Fragility of The Condorcet Jury Theorem: Information Aggregation and Preference Aggregation

Masayuki Odora

Fragility of The Condorcet Jury Theorem: Information Aggregation and Preference Aggregation[◊]

Masayuki Odora[⬥]

January 2024

Abstract: This study considers a binary election in which imperfectly informed voters have partially conflicting interests. There is an unambiguously correct alternative in some states, while voters disagree on the better alternative in other states. The true state is unknown to anybody, but each voter receives a private signal about the state. This study identifies the circumstances in which the probability that a society utilizing the majority rule reaches the correct decisions does not converge to 1, thus showing the failure of an asymptotic Condorcet Jury Theorem. Moreover, we show that the voting behavior never reflects voters' private information in the large elections.

Keywords: Information aggregation, partially conflicting interests, Condorcet Jury Theorem

---

1. Introduction

Classical asymptotic Condorcet Jury Theorem (CJT) states that a group of voters utilizing a majority rule identifies the correct alternative with arbitrary precision as the size of the group tends to infinity. The original CJT considered a binary election in which there is an unambiguously correct alternative (i.e., the alternative that serves the interests of all voters), but voters are unaware of the identity of it. Each voter, not knowing which alternative is correct, has some imperfect information about the merit of the alternatives.

An important premise for the CJT to hold is that a group of voters have common interests in the following sense. Voters never disagree on the better alternative once the underlying uncertainty is resolved (Austen-Smith & Banks, 1996; Gerardi, 2000; Wit, 2000). For instance, consider the jury setup. A group of jurors, not knowing whether the defendant is guilty or innocent, may disagree about whether to convict or not simply because each juror has different information. However, once the uncertainty about the defendant is resolved, they never disagree on the correct decision, i.e., convicting the guilty and acquitting the innocent.

However, many real-world elections in which large number of voters participate, such as national elections or referendums, do not fit the common-value setting. In real-world elections, voters may disagree even in the absence of uncertainty, in which case there is no longer any *unambiguously correct* alternative. This study considers a binary election in which voters may or may not agree on the better alternative, depending on the state, and shows the following: the probability that the group of voters identifies the correct alternative under majority rule does not converge to 1 even when the number of voters tends to infinity if a prior probability that voters disagree is large enough. Thus, this study provides circumstances in which the asymptotic CJT fails to hold.

As an example of our binary elections, consider elections in which two candidates who differ in their ideological positions and qualities (or valence) compete for a single office and voters care about both their ideological positions and qualities. If the qualities of the candidates differ significantly, then voters would *unanimously* prefer the higher-quality one, which is an *unambiguously correct* alternative in this case. By contrast, if the difference in the qualities is small, voters would prefer the candidate whose ideological positions are closer to their ideal point, implying that there is no *unambiguously correct* alternative. The qualities are often private information, while the ideological positions are not, so voters cannot tell whether the elections are matter of truth or taste. We call this type of preference heterogeneity as *partially conflicting interests* (Schulte 2012; Odora, 2023).

To be more precise, this study considers the following setup. A group of imperfectly informed voters must choose one of two alternatives (denoted by A and B) by the majority rule, where the alternatives correspond to the candidates in the previous example. There is an unknown state; in the words of the previous example, the state corresponds to the quality of the candidate B minus that of the candidate A. The voters' payoff depends on the collective decision, the state, and a binary preference parameter (labeled $\mathcal{A}$ and $\mathcal{B}$). The preference parameter captures voters' preferred alternatives in the event that there is no correct alternative and therefore it is interpreted as the ideal point. In particular, both type $\mathcal{A}$ and type $\mathcal{B}$ voters agree on the better alternative when the state is extreme value (i.e., the absolute value of quality

difference is large). By contrast, when the state is intermediate value (i.e., the quality difference is around zero) type $\mathcal{A}$ (type $\mathcal{B}$) voters prefer A (B). Each voter receives a binary signal, correlated with the true state. This study focuses on the symmetric equilibria, in which (1) all voters use the same strategies and (2) both types of voters use the same strategies.

In this setup, this study investigates whether the group of voters collectively identifies correct alternatives with arbitrary precision when the number of voters goes to infinity. This study shows that the correct alternative is not identified asymptotically when the prior probability that voters disagree on the better alternatives is large enough (Theorem 3). Moreover, in large elections, an equilibrium voting behavior is based on their preference type, rather than their signals (Theorem 2). In other words, any sequence of symmetric equilibrium converges to a partisan voting equilibrium, which is the pure strategy in which type $\mathcal{A}$ ($\mathcal{B}$) voters vote for A (B). This is true irrespective of the size of the state in which voters disagree.

1.1. Related Literature

Early literature on CJT investigated the statistical property of the likelihood that the majority of voters vote correctly, presuming (1) the common value setting and (2) a sincere voting, where sincere voting means that each voter votes as if she alone could determine the collective outcome (see, e.g., Berend & Paroush, 1998 and references therein). An important exception is Miller (1986) who considered a binary election in which voters disagree on the better alternative, but Miller maintained the sincere voting assumption.

Austen-Smith and Banks (1996) shows that the sincere voting assumption may be inconsistent with equilibrium behavior even when voters have completely identical preference, which left us with the questions about the validity of CJT in a strategic environment. Since then, a game theoretic version of CJT has been reestablished by allowing mixed strategy equilibrium (Wit, 2000; McLennan, 1998, Gerardi, 2000). However, many game theoretic literature on CJT maintained the common-value assumptions, where this assumption means that voters never disagree once the uncertainty is resolved (Austen-Smith & Banks 1996; McLennan, 1998; Gerardi, 2000; Wit, 2000).

