Elections with Opinion Polls: Information Acquisition and Aggregation*

Andrei Gomberg† Tetsuya Hoshino‡

March 11, 2025

Abstract

We study elections with opinion polls, which contain information about their likely margins. Rationally inattentive voters acquire information about both the alternatives and the polls, subject to entropy-based costs. We characterize the unique symmetric equilibrium with information acquisition. Our main results are: (i) the probability of making the correct choice is independent of electorate size, voters' prior, and the presence of partisans; (ii) elections become closer as electorate size grows. Furthermore, elections with polls can achieve a higher probability of making the correct choice than those without. We offer novel implications for regression discontinuity design in close elections.

Keywords: Rational inattention, rational ignorance, voting, opinion polls, regression discontinuity designs.

JEL codes: D72, D82, D83

*We would like to thank Yu Awaya, Gadi Barlevy, Ying Chen, Henrique de Oliveira, Mark Dean, John Duffy, John Duggan, Ruben Enikolopov, Chishio Furukawa, Fei Li, Srihari Govindan, Emilio Gutierrez, Chiaki Hara, Kosuke Imai, Edi Karni, Takuma Kamada, Hiroto Katsumata, Ali Khan, Margarita Kirneva, Horacio Larreguy, Shaun McRae, César Martinelli, Tetsuya Matsubayashi, Jaakko Meriläinen, Romans Pancs, Mauricio Romero, Yuliy Sannikov, Alberto Simpser, Katsutoshi Wakai, and Xinyang Wang for helpful comments. We are also grateful to the comments from the conference/seminar participants at ITAM, Kyoto University, São Paulo School of Economics, Johns Hopkins University, the 2023 Kyoto Autumn Workshop, the 2024 SAET Conference, the 2024 SSCW Meeting, the SING19 Conference, the 2024 SEA Conference, and the 2025 JSQPS Winter Meeting. All remaining errors are our own.

†ITAM. Email: gomberg@itam.mx

‡ITAM. Email: tetsuyahoshino137@gmail.com

1 Introduction

Existing studies present two contrasting perspectives on the ability of voting to aggregate information. Condorcet's (1785) Jury Theorem offers an optimistic view: when voters share common interests but possess dispersed information, majority voting can effectively select the commonly preferred—or "correct"—alternative. This optimism is challenged by Downs's (1957) hypothesis of **rational ignorance**, which posits that voters must incur costs to acquire information. In a large electorate, where an individual vote is unlikely to be pivotal, each voter may rationally choose to remain uninformed, leading to an under-provision of information and casting doubt on the ability of voting to aggregate information efficiently. This pessimism has spurred extensive research on voters' information acquisition and aggregation (e.g., Persico, 2004; Martinelli, 2006, 2007; Koriyama and Szentes, 2009; Oliveros, 2013; Triossi, 2013). The key trade-off highlighted in the literature is between the cost of acquiring information and the probability of being pivotal.

In this paper, we examine the possibility that information about pivot probability may be available to voters through public **opinion polls** (or simply **polls**). Polls may help voters more accurately assess whether it is worthwhile to learn about the alternatives. If a poll suggests a one-sided race, voters perceive a lower pivot probability and acquire less information; if it indicates a close race, they perceive a higher pivot probability and acquire more.

Our main question is how polls shape both the acquisition and the aggregation of information. Although the closeness of an election may affect voters' incentives to learn, that closeness is itself endogenously determined by their collective behavior. We seek to understand this equilibrium interaction and its implications for the quality of collective decision-making.

Model Consider an electorate of N voters, with N odd, who must choose between two alternatives, 1 and 0, by majority rule. Each voter must cast a vote for one of the alternatives, with no abstention allowed. There is a state θ that takes one of two values, 1 or 0, according to a prior μ. At state θ, alternative θ is the "correct" choice; each voter receives a payoff of 1 if the collective choice matches the state, and 0 otherwise.

Voters may acquire costly information about both the state and a poll, which indicates a likely vote share of the two alternatives. We model their information acquisition using the framework of **rational inattention**, in which a voter chooses any (noisy) signal that may correlate with the state and the poll, subject to a cost proportional to the expected reduction in uncertainty, measured by entropy (e.g., Sims, 2003; Matějka and McKay, 2015). This approach allows her to pay close attention to the state when a poll suggests a close race, but ignore it otherwise. Each voter maximizes her expected payoff from the election outcome, net of the (private) cost of information.[1]

Voting and polling are interdependent: voters respond to polling information, while the poll reflects the aggregate behavior of the electorate. An equilibrium thus requires the following triadic conditions: (i) voters' behavior is individually optimal under their beliefs; (ii) their beliefs are

[1]The poll itself contains no noise in aggregating votes, but voters will typically acquire noisy signals about it.

