta_1)\frac{\Pr(\beta_1|s,b)}{\Pr(s,b)}y.$$

We consider equilibria where $\frac{\rho(A,\beta_1)}{\rho(A,\alpha_2)} \to 0$, leading to

$$\bar{q}(A,s,b) = \frac{\Pr(\alpha_2|s,b)}{\Pr(s,b)} \left(1 - \mathbb{E}[y]\right),$$

and by symmetry, $\bar{q}(A,s,b) = \bar{q}(B,-s,a)$.

Given that $\mathbb{E}[y] = \frac{1}{2}$, we have

$$\bar{q}(A,s,b) = \frac{\Pr(\alpha_2|s,b)}{\Pr(s,b)} \frac{1}{2}.$$

Define $\bar{p}_A^\Delta(\alpha_2)$ and $\bar{p}_B^\Delta(\alpha_2)$ as the normalized participation rates for A and B in state α_2, respectively. For sufficiently small Δ, $\frac{\bar{p}_A^\Delta}{\bar{p}_B^\Delta}$ can be made arbitrarily large, ensuring that the SVC is stronger than $\bar{M}_0$.

Now, consider any such Δ and g. We construct an equilibrium for large n by defining a box of pivotal probabilities $\Pi^\Delta(K)$ and the corresponding best responses $\hat{\sigma}^\Delta(\rho)$, which include the abstention of types outside $[\bar{y}, 1-\bar{y}]$ when receiving a z message. The truncated pivotal probabilities ρ_T^Δ are defined accordingly.

Since $\Pi^\Delta(K)$ is contractible, there exists a fixed point ρ_T^Δ, and we consider sequences of such fixed points for which the normalized participation rates converge, denoted by $\bar{p}(x,\omega_i)$.

For large enough K, we have $\bar{p}(A,\alpha_2) \approx \bar{p}_A^\Delta(\alpha_2)$, and similarly for B. This implies that for large enough n, the SVC is stronger than $\bar{M}_0$, and the K bound from the box does not bind.

Moreover, by construction, the SVC is strong enough to discourage voters with intermediate types and a z message from participating, leading to an equilibrium where essentially only voters with a and b messages participate in the ω_1 and ω_2 states, respectively.

Applying the law of large numbers, we conclude that for large enough n, the probability of B winning in state α and A winning in state β approaches 1, completing the proof of Theorem 2. ■

B Appendix: Condensed Outline and Idea of the Proof

We construct an equilibrium in which the voting rates $\bar{p}(x, \omega_i)$ are strictly positive in the substates α_2 and β_2 (when the sender provides low quality information). They are strictly positive because of the presence of a strictly positive share of partisans, that is, $m_A > 0$ and $m_B > 0$. The ability to use the vector $\bar{p}(\cdot, \cdot)$ to describe the limit is one of the main ideas of the proof. Because the $\bar{p}$ are strictly positive and bounded, we can use normal arithmetic operations (adding, subtracting, dividing, etc.) with the limit objects.

In states α_2 and β_2, the usual CJT argument implies an equal-margins-condition, and the margins are strictly positive:

$$EM := \sqrt{\bar{p}(A, \alpha_2)} - \sqrt{\bar{p}(B, \alpha_2)} = \sqrt{\bar{p}(B, \beta_2)} - \sqrt{\bar{p}(A, \beta_2)} > 0.$$

In the constructed equilibrium, most voters with an a message vote B (meaning, $\bar{p}(B, \beta_1) > \bar{p}(A, \beta_1) = 0$) and most voters with a b message vote A (meaning, $\bar{p}(B, \alpha_1) > \bar{p}(A, \beta_1) = 0$).

The equilibrium holds together because a vote for B is much more likely to be pivotal in α_2 than in β_1 and a vote for A is much more likely to be pivotal in β_2 than a vote for A in α_1 (or a vote for B in α_1, so these voters do not worry about the SVC)–plus, we don't let the ex-ante probability of ω_2 vanish. This latter point is important: If the ex-ante probability of ω_2 would vanish, then voters with a/b messages would have much lower participation incentives than voters with z messages.

This claim about the relative pivotal probabilities follows from the fact that the **absolute** winning margins are ordered as,

$$\sqrt{\bar{p}(B, \beta_1)} - \underbrace{\sqrt{\bar{p}(A, \beta_1)}}_{=0} > \sqrt{\bar{p}(A, \alpha_2)} - \sqrt{\bar{p}(B, \alpha_2)} = EM.$$

Given this ordering of the winning margins, an "approximation lemma" will imply that, indeed, $\lim_{n\to\infty} \frac{\Pr(PivB|\beta_1)}{\Pr(PivA|\alpha_2)} = 0$.

Ensuring this latter ordering of the absolute winning margin is not trivial, and this is where most of the work goes into (apart from the approximation lemma): The problem is that the a signals (who are voting for B) must turn out at sufficient numbers, even though (i) voters with an a signal will put a high probability on β_1 and, in addition, (ii) a vote is unlikely to be pivotal in β_1. Thus, their expected probability of being pivotal is small, and so their turnout incentives can be weak.

The reason the signal precision r must be small is the following. Recall from Heese and Lauermann (2023) that, in the absence of voting costs, we constructed a message and an equilibrium in which the voters with a z message behaved as in the equilibrium without a sender, as in the "modern CJT." As part of that proof, we showed an equal winning margin in α and in β. The size of the total winning margin depends on r: it is strictly positive for all $r > 1/2$ and, the larger r, the larger the winning margin. At the same time, when ε and g are small, and persuasion is effective in making voters with an a or b-message believe they are in ω_2, then the winning margin in ω_1 becomes relatively small because a and b messages turn out in very low numbers.