Exceptions are Feddersen and Pesendorfer (1997) (henceforth FP) and Bhattacharya (2013), both of which considered the game theoretic version of CJT assuming that voters have heterogenous preferences. In Bhattacharya (2013), voters have a binary preference type, and the interests of each type are diametrically opposed in a sense that both types of voters always disagree on better alternatives once the uncertainty is resolved. He shows that a monotonicity of preference is crucial in order for the voting mechanism to approximate the outcome that would have been chosen if the true state were common knowledge. In particular, he shows that information aggregation is guaranteed only when voters' preference satisfies the Strong Preference Monotonicity (SPM), where SPM requires voters to respond in the same direction towards the change in likelihood of the state. Voters' preferences in the model of this study satisfy the SPM because the expected payoff difference increases as the belief put heavier weight towards the high state. However, this study shows that the information aggregation may fail even if voters' preferences satisfy SPM.

FP considered the voters whose preferences are heterogenous, and these preferences satisfy the SPM, like this study. In contrast to this study, FP shows in a fairly general setting

that information is aggregated asymptotically even when preferences are heterogenous. Thus, this study proposes the tractable model that produces the failure of information aggregation even when voters have FP-like preference heterogeneity, satisfying SPM.

The remainder of this paper is organized as follows. Section 2 describes the model, and Section 3 introduces the symmetric equilibrium. Section 4, 5, and 6 state the results, and Section 7 concludes. The appendices contain proofs omitted in texts.

2. The Model

A set of voters, $\{1, 2, \ldots, n\}$ ($n \geq 3$ odd), makes a collective decision $o \in \{A, B\}$ by a majority rule. Each voter simultaneously votes for one of two alternatives, A or B, where abstentions are not allowed and there is no cost of voting. The collective decision o will be A if A gets majority of votes and B if otherwise. There is an unknown state θ, which is uniformly distributed on $\Theta = (0,1)$, where Θ is partitioned into three subsets $\Theta_A = (0, \pi)$, $\Theta_N = (\pi, 1 - \pi)$, and $\Theta_B = (1 - \pi, 1)$, with $\pi \in (0, 1/2)$. The underlying state is unknown to anybody. Each voter has a binary preference type $t_i \in \{\mathcal{A}, \mathcal{B}\}$, and voters' payoff depends on the collective decision $o \in \{A, B\}$, the state $\theta \in \Theta$, and preference type $t_i \in \{\mathcal{A}, \mathcal{B}\}$. Let $u_{t_i}(o; \theta)$ be the payoff of voter, and $u_{t_i}(\theta) := u_{t_i}(B; \theta) - u_{t_i}(A; \theta)$ denote the payoff difference of a type t_i voter between alternative B and alternative A in state θ. To capture the partially conflicting interests, we suppose that

$$u_{\mathcal{B}}(\theta) = \theta - \pi,$$

and

$$u_{\mathcal{A}}(\theta) = \theta - (1 - \pi).$$

With this utility function, type $\mathcal{A}$ voters and type $\mathcal{B}$ voters have common interests when θ lies in $\Theta_B \cup \Theta_A$, while they have conflicting interests when θ lies in Θ_N. To see this, observe that

$$\begin{cases} u_{\mathcal{B}}(\theta), u_{\mathcal{A}}(\theta) < 0, & \text{if } \theta \in \Theta_A \\ u_{\mathcal{B}}(\theta), u_{\mathcal{A}}(\theta) > 0, & \text{if } \theta \in \Theta_B \end{cases},$$

and

$$u_{\mathcal{A}}(\theta) < 0 < u_{\mathcal{B}}(\theta), \quad \text{if } \theta \in \Theta_N.$$

The preference types are private information, independent across individuals and are identically distributed as $\Pr(t_i = \mathcal{B}) = 1/2$. Before voting, each voter receives conditionally independent, identically distributed private signal $s_i \in \{0, 1\}$, satisfying $\Pr(s_i = 1 | \theta) = \theta$. That is, conditional on the true state being θ, each voter gets 1-signal with probability θ and 0-signal with remaining probability $1 - \theta$.

3. Symmetric Strategy

A voting strategy of a voter, $\mu = \left(\mu_{B,1}, \mu_{B,0}, \mu_{\mathcal{A},1}, \mu_{\mathcal{A},0}\right)$, specifies a probability of voting for alternative B for each realization of preference type and signal $(t_i, s_i) \in \{\mathcal{A}, \mathcal{B}\} \times \{0,1\}$. A strategy profile is *symmetric* if (1) all voters use the same strategies and (2) the probability of following the own signal is the same between type $\mathcal{A}$ and type $\mathcal{B}$. The latter condition means that the probability of following the signal that conflict or matches with their type is the same between type $\mathcal{A}$ and type $\mathcal{B}$ voters. That is, we require $\mu_{\mathcal{A},1} = 1 - \mu_{B,0}$ and $\mu_{B,1} = 1 - \mu_{\mathcal{A},0}$.

When choosing which alternative to vote for, any voter must consider the event in which his/her vote is pivotal, i.e., exactly $(n-1)/2$ voters voted for B. Let $\tau(\theta)$ denote the probability that any voter votes for the alternative B in state θ. Then the probability that any voter becomes pivotal in state θ is given by

$$\mathbb{P}_n(piv|\theta) = \binom{n-1}{(n-1)/2} \left[\tau(\theta)(1 - \tau(\theta))\right]^{(n-1)/2}.$$

Consider a type $\mathcal{B}$ voter holding signal s. Given that other voters are following symmetric strategy profile μ, voting for B gives him/her higher expected payoff than voting for A if and only if $\mathbb{E}_n[u_B(\theta)|piv, s; \mu] \geq 0$, where $\mathbb{E}_n[u_B(\theta)|piv, s; \mu]$ denotes the expected payoff difference when there are n voters. Observe that

$$\mathbb{E}_n[u_B(\theta)|piv, s; \mu] \geq 0 \Leftrightarrow \int_0^1 \theta f(\theta|s, piv; \mu)\, d\theta \geq \pi \Leftrightarrow \mathbb{E}_n[\theta|s, piv; \mu] \geq \pi,$$

where $f(\theta|s, piv; \mu)$ is a posterior density of θ conditional on being pivotal and receiving signal s and $\mathbb{E}_n[\theta|s, piv; \mu]$ denotes a posterior mean of the state. Similarly, voting for B gives higher expected payoff to a type $\mathcal{A}$ voter with signal s if and only if

$$\int_0^1 \theta f(\theta|s, piv; \mu)\, d\theta \geq 1 - \pi \Leftrightarrow \mathbb{E}_n[\theta|piv, s; \mu] \geq 1 - \pi.$$