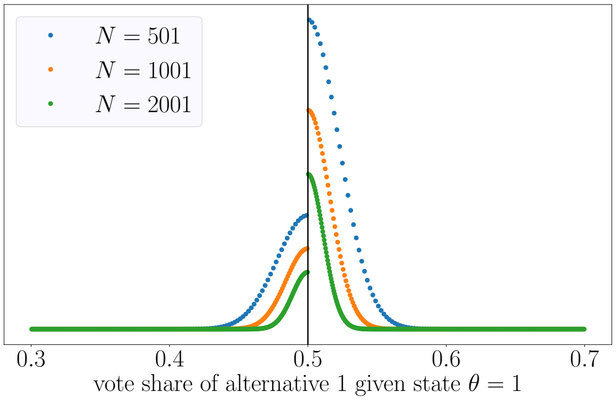

Figure 1: With opinion polls

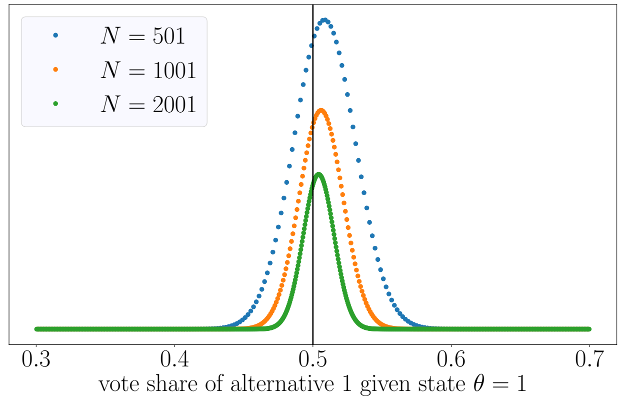

Figure 2: Without opinion polls

Note: The equilibrium vote-share distributions at state $\theta = 1$ under the prior $\mu(1) = 0.5$

consistent with the acquired information about the state and the poll; and (iii) the poll is consistent with actual voting behavior. Such an equilibrium is a fixed point of this interactive system.

This equilibrium concept, built on Denti's (2023) notion of robustness to information acquisition, has two interpretations in our setting. First, it parallels a rational-expectations equilibrium: voters acquire (noisy) endogenous information about collective behavior and respond accordingly (Grossman and Stiglitz, 1980; Hébert and La'O, 2023). Here, polls play a crucial role in aggregating voting intentions, allowing them to form rational expectations. Second, although our model is fully static, this equilibrium can be viewed as the long-run outcome of a sequential electoral process in which voters repeatedly revise their votes upon observing updated polls. (See Section 2.2 for details.) In this Introduction, we focus on **informative equilibria**, in which voters acquire information, leaving uninformative equilibria to later analysis. As we show, an informative equilibrium exists and is unique unless the prior μ is heavily biased toward one state.

Of particular interest is the equilibrium vote share. Figure 1 depicts the probability mass functions of the vote shares of alternative 1 given state $\theta = 1$ (at which alternative 1 is correct) under the prior $\mu(1) = 0.5$, for electorate sizes $N = 501, 1001, 2001$.[2] At state $\theta = 1$, any vote share above 0.5 yields the correct choice. Observe that each distribution "jumps" at 0.5, indicating that the election likely results in favor of the correct choice. This jump reflects our intuition that polls encourage information acquisition when the election is close.

Main Results Our main results—formally stated as Theorems 1 and 2—are as follows:

> In an election with an opinion poll, the equilibrium probability that majority vote chooses the correct alternative is independent of the electorate size N and the prior μ. Moreover, as N increases, the equilibrium vote share converges in probability to the winning threshold of 0.5 even if μ is biased toward one state.

[2]The heights of the probability mass functions decrease as N increases, since the total probability 1 is distributed across a larger number of possible vote share values.

Theorem 1 states that Condorcet's optimism (that larger electorates should choose more accurately) is perfectly canceled by Downs's pessimism (that larger electorates might choose less accurately). Furthermore, the probability of making the correct choice does not depend on μ. Theorem 2 shows that as N grows, the vote-share distribution concentrates at 0.5, as illustrated in Figure 1. This implies that large elections tend to be close. Notably, this result continues to hold even if μ is biased toward one state. Despite this closeness, the probability of choosing the correct alternative remains considerably above one half.

Theorems 1 and 2 hold *exactly* as stated even when there are **partisans** in the electorate. Suppose that N_1 voters always vote for alternative 1 and N_0 always vote for alternative 0, with $N_1, N_0 < N/2$. Then, the probability that majority vote chooses the correct alternative is independent of N_1 and N_0;[3] similarly, as N increases, the vote share converges to 0.5 for any $N_1, N_0 < N/2$. Furthermore, both theorems hold under **supermajority rule**, including **unanimity rule**.[4]

In contrast to settings with exogenous information, we do not find a swing voter's curse, where less-informed voters ignore their private signals to defer to (potentially) better-informed majorities (Feddersen and Pesendorfer, 1996). In exogenous-information models, being pivotal can itself be more informative than any private signal, causing voters to discount their own signals. In our endogenous-information model, voters strategically acquire whatever signals they anticipate using, knowing the likelihood of being pivotal and the presence of partisans. Consequently, no (non-partisan) voter would acquire any costly information that she then ignores in equilibrium.

Comparison between Elections with and without Opinion Polls Do polls help the electorate make more accurate choices? To address this question, consider an election without a poll, in which voters learn only about the state (as in earlier rational-ignorance studies). Figure 2 plots the probability mass function of the equilibrium vote share of alternative 1 in this setting. Unlike Figure 1, there is no jump at 0.5, though the distribution shifts favorably toward the correct alternative.

We show that, for a large enough electorate size N and a nearly symmetric prior $\mu(1) \approx 0.5$, the probability that majority vote chooses the correct alternative is strictly higher *with* a poll than *without* one (Proposition 1).[5] This result underscores the role of polling in both information acquisition and aggregation.

Regression Discontinuity Design in Close Elections Our findings offer novel implications for the **regression discontinuity (RD) design** in close elections. Researchers often identify treatment effects in electoral environments by comparing candidates who barely win against those who barely lose (e.g., Hahn, Todd, and Van der Klaauw, 2001). Although such designs are widely applied,

[3] If either group of partisans is at least half of the electorate, their preferred alternative trivially wins.

[4] These extensions are essentially the same. For example, among $N = 101$ voters, with $N_1 = 10$ partisans for alternative 1 and $N_0 = 20$ for alternative 0, the remaining 71 non-partisans face a 41-out-of-71 supermajority requirement to elect alternative 1. Likewise, if $N_1 = 0$ and $N_0 = 50$, the remaining 51 non-partisans effectively face a unanimity requirement.