Thus, when r is large and ε (and g) are small, then the total margin of victory in ω_2 can then easily be larger than the total turnout of the a/b voters in ω_1 (recall their weak participation incentives), and so the approximation lemma would imply that the probability of being pivotal in ω_1 may be *higher* than in ω_2—even if almost all of the actually participating voters vote for A (or almost all vote for B).

From a technical side, what helps in this proof is the presence of partisans and the normalization (expressing everything relative to $\rho(A, \alpha_2)$). With the normalization, we have a we have well-defined limit objects. In particular, we can talk about "absolute" winning margins. The presence of partisans ensures that most of these normalized participation rates are strictly positive and all are bounded.

The construction of the equilibrium works as follows. First, we consider sequences of truncated best response pivot probabilities, that is, sequences of vectors of pivotal probabilities $\rho \in [0,1]^8$ such that these are in the "box" Π defined before and such that they are a fixed point of the truncated best response: Formally, we have a sequence with elements ρ such that $\rho \in \Pi(K, M, W)$ and $\hat{\rho}_T(\hat{\sigma}^\rho) = \rho$, that is, given the voters best response to ρ, $\hat{\sigma}^\rho$, the truncation of the induced pivot probabilities to the box, $\hat{\rho}_T(\cdot)$, equals ρ.

(For the proof, we do not include indices for the sequence, and we also do not explicitly use $\hat{\rho}_T$.)

We establish the following results:

1. Lemma 2: The approximation lemma characterizes the SVC (the relative probabilities of being pivotal with an A and a B vote in a given substate) and it shows that the ordering of the absolute margin for victory across substates determines the relative probabilities of being pivotal.

   The hard part of the lemma was dealing formally with the possibility that the vote rate in one (sub-)state may vanish. This is because we have no general explicit asymptotic approximation for this case (we cannot so easily say $\Pr(\text{Piv}) \approx \text{Expression}$). However, we verify and use the intuitive fact that the probability of being pivotal is bounded by the probability of being pivotal when the vote rate is arbitrarily small but fixed.

2. Lemma 1: For every sequence of truncated best responses ρ, the participation rates $\bar{p}$ are bounded and bounded away from 0, except possibly $\bar{p}(A, \beta_1)$ and $\bar{p}(B, \alpha_1)$. These bounds are explicitly stated and uniform in all parameters except W. This lemma utilizes the presence of partisans.

3. Lemma 3: The number of participants in ω_2 is unbounded as $n \to \infty$.

4. Lemma 4: The fact that the vote rates are uniformly bounded implies that the SVC is bounded, and so we can choose M large enough so that the M bound of the box does not bind (recall that $\frac{1}{M} \leq \frac{\rho(B,\alpha_2)}{\rho(A,\alpha_2)} \leq M$).

5. Lemma 5: The vote rates are ordered. Voters with a z message are more likely to vote A with an u signal (when they believe α is more likely). When K is large enough, voters with a/b messages are more likely to vote against that message than in favor of it.

6. Lemma 6: Using arguments as in the original CJT, we show that the winning margins in ω_2 must be ordered, A wins in α_2 and B in β_2, strictly positive, and equal. This lemma also establishes that the W bound does not bind for W large enough (recall that $\frac{1}{W} \leq \frac{\rho(B,\beta_2)}{\rho(A,\alpha_2)} \leq W$).

   The lemma again uses the presence of partisans: If the upper W bound would bind for W large, then the B partisans would participate at much higher rates than the A partisans—and in fact, there would be many more B votes than A votes. So, B would win in α_2 and in β_2. However, a vote would be much less likely to be pivotal in β_2 than in α_2—in contradiction to W being large. The same argument implies the equal-margins-condition (the proof shows the EMC first).

As in the original CJT, we need some ordering of voting rates to argue that it cannot be that the margins are equal in the trivial sense (all the non-partisans vote in a particular way). For this, we use Lemma 5.

7. Lemma 8: When r is small, the voting rates of a given type of voter (y preference and m signal) are not very different across states because they don't condition much on their signal (u versus d).

   Also, when K is large, most voters with a and b messages vote the opposite of their message (rather than equal to it)–if they vote at all.

8. Lemma 9: Everything comes together. We show:

$$\sqrt{\bar{p}\left(B,\beta_1\right)}-\underbrace{\sqrt{\bar{p}\left(A,\beta_1\right)}}_{=0}>\sqrt{\bar{p}\left(A,\alpha_2\right)}-\sqrt{\bar{p}\left(B,\alpha_2\right)}=EM.$$

   As a consequence, the K bound does not hold.

   For g small enough, there are few a and b message voters in ω_2 whereas all have such messages in ω_1. Importantly, the winning margin in ω_2 is proportional to the number of a and b messages. This is because the EMC implies that the z signals must balance simultaneously the A votes in α_2 and the B votes in β_2, while A wins in α_2 and B in β_2. Now, the key is that, for small r, the vote rates among the z signals are essentially the same in α_2 and β_2. Therefore, among the z signals the difference of votes for A and B must be relatively small and of the order of the number of participating voters with a/b messages, that is, the margin is on the order of $g\bar{q}\left(A,b\right)$ and $g\bar{q}\left(B,a\right)$, and this is smaller than the margin in ω_1 states, which is essentially $\bar{q}\left(B,a\right)$ and $\bar{q}\left(A,b\right)$ in α_1 and β_1, respectively.

   Remark: A (minor) problem in the proof is ensuring that vote rates for A and B among the a, b messages are of similar magnitude. This is what Lemma 7 is for.

9. Since we already proved that the M and W bounds don't bind, we have established that none of the constraints from the Box bind. Thus, for n large enough, the truncated best responses are, in fact, equilibria.

10. Since $K \to \infty$, the a and b messages are voting almost exclusively for B and A and hence, the outcome is the manipulated one in which the full information outcome is inverted.