Note that posterior distribution of the state, conditional only on the signal s, is again Beta distribution with parameters $1 + \mathbb{I}_{s=1}$ and $1 + \mathbb{I}_{s=0}$, where $\mathbb{I}_{s=1}$ and $\mathbb{I}_{s=0}$ are indicator functions taking value of one if $s = 1$ and $s = 0$, respectively, and zero if otherwise. Therefore, the posterior mean of the state conditional on being pivotal and receiving signal s is given by

$$
\mathbb{E}_n[\theta|piv,s;\mu] =
\begin{cases}
\dfrac{\int_0^1 \theta^2 \big[\tau(\theta)(1-\tau(\theta))\big]^{(n-1)/2}\,d\theta}{\int_0^1 \theta\big[\tau(\theta)(1-\tau(\theta))\big]^{(n-1)/2}\,d\theta}, & \text{if } s = 1 \\[2em]
\dfrac{\int_0^1 \theta(1-\theta)\big[\tau(\theta)(1-\tau(\theta))\big]^{(n-1)/2}\,d\theta}{\int_0^1 (1-\theta)\big[\tau(\theta)(1-\tau(\theta))\big]^{(n-1)/2}\,d\theta}, & \text{if } s = 0
\end{cases}.
$$

LEMMA 1:
In any symmetric voting equilibrium μ^*, we have $\mu_{\mathcal{B},1}^* = 1 - \mu_{\mathcal{A},0}^* = 1$, or equivalently, probability of following the signal that matches with their types is 1.

PROOF:
Case 1: Probability of following the signal that conflicts with the type is positive.
Suppose we have $\mu_{\mathcal{A},1}^* = 1 - \mu_{\mathcal{B},0}^* > 0$ in the symmetric equilibrium. That is, type $\mathcal{A}$ voter with 1-signal follows their signal (i.e., voting for B) with positive probability and type $\mathcal{B}$ voter with 0-signal follows their signal (i.e., voting for A) with the same positive probability. Then, it follows that

$$
1 - \pi \le \int_0^1 \theta f(\theta|s = 1, piv; \mu^*)\,d\theta,
$$

and

$$
\pi \ge \int_0^1 \theta f(\theta|s = 0, piv; \mu^*)\,d\theta,
$$

implying that

$$
\pi < \int_0^1 \theta f(\theta|s = 1, piv; \mu^*)\,d\theta,
$$

and

$$
1 - \pi > \int_0^1 \theta f(\theta|s = 0, piv; \mu^*)\,d\theta,
$$

thus, proving the Lemma 1.

Case 2: Probability of following the signal that conflicts with the type is zero.
Suppose instead we have $\mu_{\mathcal{A},1}^* = 1 - \mu_{\mathcal{B},0}^* = 0$, so that type $\mathcal{A}$ voter with 1-signal follows their signal (i.e., voting for B) with probability zero and type $\mathcal{B}$ voter with 0-signal follows their signal (i.e., voting for A) with probability zero. Then, it follows that

$$1 - \pi \geq \int_0^1 \theta f(\theta|s = 1, piv; \mu^*) \, d\theta,$$

and

$$\pi \leq \int_0^1 \theta f(\theta|s = 0, piv; \mu^*) \, d\theta.$$

Thus, it suffices to prove the following:

$$\int_0^1 \theta f(\theta|s = 1, piv; \mu^*) \, d\theta > \int_0^1 \theta f(\theta|s = 0, piv; \mu^*) \, d\theta.$$

Note that the family of conditional densities of the signal $\{f(s|\theta)\}_\theta$ satisfies the strict Monotone Likelihood Ratio Property (Milgrom, 1981, p. 383). Given that the prior distribution of θ is now $f(\theta|piv; \mu^*)$, posterior distribution resulting from $s = 1$ first order stochastically dominates the posterior distribution resulting from $s = 0$ (Proposition 2 in Milgrom, 1981, p.383), thus proving the Lemma 1. ∎

Due to Lemma 1, we can simply denote by $\mu^* = \mu_{\mathcal{A},1}^* = 1 - \mu_{\mathcal{B},0}^*$ the symmetric voting equilibrium, which is the probability of following the conflicting signals.

4. Equilibrium Analysis: Pure Strategies

In an equilibrium with $\mu^* = 1$, all voters vote according to their signals, i.e., vote for A if they receive signal $s = 0$ and for B otherwise. In an equilibrium with $\mu^* = 1$, all voters vote according to their preference type, i.e., vote for A if they are type $\mathcal{A}$ and for B otherwise. I refer to the former equilibrium as an informative voting equilibrium, and the latter one as a partisan voting equilibrium.

PROPOSITION 1:
Fix any π. Informative voting profile cannot be an equilibrium for every sufficiently large n.