[5] While this proposition formally applies only to nearly symmetric priors, we conjecture that it holds more generally.

their validity has been called into question. For example, Caughey and Sekhon (2011) show that winners and losers in close U.S. House elections differ significantly on pretreatment covariates (e.g., campaign finance and incumbency); Snyder (2005) finds that incumbents won a disproportionate share of close U.S. House elections. By contrast, Eggers, Fowler, Hainmueller, Hall, and Snyder (2015) do not detect such sorting in many close elections, yet confirm Snyder's findings for the U.S. House elections.

One possible cause of sorting is post-election manipulation of votes (Snyder, 2005); another is precisely measured campaign effort (Caughey and Sekhon, 2011). Both mechanisms hinge on having reliable information about how the vote is trending, which brings to mind polls. Indeed, Eggers et al. (2015) note that close U.S. House elections, where sorting is repeatedly documented, tend to be polled more often than other elections commonly used in RD designs.

We propose a novel mechanism for sorting in close elections that arises from individual voter decisions alone. As illustrated in Figures 1 and 2, the equilibrium vote shares have a jump at the 0.5 threshold if and only if there is a poll. This suggests that *even when alternatives are ex-ante identical, voters may coordinate, through polling, to choose the (ex-post) correct alternative by a narrow margin*. This polling-based mechanism complements existing explanations and could help reconcile the mixed evidence on sorting in close elections.

Related Literature Our study bridges two strands of the literature: rational ignorance and rational inattention. Downs's (1957) classical hypothesis of rational ignorance maintains that voters acquire information only when the expected benefits exceed the costs. This hypothesis helps explain why voters in a large electorate (with a small pivot probability) may rationally remain uninformed. By endogenizing information acquisition, it also foreshadows the approach of rational inattention.

Existing studies on rational ignorance adopt a "rigid" information structure, maintaining parametric assumptions about signals available to voters. For instance, Martinelli (2006) assumes that voters choose a particular level of signal precision about a state; he then shows that the diminishing incentives to acquire precise signals (due to small pivotal probabilities) can outweigh the value of having more signals in a larger electorate. Oliveros (2013) and Triossi (2013) introduce voter heterogeneity, while other related work focuses on a binary choice of whether to purchase a fixed-quality signal or not (e.g., Mukhopadhaya, 2003; Persico, 2004; Martinelli, 2007; Gerardi and Yariv, 2008; Koriyama and Szentes, 2009). Experiments also implement similar binary-choice information acquisition models (e.g., Bhattacharya, Duffy, and Kim, 2017; Elbittar, Gomberg, Martinelli, and Palfrey, 2020).

We are not the first to posit that voters may learn about their pivot probabilities. Ekmekci and Lauermann (2022) show that when voters receive information about electorate size, they may infer their pivotality and potentially generate inefficient information aggregation. Because their environment assumes exogenous information, it does not explore the endogenous information acquisition that we study. Taylor and Yildirim (2010) examine how polling may lead to closer elections when turnout is endogenous, but they do not consider endogenous information acquisition.

Another line of work examines cheap-talk communication before voting, focusing on how exogenous signals are shared. Feddersen and Pesendorfer (1998) and Austen-Smith and Banks (1996) show that under unanimity rule, voters may rationally choose to disregard their private signals, leading to inefficiency. Coughlan (2000) demonstrates that a single round of communication can enable full information sharing; Gerardi and Yariv (2007) study broader communication protocols, showing that many voting rules yield equivalent equilibrium outcomes after communication. In our setting, opinion polls induce a similar equivalence across voting rules, but with endogenous acquisition of information.[6]

The rational inattention approach has found applications in political economy, though it has not yet been used to examine the implications of the rational ignorance hypothesis. Matějka and Tabellini (2021) examine a spatial model of electoral competition with rationally inattentive voters. Assuming that voters are rationally inattentive, Yuksel (2022) explores political polarization, while Li and Hu (2023) investigate politicians' accountability.[7]

Layout The paper is organized as follows. Section 2 develops the model of elections with polls and presents our main results. Section 3 compares these outcomes to those in elections without polls, focusing on the probability of choosing the correct alternative. Section 4 then discusses the implications for regression discontinuity designs in close elections. Section 5 concludes by outlining extensions and limitations of our analysis. Proofs are found in Appendix A.

The replication code for all figures in this paper is available at this webpage.

2 Elections with Opinion Polls

In this section, we study elections with opinion polls. We begin by introducing an election model and an equilibrium concept, then characterize equilibria in finite-voter elections, and finally analyze the limit case as the number of voters tends to infinity.

2.1 Model

Base Environment There are $N = 2n + 1$ voters, denoted $i = 1, \ldots, N$, for an integer $n \geq 0$. There are two alternatives $a \in A = \{0, 1\}$, and each voter i votes for an alternative $a_i \in A$, which we call action a_i. No abstention is allowed. Let $a = (a_1, \ldots, a_N)$ be an action profile and a_{-i} be

[6] A number of experimental studies examine the effect of polls. Sinclair and Plott (2012) find that voter errors decline when polls inform people about others' intentions; Agranov, Goeree, Romero, and Yariv (2018) observe that availability of polling data shapes voters' perceptions of pivotality and thus their choices.

[7] These studies assume a continuum of voters. Although this approach may have certain advantages, it would not adequately capture the interplay between Condorcet-type information aggregation and the free-rider problem emphasized by rational ignorance. In a continuum-voter model, each voter's influence is infinitesimal, so pivotality (and thus fully endogenous information acquisition) never arises.

an action profile of all voters but i. Given a, the **vote share (of alternative 1)** is

$$\bar{a}_N = \frac{1}{N}\sum_{i=1}^{N} a_i,$$

while the vote share of alternative 0 is $1-\bar{a}_N$. Under **simple majority rule**, alternative 1 is chosen if and only if $\bar{a}_N > \frac{1}{2}$, and alternative 0 is chosen otherwise; thus, the chosen alternative is denoted by $\mathbb{1}\{\bar{a}_N > \frac{1}{2}\}$, where $\mathbb{1}$ is the indicator function. There is no tie, since N is odd.