PROOF:
Given that other voters are following the informative voting profile, any voter can infer from a pivotal event that there are exactly $(n - 1)/2$ 1-signals among $n - 1$ voters. Therefore, a type $\mathcal{A}$ voter who privately observes 1-signal knows in the pivotal event that the number of 1-signals among n voters are now $(n - 1)/2 + 1$. In order for the informative voting profile to be an equilibrium, following the signal $s = 1$ must be beneficial for this type $\mathcal{A}$ voter. Thus, the following must be satisfied:

$$\int_0^1 \theta f(\theta|s = 1, piv; \mu^*) \, d\theta \geq 1 - \pi,$$

which is equivalent to

$$\underbrace{\frac{\frac{n-1}{2}+2}{n+2}}_{\to 1/2 \ (n\to\infty)} \geq 1-\pi.$$

Notice that the left-hand side this inequality is a mean of Beta distribution with parameters $(n-1)/2+2$ and $(n-1)/2+1$. However, since this converges to $1/2$, this inequality cannot be true for every sufficiently large n, proving the proposition. ∎

Proposition 1 says that the informative voting cannot be an equilibrium even if $\pi = 1/2 - \varepsilon$ so that the prior probability that voters disagree on the better alternative is arbitrarily small but positive. Note that informative voting becomes equilibrium for any n when $\pi = 1/2$ so that the voters never disagree (Austen-Smith & Banks, 1996).

Next proposition provides the necessary and sufficient condition for the existence of the partisan equilibrium.

PROPOSTION 2:
Fix any n. Partisan voting profile becomes an equilibrium if and only if $\pi \leq 1/3$.

PROOF:
When other voters are following partisan voting profile, a pivotal event conveys no information about the underlying state. Thus, the only information that any voter can rely on is his/her private signal. In order for the partisan voting profile to be an equilibrium, voting for A must be a best response for type $\mathcal{A}$ voter holding 1-signal:

$$\int_0^1 \theta f(\theta|s=1)\,d\theta \leq 1-\pi \iff \pi \leq 1/3,$$

thus, proving the Proposition 2 ∎

In partisan voting equilibrium no information is conveyed by being pivotal, and therefore voters must rely only on their private information. The partisan voting, however, tells them to ignore it, which can be beneficial only when the size of the disagreement is large enough.

5. Equilibrium Analysis: Mixed Strategies

This section investigates the existence of the symmetric equilibrium in mixed strategy and its asymptotic properties. To highlight a dependence on n, we let μ_n^* denote the symmetric equilibrium when there are n voters. The following theorem provides the necessary and sufficient condition for the existence of mixed strategy symmetric equilibrium.

THEOREM 1:

(1) If $\pi \leq 1/3$ then for any n there does not exist mixed strategy symmetric voting equilibrium.

(2) If $1/3 < \pi$, then there exists a symmetric voting equilibrium μ_n^* with $\mu_n^* \in (0,1)$ for every sufficiently large n. Furthermore, such symmetric equilibrium in mixed strategy, if it exists, is unique.

PROOF:
See Appendix A.

Theorem 1 asserts that the symmetric voting equilibrium which is neither the partisan voting nor the informative voting exists only when $1/3 < \pi$ so that the prior probability of disagreement is small. In contrast, if $\pi \leq 1/3$ so that such probability is large enough, then the only symmetric equilibrium is the partisan equilibrium.

Next theorem asserts that any symmetric equilibrium, including mixed strategy one, eventually converges to the partisan voting equilibrium.

THEOREM 2:

Suppose $1/3 < \pi$ so that, for every sufficiently large n, the unique symmetric voting equilibrium μ_n^* is in mixed strategy, $\mu_n^* \in (0,1)$. Then, the sequence of such equilibrium converges to the Partisan voting equilibrium. In other words,

$$\lim_{n \to \infty} \mu_n^* = 0.$$

PROOF:
See Appendix B.

Theorem 2, together with the first part of the Theorem 1, imply that in large elections voting behavior is determined solely by voters' preferences rather than information, which is formally stated in the next corollary. Note that this is true for any π.

COROLLARY 1:
Any sequence of symmetric voting equilibrium converges to the partisan voting equilibrium.

6. Asymptotic Condorcet Jury Theorem

Now we are ready to examine the validity of the Condorcet Jury Theorem. Let us begin by formal definition of a probability that the society reaches the correct decision under majority voting. Recall that in our context the correct alternative exists only when the true state θ lies in Θ_A or Θ_B.

DEFINITION 1:

A sequence of symmetric voting equilibrium $(\mu_n^*)_n$ asymptotically identifies the correct alternatives if the following two properties are satisfied:

1. *Fix any $\theta \in \Theta_B = (1 - \pi, 1)$. The probability of an alternative B being chosen, conditional on the true state being θ, converges to 1 as $n \to \infty$.*
2. *Fix any $\theta \in \Theta_A = (1 - \pi, 1)$. The probability of an alternative A being chosen, conditional on the true state being θ, converges to 1 as $n \to \infty$.*

Next theorem asserts that CJT is not true if the size of disagreement is large enough.

THEOREM 3:

If $\pi \leq 1/3$, then there is no sequence of symmetric equilibrium that asymptotically identifies the correct alternatives.

PROOF:

Theorem 3 immediately follows from the Proposition 1, Proposition 2, and Theorem 1.　　　■

7. Conclusion

Most of the game-theoretic literature on Condorcet voting presumes that the voters are engaging in the world of *"guilty or innocent"* (Austen-Smith & Banks, 1996; Gerardi, 2000; Wit 2000). Although voters are unaware of which alternative is correct, they know that there *always* is the correct alternative, and hence they can reach an agreement once the uncertainty is resolved.

In contrast, this study considered a binary election in which voters have partially conflicting interests (Schulte, 2012; Odora, 2023). Specifically, there is a correct alternative for all voters is some states, so the collective decisions are matter of truth, such as guilty or innocent. In other states, however, voters disagree on the better alternative and therefore the elections are matter of taste. This study investigated whether a group of voters using the majority rule can reach the *correct* decisions asymptotically when the size of the group tends to infinity, which is the classical statement known as Condorcet Jury Theorem (CJT). This study shows that if the prior probability that voters disagree is large enough, then the information aggregation through voting does not work, i.e., CJT does not hold. Moreover, this study shows that the voting behavior does not reflect signals and it is solely determined by the preference type in the large elections, as long as the prior probability of disagreement is positive.

A question unanswered yet is whether a pre-voting communication helps identify the correct alternatives when the voters have partially conflicting interests. Schulte (2012) examined a deliberative committee with partially conflicting interests, but her primary interest lies in the mechanism of the deliberation in small committees, and therefore she did not consider the asymptotic efficiency. Odora (2023), who extend our model to incorporate the deliberation stage, shows that it is generically impossible for all voters to tell their signals truthfully in the

deliberation stage, but the validity of the CJT is still remain unanswered.

Declaration of Interest
None.

Appendix A: Proof of Theorem 1

The symmetric strategy $\mu \in (0,1)$ is an equilibrium if and only if any voter whose signal conflicts with their type is indifferent between the two alternatives, conditional on being pivotal. In other words, both type $\mathcal{A}$ voter with 1-signal and type $\mathcal{B}$ voter with 0-signal are indifferent between voting for A and voting for B. That is, $\mu \in (0,1)$ is an equilibrium if and only if the followings are satisfied:

$$\mathbb{E}_n[\theta|s = 1, piv; \mu] = 1 - \pi.$$

$$\mathbb{E}_n[\theta|s = 0, piv; \mu] = \pi.$$

We define two functions of symmetric strategy $\mu \in [0,1]$, $\Gamma_1^n(\mu)$ and $\Gamma_0^n(\mu)$ by $\Gamma_1^n(\mu) = \mathbb{E}_n[\theta|s = 1, piv; \mu]$ and $\Gamma_0^n(\mu) = \mathbb{E}_n[\theta|s = 0, piv; \mu]$, respectively. Then proving or disproving the existence of symmetric equilibrium amounts to showing the existence of a solution μ^*, satisfying both $\Gamma_1^n(\mu^*) = 1 - \pi$ and $\Gamma_0^n(\mu^*) = \pi$ at the same time.

First part of the Theorem 1 (i.e., necessity) immediately follows from the fact that $\Gamma_1^n(\mu)$ is strictly decreasing in μ and the observation that $\Gamma_1^n(0) = 2/3$.

Next, we shall prove the second part, i.e., sufficiency. Suppose $1/3 < \pi$. We start from the following lemma, showing the symmetry between type $\mathcal{A}$ voter holding 1-signal and type $\mathcal{B}$ voter holding 0-signal.

LEMMA A.1:
For any n and μ, we have $\Gamma_1^n(\mu) + \Gamma_0^n(\mu) = 1$.

PROOF:
By definition, we observe that

$$\mathbb{E}_n[\theta|s = 1, piv; \mu] = \frac{\int_0^1 \theta^2 \mathbb{P}_n(piv|\theta; \mu)\, d\theta}{\int_0^1 \theta \mathbb{P}_n(piv|\theta; \mu)\, d\theta},$$

and

$$\mathbb{E}_n[\theta|s = 0, piv; \mu]$$

$$= \frac{1}{\int_0^1 (1-\theta)\mathbb{P}_n(piv|\theta;\mu)\,d\theta}\left\{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\hat\theta - \int_0^1 \theta^2\mathbb{P}_n(piv|\theta;\mu)\,d\theta\right\}.$$

Furthermore, posterior mean of the state, conditional only on being pivotal, $\mathbb{E}_n[\theta|piv;\mu]$, is given by

$$\mathbb{E}_n[\theta|piv;\mu] = \frac{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta}{\int_0^1 \mathbb{P}_n(piv|\theta;\mu)\,d\theta} = \frac{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta}{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta + \int_0^1 (1-\theta)\mathbb{P}_n(piv|\theta;\mu)\,d\theta}.$$

Therefore, it follows that

$$\mathbb{E}_n[\theta|piv;\mu] = \frac{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta}{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta + \int_0^1 (1-\theta)\mathbb{P}_n(piv|\theta;\mu)\,d\theta}\mathbb{E}_n[\theta|s=1,piv;\mu]$$

$$+ \frac{\int_0^1 (1-\theta)\mathbb{P}_n(piv|\theta;\mu)\,d\theta}{\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta + \int_0^1 (1-\theta)\mathbb{P}_n(piv|\theta;\mu)\,d\theta}\mathbb{E}_n[\theta|s=0,piv;\mu] \qquad (A.1)$$

Now, observe that

$$\mathbb{P}_n(piv|\theta;\mu) = \binom{n-1}{(n-1)/2}\left[\tau(\theta)(1-\tau(\theta))\right]^{(n-1)/2},$$

where

$$\tau(\theta) = \frac{1}{2} + \frac{\mu}{2}(2\theta - 1),$$

implying the following identity

$$\int_0^1 \theta\mathbb{P}_n(piv|\theta;\mu)\,d\theta = \int_0^1 (1-\theta)\mathbb{P}_n(piv|1-\theta;\mu)\,d\theta = \int_0^1 (1-\theta)\mathbb{P}_n(piv|\theta;\mu)\,d\theta.$$

Thus, combining this identity with eq. (A.1) yields $\mathbb{E}_n[\theta|piv;\mu] = 1/2$, and therefore eq. (A.1) can be rewritten as

$$\frac{1}{2} = \frac{1}{2}\mathbb{E}_n[\theta|s=1,piv;\mu] + \frac{1}{2}\mathbb{E}_n[\theta|s=0,piv;\mu],$$

which completes the proof of Lemma A.1. ∎

The Lemma A.1 allows us to focus on the equation $\Gamma_1^n(\mu^*) = 1 - \pi$. We use the Intermediate Value Theorem to show the existence of the solution $\mu^* \in (0,1)$ to this equation. To

do this, we first establish that $\Gamma_1^n(\cdot)$ is continuous on the interval $[0,1]$. To see this, for arbitrary $\mu \in [0,1]$ and arbitrary sequence $(\mu^k)_k$ in $[0,1]$ with $\mu^k \to \mu$, we observe that