There is a true state θ drawn from a state space $\Theta = \{0,1\}$ according to a prior $\mu \in \Delta(\Theta)$. At state θ, alternative θ is considered "correct." For now, we assume that all voters share the common interest of matching the chosen alternative $\mathbb{1}\{\bar{a}_N > \frac{1}{2}\}$ with state θ. Specifically, we define each voter's payoff function $u : [0,1]\times\Theta \to \{0,1\}$ by

$$u(\bar{a}_N, \theta) = \begin{cases} 1 & \text{if } \mathbb{1}\{\bar{a}_N > \frac{1}{2}\} = \theta \\ 0 & \text{if } \mathbb{1}\{\bar{a}_N > \frac{1}{2}\} \neq \theta. \end{cases}$$

In Section 2.5, we extend the model to incorporate supermajority rule (including unanimity rule) and partisan voters whose preferences over the alternatives are independent of state θ.

Information Acquisition Voters acquire information at a cost. Each voter can learn about two variables: the state θ and an **opinion poll** indicating how the remaining electorate is likely to vote. Specifically, voter i may learn about the vote share $\bar{a}_{-i} = \sum_{j\neq i} a_j/(N-1)$. We model information acquisition using the framework of rational inattention (Sims, 2003). In this approach, each voter can flexibly choose what to learn, not only how much to learn, allowing, for example, more attention to the state θ when an election is likely to be close and less attention otherwise.

Each voter i selects a signal structure consisting of a signal space S_i and a conditional distribution $\sigma_i(\cdot \mid \bar{a}_{-i}, \theta) \in \Delta(S_i)$ for each $(\bar{a}_{-i}, \theta)$. She then takes an action based on the realized signal. Without loss of generality, we can restrict ourselves to "direct" signal structures in which the signal space is $S_i = A$ and each realization a_i is interpreted as a recommendation to take action a_i.[8] Accordingly, voter i's strategy is represented by a system of conditional action distributions $P_i(\cdot \mid \bar{a}_{-i}, \theta) \in \Delta(A)$, one for each $(\bar{a}_{-i}, \theta)$.

Following much of the literature (e.g., Sims, 2003; Matějka and McKay, 2015), we assume the entropy-based cost of information.[9] It is proportional to the expected reduction in the voter's uncertainty measured by entropy. Suppose that voter i has a belief $\mu_i(\bar{a}_{-i}, \theta)$, which is endogenously determined in equilibrium. Given a strategy P_i, the expected reduction in entropy of $(\bar{a}_{-i}, \theta)$ is

$$\mathbb{I}(\bar{a}_{-i}, \theta; a_i) = \mathbb{H}(\bar{a}_{-i}, \theta) - \mathbb{H}(\bar{a}_{-i}, \theta \mid a_i),$$

[8] By a standard argument from the literature (e.g., Matějka and McKay, 2015), every pair of a signal structure and an action mapping can be replaced by a (weakly) cheaper direct signal structure that induces the same conditional action distribution. Also, voter i does not randomize over signal structures.

[9] See de Oliveira (2019) for axiomatic foundations.

where $\mathbb{H}$ denotes the entropy function. Here, $\mathbb{H}(\bar{a}_{-i}, \theta)$ is the entropy of $(\bar{a}_{-i}, \theta)$, where $(\bar{a}_{-i}, \theta)$ is distributed according to μ_i, and $\mathbb{H}(\bar{a}_{-i}, \theta \mid a_i)$ is the conditional entropy of $(\bar{a}_{-i}, \theta)$ given a_i, where $(a_i, \bar{a}_{-i}, \theta)$ is distributed according to P_i and μ_i.[10] The information cost of P_i under μ_i is $\lambda\mathbb{I}(\bar{a}_{-i}, \theta; a_i)$, where $\lambda > 0$ is the unit cost of information.

We denote by $\mathcal{P}_N$ the **election with a poll** (omitting the prior μ and the unit cost λ).

2.2 Equilibrium

Voter i's strategy P_i is called optimal under her belief μ_i if it maximizes her expected payoff net of information costs: $\mathbb{E}[u(\bar{a}_N, \theta)] - \lambda\mathbb{I}(\bar{a}_{-i}, \theta; a_i)$, where $\mathbb{E}[u(\bar{a}_N, \theta)]$ is the expected payoff with respect to the distribution over $(a_i, \bar{a}_{-i}, \theta)$ induced by P_i and μ_i.

We define an equilibrium via the following triadic relationship: (i) voters' strategies are optimal under their beliefs; (ii) their beliefs are consistent with their acquired information about the state and the poll; and (iii) the poll is consistent with their strategies. The equilibrium is the joint distribution over action profiles and states implied by voters' strategies and beliefs.

Definition 1. In an election $\mathcal{P}_N$, a distribution $P_N^* \in \Delta(A^N \times \Theta)$ is an **equilibrium** if it satisfies the following conditions:

1. The marginal distribution of θ equals the prior μ; namely, $\mu(\theta) = \sum_{a \in A^N} P_N^*(a, \theta)$.
2. Each voter i's strategy P_i is optimal under her belief μ_i, where μ_i is the marginal distribution of $(\bar{a}_{-i}, \theta)$ and P_i is the conditional distribution of a_i given $(\bar{a}_{-i}, \theta)$; namely,

$$\mu_i(x, \theta) = \sum_{a_i} \sum_{a_{-i}:\bar{a}_{-i}=x} P_N^*(a_i, a_{-i}, \theta),$$

$$P_i(a_i \mid x, \theta) = \frac{1}{\mu_i(x, \theta)} \sum_{a_{-i}:\bar{a}_{-i}=x} P_N^*(a_i, a_{-i}, \theta).$$

Equilibrium Interpretation This equilibrium concept builds on Denti's (2023) notion of robustness to information acquisition. We offer two interpretations. First, the equilibrium is interpreted as a blend of a Bayesian Nash equilibrium and a rational-expectations equilibrium (Hébert and La'O, 2023). It is a Bayesian Nash equilibrium in that agents behave optimally under uncertainty, and it is a rational-expectations equilibrium in that agents learn from endogenous aggregate behavior while simultaneously choosing their strategies as in Grossman and Stiglitz (1980). In our setting, the poll acts as an institutional device to aggregate voters' actions, thus facilitating the rational-expectations interpretation. Voters make decisions while learning about the state and the poll; their behavior is consistent with the poll because otherwise, the poll would not reflect the voting behavior.

[10] For a discrete random variable Y, the entropy is defined as $\mathbb{H}(Y) = -\sum_y p_Y(y) \ln p_Y(y)$, where p_Y is the probability mass function of Y. For another discrete random variable Z given Y, the conditional entropy $\mathbb{H}(Z \mid Y)$ is defined as $\mathbb{H}(Z \mid Y) = -\sum_y p_Y(y) \sum_z p_{Z|Y}(z \mid y) \ln p_{Z|Y}(z \mid y)$, where $p_{Z|Y}$ is the conditional probability mass function of Z given Y. See Cover and Thomas (2006) for a comprehensive treatment.

Second, although our model is static, the equilibrium can be interpreted as the long-run outcome of a dynamic electoral process. This interpretation aligns with the stochastic best-response dynamics interpretation. Imagine a sequential procedure in which, each period, a randomly chosen voter can acquire information and myopically revise her vote. Any revision is then reflected in a new poll, and the process repeats. This protocol yields a Markov chain on action profiles that converges to a unique stationary distribution, which coincides with our equilibrium, as shown in Denti (2020, working paper version) and Hoshino (2018). This interpretation captures an environment of frequent polling during an electoral campaign. Moreover, the assumption of myopic updating is sensible in a large electorate, where individual votes have a negligible impact.

We focus on symmetric equilibria, where the strategies and beliefs are identical across all voters. Even in a symmetric equilibrium, voters may receive distinct realized signals and take different actions. A symmetric equilibrium P_N^* is called **uninformative** if all voters choose to acquire no information; otherwise, it is called **informative**.[11]

2.3 Equilibrium Characterization

Given a symmetric equilibrium P_N^*, we define the **equilibrium vote share** as a random variable $\bar{a}_N^* = \sum_{i=1}^N a_i^*/N$, where $a_1^*, \ldots, a_N^*$ and θ are distributed according to P_N^*. The conditional probability of $\bar{a}_N^* = k/N$ (i.e., k votes for alternative 1 and $N - k$ votes for alternative 0) given state θ is denoted by

$$\Pr\left(\bar{a}_N^* = \frac{k}{N} \mid \theta\right) = \sum_{a:\bar{a}_N = k/N} P_N^*(a \mid \theta),$$

where Pr denotes the probability.

Under entropy-based costs, equilibrium behavior is characterized by a biased logit distribution (Matějka and McKay, 2015; Caplin, Dean, and Leahy, 2019; Denti, 2023). Using these results, we characterize the equilibrium distributions as follows.

Lemma 1. *In every election $\mathcal{P}_N$, every symmetric equilibrium P_N^* has some $p_N^* \in [0,1]$ such that the equilibrium vote share $\bar{a}_N^*$ has the following distribution: for each θ and each $k = 0, 1, \ldots, N$,*

$$\Pr\left(\bar{a}_N^* = \frac{k}{N} \mid \theta\right) = \frac{1}{Z_N(p_N^*, \theta)} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k}, \tag{1}$$

where $Z_N : [0,1] \times \Theta \to \mathbb{R}$ is the function defined by

$$Z_N(p, \theta) = \sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) p^k (1-p)^{N-k}, \tag{2}$$

[11] If $N \geq 3$, there always exist two uninformative equilibria in which all voters choose the same action. Indeed, if $N - 1$ voters vote for the same alternative, the remaining voter is never pivotal and thus acquires no information.

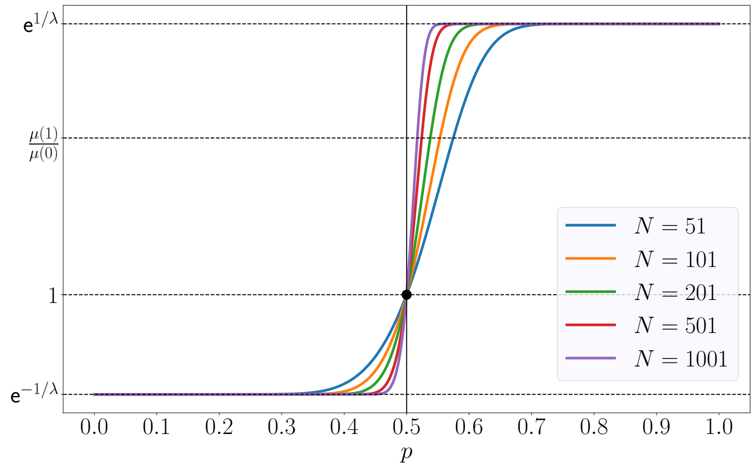

Figure 3: The graphs of $\frac{Z_N(p,1)}{Z_N(p,0)}$

and p_N^* equals the equilibrium marginal probability of each voter choosing action 1.[12].