$$\lim_{k\to\infty} \int_0^1 \theta^2 \mathbb{P}_n(piv|\theta;\mu^k)\, d\theta = \int_0^1 \left(\lim_{k\to\infty} \theta^2 \mathbb{P}_n(piv|\theta;\mu^k)\right) d\theta = \int_0^1 \theta^2 \mathbb{P}_n(piv|\theta;\mu)\, d\theta,$$

and

$$\lim_{k\to\infty} \int_0^1 \theta \mathbb{P}_n(piv|\theta;\mu^k)\, d\theta = \int_0^1 \left(\lim_{k\to\infty} \theta \mathbb{P}_n(piv|\theta;\mu^k)\right) d\theta = \int_0^1 \theta \mathbb{P}_n(piv|\theta;\mu)\, d\theta,$$

where we used the Arzela's Bounded Convergence Theorem to exchange the integration and the pointwise convergence of the sequence of the functions (see, e.g., Silva, 2010). Thus, we have

$$\lim_{k\to\infty} \Gamma_1^n(\mu^k) = \lim_{k\to\infty} \frac{\int_0^1 \theta^2 \mathbb{P}_n(piv|\theta;\mu^k)\, d\theta}{\int_0^1 \theta \mathbb{P}_n(piv|\theta;\mu^k)\, d\theta} = \frac{\int_0^1 \theta^2 \mathbb{P}_n(piv|\theta;\mu)\, d\theta}{\int_0^1 \theta \mathbb{P}_n(piv|\theta;\mu)\, d\theta} = \Gamma_1^n(\mu),$$

which proves the continuity of $\Gamma_1^n(\cdot)$.

Furthermore, we have $\Gamma_1^n(0) = 2/3$ and $\Gamma_1^n(1) = (n+3)/(2(n+2))$, where $\Gamma_1^n(1)$ converges to $1/2$ as $n \to \infty$, implying that we have $\Gamma_1^n(0) < 1 - \pi < \Gamma_1^n(1)$ for every sufficiently large n. Thus, it follows from the Intermediate Value Theorem that there exists a solution $\mu_n^* \in (0,1)$ such that $\Gamma_1^n(\mu_n^*) = 1 - \pi$ for every sufficiently large n. ∎

Appendix B: Proof of Theorem 2

We will prove Theorem 2 by contradiction; thus, assume that the sequence of symmetric voting equilibrium $(\mu_n^*)_n$ does not converge to 0. Then, there exists $\varepsilon > 0$ and a subsequence $(\mu_m^*)_m$ such that $\mu_m^* \geq \varepsilon > 0$ for every m. We define the function of symmetric strategy, $\Gamma_1^n(\mu)$, by $\Gamma_1^n(\mu) = \mathbb{E}_n[\theta|s = 1, piv; \mu]$. For this ε, we have the following Lemma B.1

LEMMA B.1:

$$\lim_{n\to\infty} \Gamma_1^n(\varepsilon) = \frac{1}{2}$$

Lemma B.1 implies that for sufficiently large m,

$$\Gamma_1^m(\varepsilon) < 1 - \pi = \Gamma_1^m(\mu_m^*).$$

However, since $\Gamma_1^m(\cdot)$ is strictly decreasing, we must have $\mu_m^* > \varepsilon$ whenever $\Gamma_1^m(\varepsilon) < \Gamma_1^m(\mu_m^*)$. Thus, we see that $\mu_m^* > \varepsilon$ for sufficiently large m, which contradicts the definition of the subsequence $(\mu_m^*)_m$.

What remains is to show the Lemma B.1.

Proof of Lemma B.1

To show that

$$\lim_{n \to \infty} \mathbb{E}_n[\theta | s = 1, piv; \varepsilon] = \frac{1}{2},$$

we use the Laplace Approximation Methods, which is an asymptotic approximation for the posterior mean (Tierney & Kadane, 1986; Bernardo & Smith, 1994; Robert, 2007). In general, posterior mean given a vector of observation $x = (x_1, .., x_n)$, $\mathbb{E}(\theta | x)$, can be written as

$$\mathbb{E}(\theta | x) = \frac{\int_0^1 b_N(\theta) \exp\{-n \cdot h_N(\theta)\} \, d\theta}{\int_0^1 b_D(\theta) \exp\{-n \cdot h_D(\theta)\} \, d\theta}.$$

In our context, our posterior mean $\mathbb{E}_n[\theta | s = 1, piv; \varepsilon]$ can be written in this manner by letting

$$-n \cdot h_N(\theta) = \log \theta^2 \big[\tau(\theta)\big(1 - \tau(\theta)\big)\big]^{(n-1)/2},$$

$$-n \cdot h_D(\theta) = \log \theta \big[\tau(\theta)\big(1 - \tau(\theta)\big)\big]^{(n-1)/2},$$

and $b_N(\theta) = b_D(\theta) = constant$, where

$$\tau(\theta) = \frac{1}{2} + \frac{\varepsilon}{2}(2\theta - 1)$$

is the probability that any voter vote for alternative B in state θ when voters follow symmetric strategy ε.

In general, in the words of Tierney and Kadane (1986), the posterior mean, which can be expressed in the form of

$$\frac{\int_0^1 b_N(\theta) \exp\{-n \cdot h_N(\theta)\} \, d\theta}{\int_0^1 b_D(\theta) \exp\{-n \cdot h_D(\theta)\} \, d\theta},$$

is said to be *written in fully exponential form* if $b_N(\theta) = b_D(\theta)$. We immediately see that our posterior mean, $\mathbb{E}_m[\theta | s = 1, piv; \varepsilon]$, is written in fully exponential form.