The following properties hold:

1. *P_N^* is an uninformative equilibrium if and only if $p_N^* \in \{0,1\}$.*
2. *P_N^* is an informative equilibrium if and only if $p_N^* \in (0,1)$ and it is a solution to equation*

$$\frac{Z_N(p,1)}{Z_N(p,0)} = \frac{\mu(1)}{\mu(0)}. \tag{3}$$

Informative Equilibrium By Lemma 1, an informative equilibrium exists if and only if (3) has a solution $p_N^* \in (0,1)$. The function $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p and ranges from $\mathrm{e}^{-1/\lambda}$ to $\mathrm{e}^{1/\lambda}$, as depicted in Figure 3.[13] This implies that there is a unique informative equilibrium if $\frac{\mu(1)}{\mu(0)}$ is between $\mathrm{e}^{-1/\lambda}$ and $\mathrm{e}^{1/\lambda}$. We formalize this observation as follows.

Condition 1. An election $\mathcal{P}_N$ has the prior μ and the unit cost of information λ such that

$$\mathrm{e}^{-1/\lambda} < \frac{\mu(1)}{\mu(0)} < \mathrm{e}^{1/\lambda}.$$

Lemma 2. *An election $\mathcal{P}_N$ has an informative equilibrium if and only if it satisfies Condition 1. The informative equilibrium is unique whenever it exists.*

Condition 1 depends only on the prior μ and the unit cost λ, and not on the electorate size N. Under the prior $\mu(1) = 0.5$, Condition 1 is satisfied for any $\lambda > 0$; under any $\mu(1) \in (0,1)$, it is also satisfied when $\lambda > 0$ is sufficiently small.

We illustrate the vote-share distribution (1) at the informative equilibrium. Figure 1 (in Section 1) plots the probability mass functions at state $\theta = 1$ under the prior $\mu(1) = 0.5$, with the unit cost

[12]That is, $p_N^* = \sum_{a_{-i},\theta} P_N^*(1, a_{-i}, \theta)$ for each voter i

[13]These properties are shown in the proof of Lemma 2.

$\lambda = 1$. Figures 4 and 5 show these probability mass functions under $\mu(1) = 0.6$ and 0.4 respectively.[14] In each case, the equilibrium vote-share distribution "jumps" at the winning threshold of 0.5, thereby increasing the probability that majority vote chooses the correct alternative 1 (at state $\theta = 1$). Intuitively, when the poll indicates a close election, voters have stronger incentives to acquire information, helping them to choose the correct alternative.

2.4 Main Results

2.4.1 Equilibrium Probability of Correct Choice

A key question is whether a larger electorate can choose the correct alternative more (or less) accurately. To address this question, we consider the probability that majority vote chooses the correct alternative. We then define the **(unconditional) probability of correct choice** as

$$\Pr(u(\bar{a}_N^*, \theta) = 1) = \mu(1) \Pr\left(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1\right) + \mu(0) \Pr\left(\bar{a}_N^* < \frac{1}{2} \mid \theta = 0\right), \tag{4}$$

where $\Pr(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1)$ is the conditional probability of choosing alternative 1 given state $\theta = 1$ and $\Pr(\bar{a}_N^* < \frac{1}{2} \mid \theta = 0)$ is the conditional probability of choosing alternative 0 given state $\theta = 0$.

Our first main result is that the probability of correct choice is independent of both the electorate size N and the prior μ, although the vote-share distribution depends on both.

Theorem 1. *Every election $\mathcal{P}_N$ satisfies the following properties:*

1. *In an uninformative equilibrium, with all voters choosing the same alternative a, the probability of correct choice is $\mu(a)$.*
2. *In an informative equilibrium, which exists (and is unique) if and only if Condition 1 is satisfied, the probability of correct choice is*

$$\Pr(u(\bar{a}_N^*, \theta) = 1) = \frac{\mathsf{e}^{1/\lambda}}{1 + \mathsf{e}^{1/\lambda}}. \tag{5}$$

Moreover, the informative equilibrium, if it exists, yields a strictly higher probability of correct choice than all uninformative equilibria: $\mathsf{e}^{1/\lambda}/(1 + \mathsf{e}^{1/\lambda}) > \max\{\mu(1), \mu(0)\}$.

Theorem 1 reflects a perfect cancellation between Condorcet's optimism and Downs's pessimism. Increasing N has two opposing effects on decision quality: a positive effect that the electorate aggregates more individual signals, and a negative effect that each voter's signal becomes less informative. The classic Condorcet-type results focus on exogenous, independent signals, capturing the positive effect while ignoring the negative. In contrast, the rational ignorance hypothesis emphasizes endogenous (often independent) signals, highlighting the negative effect. In our model, signals are both endogenous and correlated, offsetting these effects.

[14] Condition 1 is satisfied since $\mu(1)/\mu(0) = 3/2$ and $2/3$ for $\mu(1) = 0.6$ and 0.4 respectively, and these ratios lie between $\mathrm{e}^{-1} \approx 0.368$ and $\mathrm{e} \approx 2.718$.

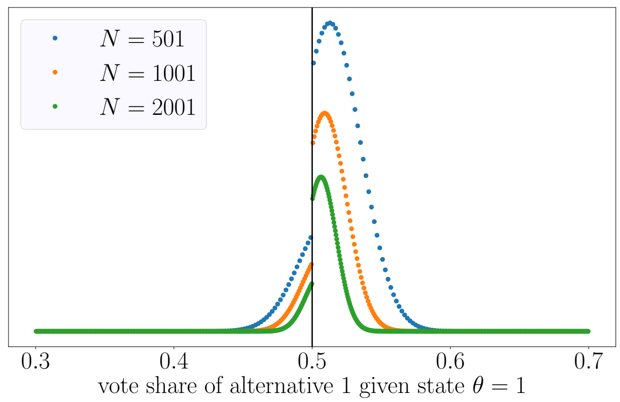

Figure 4: Prior $\mu(1) = 0.6$

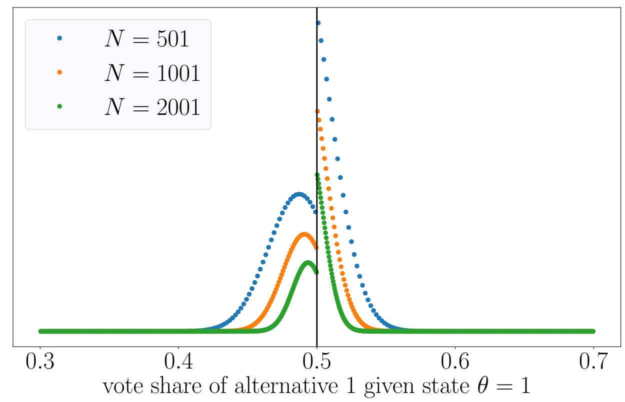

Figure 5: Prior $\mu(1) = 0.4$

Note: The equilibrium vote shares at state $\theta = 1$ under priors $\mu(1) = 0.6$ and 0.4

Here is a proof sketch for Theorem 1. The probability of choosing alternative 1 given state $\theta = 1$ is rewritten as $\Pr(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1) = \sum_{k=n+1}^{N} \Pr(\bar{a}_N^* = \frac{k}{N} \mid \theta = 1)$. By Lemma 1, we have

$$\Pr\left(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1\right) = \frac{\mathsf{e}^{1/\lambda} - \frac{\mu(0)}{\mu(1)}}{\mathsf{e}^{1/\lambda} - \mathsf{e}^{-1/\lambda}}. \tag{6}$$

Analogously, the probability of choosing alternative 0 given state $\theta = 0$ is

$$\Pr\left(\bar{a}_N^* < \frac{1}{2} \mid \theta = 0\right) = \frac{\mathsf{e}^{1/\lambda} - \frac{\mu(1)}{\mu(0)}}{\mathsf{e}^{1/\lambda} - \mathsf{e}^{-1/\lambda}}. \tag{7}$$

Note that both (6) and (7) are independent of N. By substituting them into (4), we obtain (5).

2.4.2 Large Elections

We examine the behavior of the equilibrium vote share $\bar{a}_N^*$ as the electorate size N grows (while we fix the prior μ and the unit cost λ). This analysis is nontrivial because the behavior of any two voters is possibly correlated and each voter's strategy changes with N.

We show that the marginal probability p_N^* of each voter choosing action 1 converges to $\frac{1}{2}$, even if the prior μ is biased toward one state. Recall that p_N^* is the (unique) solution to (3) in any informative equilibrium (Lemma 1), determined by the intersection of $\frac{Z_N(p,1)}{Z_N(p,0)}$ and $\frac{\mu(1)}{\mu(0)}$. This convergence is illustrated in Figure 3.

Lemma 3. *For every election $\mathcal{P}_N$ that satisfies Condition 1, let P_N^* be the informative equilibrium and p_N^* be the equilibrium marginal probability of each voter choosing action* 1*. Then,*

$$\lim_{N\to\infty} p_N^* = \frac{1}{2}.$$

Our second main result shows that a large election tends to be close. Figure 1 (in Section 1) illustrates that under the prior $\mu(1) = 0.5$, the informative equilibrium vote share converges in probability to the threshold of $\frac{1}{2}$ as the electorate size N increases. Figures 4 and 5 show that even under $\mu(1) = 0.6$ or 0.4, the equilibrium vote share still converges in probability to $\frac{1}{2}$.

Theorem 2. *For every election $\mathcal{P}_N$ that satisfies Condition 1, let P_N^* be the informative equilibrium and $\bar{a}_N^*$ be the equilibrium vote share. For each $\varepsilon > 0$,*

$$\lim_{N\to\infty} \Pr\left(\left|\bar{a}_N^* - \frac{1}{2}\right| < \varepsilon\right) = 1.$$

Here is the intuition behind Theorem 2. In a large electorate, each voter's probability of being pivotal is very small, so she acquires almost no information. Consequently, her conditional probabilities of choosing action 1 or 0 given state θ are close to the marginal probabilities p_N^* and $1 - p_N^*$, respectively. Since both are close to $\frac{1}{2}$ (Lemma 3), the equilibrium actions $a_1^*, \dots, a_N^*$ are approximated by i.i.d. random variables that take values 1 and 0 with equal probabilities. If the equilibrium actions were i.i.d., the law of large numbers would guarantee that the vote share converges to $\frac{1}{2}$. Since they are actually correlated (Lemma 1), we use the law of large numbers for an approximating i.i.d. process.[15]

2.5 Extensions

The main model in Section 2.1 imposes two assumptions: it uses simple majority rule; and all voters have common interests. Neither assumption is necessary to our main results.

2.5.1 Elections under Supermajority Rule

Consider the same model as in Section 2.1, except we now adopt a **supermajority rule**: alternative 1 is chosen if and only if the vote share $\bar{a}_N$ is at least a given threshold $\alpha \in (\frac{1}{2}, 1]$. This rule includes the **unanimity rule** as the special case of $\alpha = 1$. Each voter's payoff is 1 if the chosen alternative is correct and 0 otherwise. We write $\mathcal{P}_{N,\alpha}$ for this supermajority election with a poll (omitting the prior μ and the unit cost λ). We use the same equilibrium concept (Definition 1).