Next, define θ_N and θ_D by

$$-h_N(\theta_N) = \sup_\theta\{-h_N(\theta)\},$$

$$-h_D(\theta_D) = \sup_\theta\{-h_D(\theta)\},$$

and define σ_N^2 and σ_D^2 such that

$$\sigma_N^2 = \left[h_N''(\theta)|_{\theta=\theta_N} \right]^{-1}.$$

$$\sigma_D^2 = \left[h_D''(\theta)|_{\theta=\theta_D} \right]^{-1}.$$

Then Tierney and Kadane (1986)[1] have shown that if the posterior mean, $\mathbb{E}(\theta|\boldsymbol{x})$, is written in fully exponential form,

$$\mathbb{E}(\theta|\boldsymbol{x}) = \frac{b_N(\theta_N)}{b_D(\theta_D)} \cdot \frac{\sigma_N^2}{\sigma_D^2} \cdot \exp\{-n\big(h_N(\theta_N) - h_D(\theta_D)\big)\} + O(n^{-2}).$$

In our context,

$$\mathbb{E}_n[\theta | s = 1, piv; \varepsilon] = \frac{\sigma_N^2}{\sigma_D^2} \cdot \exp\{-n\big(h_N(\theta_N) - h_D(\theta_D)\big)\} + O(n^{-2}).$$

Next lemma concludes the proof for Theorem 2.

LEMMA B.2:

$$\frac{\sigma_N^2}{\sigma_D^2} \cdot \exp\{-n\big(h_N(\theta_N) - h_D(\theta_D)\big)\} \to \frac{1}{2} \ (n \to \infty).$$

PROOF:
In our context, $-h_N(\theta)$ and $-h_D(\theta)$ are defined by

$$-h_N(\theta) = \frac{\log \theta^2 \big[\tau(\theta)\big(1 - \tau(\theta)\big)\big]^{(n-1)/2}}{n},$$

---

[1] See also Bernardo and Smith (1994) Chapter 5.5.1, and Robert (2007) Chapter 6.2.3.

and

$$-h_D(\theta) = \frac{\log\theta\big[\tau(\theta)(1-\tau(\theta))\big]^{(n-1)/2}}{n},$$

implying that θ_N and θ_D are given by

$$\theta_N = \frac{\varepsilon^2(n+3) + \{8(n+1)\varepsilon^2 + \varepsilon^4(n-1)^2\}^{1/2}}{4(n+1)\varepsilon^2}, \qquad (B.1)$$

and

$$\theta_D = \frac{\varepsilon^2(n+1) + \{4n\varepsilon^2 + \varepsilon^4(n-1)^2\}^{1/2}}{4n\varepsilon^2}. \qquad (B.2)$$

Notice that both θ_N and θ_D converge to $1/2$ as $n \to \infty$, implying that both $\tau(\theta_N)$ and $\tau(\theta_D)$ converges to $1/2$.

We first establish that σ_N^2/σ_D^2 converges to 1 as $n \to \infty$. To see this, observe that

$$\frac{\sigma_N^2}{\sigma_D^2} = \frac{[h_N''(\theta_N)]^{-1}}{[h_D''(\theta_D)]^{-1}} = \frac{h_D''(\theta_D)}{h_N''(\theta_N)} = \frac{\dfrac{1}{\theta_D^2} + \dfrac{n-1}{2}\left(\dfrac{\varepsilon}{\tau(\theta_D)}\right)^2 + \dfrac{n-1}{2}\left(\dfrac{\varepsilon}{1-\tau(\theta_D)}\right)^2}{\dfrac{1}{\theta_N^2} + \dfrac{n-1}{2}\left(\dfrac{\varepsilon}{\tau(\theta_N)}\right)^2 + \dfrac{n-1}{2}\left(\dfrac{\varepsilon}{1-\tau(\theta_N)}\right)^2}$$

$$= \frac{\dfrac{2}{n-1}\cdot\dfrac{1}{\theta_D^2} + \left(\dfrac{\varepsilon}{\tau(\theta_D)}\right)^2 + \left(\dfrac{\varepsilon}{1-\tau(\theta_D)}\right)^2}{\dfrac{2}{n-1}\cdot\dfrac{1}{\theta_N^2} + \left(\dfrac{\varepsilon}{\tau(\theta_N)}\right)^2 + \left(\dfrac{\varepsilon}{1-\tau(\theta_N)}\right)^2}$$

$$\to 1 \ (n \to \infty),$$

where convergence to 1 follows from the observation that both $\tau(\theta_N)$ and $\tau(\theta_D)$ converges to $1/2$.

Next, we shall prove that $\exp\{-n(h_N(\theta_N) - h_D(\theta_D))\}$ converges to $1/2$. First, observe that

$$\exp\{-n(h_N(\theta_N) - h_D(\theta_D))\}$$

$$= \exp\left\{\log\left\{\theta_N^2\big[\tau(\theta_N)(1-\tau(\theta_N))\big]^{(n-1)/2}\right\} - \log\left\{\theta_D\big[\tau(\theta_D)(1-\tau(\theta_D))\big]^{(n-1)/2}\right\}\right\}$$

$$= \exp\left\{\log\left\{\theta_N \cdot \frac{\theta_N}{\theta_D} \cdot \left[\frac{\tau(\theta_N)(1 - \tau(\theta_N))}{\tau(\theta_D)(1 - \tau(\theta_D))}\right]^{(n-1)/2}\right\}\right\}$$

$$= \theta_N \cdot \frac{\theta_N}{\theta_D} \cdot \left[\frac{\tau(\theta_N)(1 - \tau(\theta_N))}{\tau(\theta_D)(1 - \tau(\theta_D))}\right]^{(n-1)/2}$$

$$= \theta_N \cdot \frac{\theta_N}{\theta_D} \cdot \left[\frac{\theta_N(1 - \theta_N)}{\theta_D(1 - \theta_D)}\right]^{(n-1)/2},$$

where the last equality follows from the observations that $\theta_N/\theta_D = \tau(\theta_N)/\tau(\theta_D)$ and $(1 - \theta_N)/(1 - \theta_D) = (1 - \tau(\theta_N))/(1 - \tau(\theta_D))$. Since we know that both θ_N and θ_D converge to $1/2$, it suffices to show that

$$\lim_{n \to \infty} \left[\frac{\theta_N(1 - \theta_N)}{\theta_D(1 - \theta_D)}\right]^{(n-1)/2} = 1.$$

To highlight the dependence on n, we denote $a_n = \theta_N$ and $b_n = \theta_D$, where θ_N and θ_D are given in the eq. (B.1) and eq. (B.2), respectively. Then, we must show that

$$\lim_{n \to \infty} \left[\frac{a_n(1 - a_n)}{b_n(1 - b_n)}\right]^{(n-1)/2} = 1.$$

It follows from the continuity of the exponential functions that this is equivalent to

$$\lim_{n \to \infty} \log\left[\frac{a_n(1 - a_n)}{b_n(1 - b_n)}\right]^{(n-1)/2} = \log 1,$$

so we will prove this. The idea is to use the definition of the derivative of some function $\log\phi(x)$ at $x = 0$. We see that reparametrizing to $t = 2/(n - 1)$ yields

$$\log\left[\frac{a_n(1 - a_n)}{b_n(1 - b_n)}\right]^{(n-1)/2} = \frac{\log\left[\frac{a_n(1 - a_n)}{b_n(1 - b_n)}\right]}{2/(n - 1)}$$

$$= \frac{\log\left[\frac{a_{1+2/t}(1 - a_{1+2/t})}{b_{1+2/t}(1 - b_{1+2/t})}\right]}{t},$$

and we observe that the numerator of the last expression can be written as

$$\log\left[\frac{a_{1+2/t}\left(1-a_{1+2/t}\right)}{b_{1+2/t}\left(1-b_{1+2/t}\right)}\right]$$

$$= \log\left\{\left(\frac{t+2}{1-t}\right)^2 \cdot \frac{\varepsilon^2\left(1+2t(t+2)\right)-2t(t+1)+\sqrt{4t(t+1)\varepsilon^2+\varepsilon^4}}{\varepsilon^2\left(2+t(t+4)\right)-t(t+2)+2\sqrt{t(t+2)\varepsilon^2+\varepsilon^4}}\right\} - \log 2 . \quad (B.3)$$

If we set

$$\phi(x) = \left(\frac{x+2}{1-x}\right)^2 \cdot \frac{\varepsilon^2\left(1+2x(x+2)\right)-2x(x+1)+\sqrt{4x(x+1)\varepsilon^2+\varepsilon^4}}{\varepsilon^2\left(2+x(x+4)\right)-x(x+2)+2\sqrt{x(x+2)\varepsilon^2+\varepsilon^4}} ,$$

then the right-hand side of eq. (B.3) can be rewritten as $\log\phi(t+0) - \log\phi(0)$, thus we have

$$\log\left[\frac{a_n(1-a_n)}{b_n(1-b_n)}\right]^{(n-1)/2} = \frac{\log\left[\frac{a_{1+2/t}\left(1-a_{1+2/t}\right)}{b_{1+2/t}\left(1-b_{1+2/t}\right)}\right]}{t} = \frac{\log\phi(t+0)-\log\phi(0)}{t}.$$

Therefore, we see that

$$\lim_{n\to\infty}\log\left[\frac{a_n(1-a_n)}{b_n(1-b_n)}\right]^{(n-1)/2}$$

$$= \lim_{t\to 0}\frac{\log\left[\frac{a_{1+2/t}\left(1-a_{1+2/t}\right)}{b_{1+2/t}\left(1-b_{1+2/t}\right)}\right]}{t}$$

$$= \lim_{t\to 0}\frac{\log\phi(t+0)-\log\phi(0)}{t} = \frac{d}{dx}\left[\log\phi(x)\right]\bigg|_{x=0} .$$

Finally, tedious calculation yields that the value of the derivative of $\log\phi(x)$ at $x=0$ equals to 0, which is $\log 1$:

$$\frac{d}{dx}\left[\log\phi(x)\right]\bigg|_{x=0} = 0 = \log 1 ,$$

which proves the Lemma B.2. ∎

References

Austen-Smith, D., and Banks, J. S. (1996). Information Aggregation, Rationality, and the Condorcet Jury Theorem. *American Political Science Review* 90(1), 34-45.

Berend, D, and Paroush, P. (1998). When Is Condorcet's Jury Theorem Valid? *Social Choice and Welfare* 15 (4): 481–88.

Bernardo, J., M., and Smith, A., F., M. (1994). *Bayesian Theory*. Wiley, New York.

Bhattacharya, S. (2013). Preference Monotonicity and Information Aggregation in Elections. *Econometrica* 81 (3): 1229–47.

Feddersen, T., and Pesendorfer, W. (1997). Voting Behavior and Information Aggregation in Elections with Private Information. *Econometrica* 65(5), 1029-1058.

Gerardi, D. (2000). Jury Verdicts and Preference Diversity. *American Political Science Review 94*(2), 395-406.

McLennan, A. 1998. Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents. *American Political Science Review* 92 (2): 413–18.

Milgrom, P., R. (1981). Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics* 12 (2): 380–91.

Miller, N., R. (1986) Information, electorates, and democracy: Some extensions and interpretations of the Condorcet Jury Theorem. In: Grofman, B., Owen, G. (eds) Information Polling and Group Decision Making. Greenwich CT: JAI Press

Odora, M. (2023). Deliberation and Voting: A Matter of Truth or Taste. *mimeo.*

Robert, C., P. (2007). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation.* 2nd edition. Berlin: Springer.

Schulte, E. (2012). Communication in Committees: Who Should Listen? *Public Choice* 150 (1): 97–117.

Silva, N. (2010). A Concise, Elementary Proof of Arzelà's Bounded Convergence Theorem. *The American Mathematical Monthly: The Official Journal of the Mathematical Association of America* 117 (10): 918.

Tierney, L., and Kadane, J.B., (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association.* 81, 82–86.

Wit, J. (1998). Rational Choice and the Condorcet Jury Theorem. *Games and Economic Behavior*. 22(2), 364-376.