We extend our main results. The proofs for supermajority elections closely mirror those for simple majority elections and are deferred to Online Appendix B. Lemmas 1 and 2 also extend. In particular, an informative equilibrium exists if and only if Condition 1 holds, and it is unique whenever it exists. The existence condition does not depend on a winning threshold α.

The extension of Theorem 1 establishes that the (unconditional) probability of correct choice at the informative equilibrium equals $\mathrm{e}^{1/\lambda}/(1 + \mathrm{e}^{1/\lambda})$, independent of N, μ, and the winning threshold $\alpha \in (\frac{1}{2}, 1]$. Moreover, the conditional probabilities of choosing the correct alternative do not depend

[15] Hoshino and Ui (2024) study games in which rationally inattentive players strategically interact, examining asymptotic behavior as the number of players tends to infinity. Their model essentially differs from ours and so does their analysis. In our model, each voter's decision has a negligible impact on her own payoff in a large electorate, whereas in their model (which covers Keynesian beauty contests, etc.), a player's own action can still affect her payoff even in large games. This difference in modeling underlies distinct analytical methods.

on N or α. Thus, any apparent bias introduced by the voting rule is "perfectly absorbed" in the equilibrium.

Theorem 1′. *Theorem 1 holds as is, in every election $\mathcal{P}_{N,\alpha}$ with a winning threshold $\alpha \in (\frac{1}{2}, 1]$. Moreover, the conditional probabilities of correct choice, $\Pr(\bar{a}_N^* \geq \alpha \mid \theta = 1)$ and $\Pr(\bar{a}_N^* < \alpha \mid \theta = 0)$, equal the right-hand sides of (6) and (7), respectively.*

The extension of Theorem 2 states that, as N tends to infinity, the informative equilibrium vote share becomes closer to the winning threshold α.

Theorem 2′. *For every election $\mathcal{P}_{N,\alpha}$ that satisfies Condition 1, let P_N^* be the informative equilibrium and $\bar{a}_N^*$ be the equilibrium vote share. For each $\varepsilon > 0$,*

$$\lim_{N\to\infty} \Pr(|\bar{a}_N^* - \alpha| < \varepsilon) = 1.$$

2.5.2 Elections with Partisan Voters

We now incorporate "partisans," who always vote for their preferred alternative regardless of a realized state θ. Suppose that N_1 and N_0 partisans always take actions 1 and 0, respectively, with $N_1, N_0 \leq n$. (If either group exceeded half of the electorate, its preferred alternative would trivially win.) The remaining $M \equiv N - N_1 - N_0 \geq 1$ voters are non-partisans, each receiving a payoff of 1 if the chosen alternative is correct and 0 otherwise. Under simple majority rule, alternative 1 is chosen if and only if at least $n + 1 - N_1$ non-partisans vote for it.

This election is equivalent to a partisan-free but supermajority election $\mathcal{P}_{M,\alpha}$ with a winning threshold $\alpha = (N - 2N_1)/(2M)$. To see why, note that the vote share of alternative 1, $\bar{a}_M = \sum_{i=1}^{M} a_i/M$, is at least α if and only if in the original election, the vote share of alternative 1, $(M\bar{a}_M + N_1)/N$, exceeds one-half. Hence, Theorems 1′ and 2′ apply directly.[16]

3 Comparison between Elections with and without Opinion Polls

Comparing elections with polls and those without, we show that polls can help large electorates make the correct choice.

3.1 Elections without Opinion Polls

Base Environment The base environment is the same as in Section 2, but we use different notation in order to avoid confusion with elections with polls. In an election without a poll, we denote voter i's action by b_i. Let $b = (b_1, \ldots, b_N)$ be an action profile and b_{-i} be the actions of all voters but i. Given b, we define the vote share of alternative 1 by $\bar{b}_N = \sum_{i=1}^{N} b_i/N \in [0, 1]$.

[16] Theorem 2′ extends if $N - N_1 - N_0 \to \infty$, even if $N_1 \to \infty$ or $N_0 \to \infty$ (or both).

Information Acquisition We model information acquisition as in Section 2, except voters now have no access to polling information; they learn only about the state θ. Without loss of generality, voter i's strategy can be represented by a system of conditional action distributions $Q_i(\cdot \mid \theta) \in \Delta(A)$ for each θ. All voters' choices of action are conditionally independent given θ.

The cost of information is modeled as before. Given voter i's strategy Q_i, the expected reduction of the entropy of her prior μ is $\mathbb{I}(\theta; b_i) = \mathbb{H}(\theta) - \mathbb{H}(\theta \mid b_i)$. Here $\mathbb{H}(\theta)$ is the entropy of θ, and $\mathbb{H}(\theta \mid b_i)$ is the conditional entropy of θ given b_i, where (b_i, θ) is distributed according to Q_i and μ. The information cost of Q_i is $\lambda \mathbb{I}(\theta; b_i)$, where $\lambda > 0$ is the unit cost of information.

We denote by $\mathcal{Q}_N$ the **election without a poll** (omitting the prior μ and the unit cost λ).

Equilibrium Voter i's strategy Q_i is called optimal given the others' strategies $Q_{-i} = (Q_j)_{j \neq i}$ if it maximizes her expected payoff net of information costs: $\mathbb{E}[u(\bar{b}_N, \theta)] - \lambda \mathbb{I}(\theta; b_i)$, where $\mathbb{E}[u(\bar{b}_N, \theta)]$ is the expected payoff with respect to the distribution over (b_i, b_{-i}, θ) induced by Q_i, Q_{-i}, and μ.

We focus on symmetric equilibria as in Section 2. A symmetric **(Nash) equilibrium** is a strategy profile $(Q_N^*, \ldots, Q_N^*)$ in which each voter's strategy Q_N^* is optimal given that all voters choose the same strategy Q_N^*. By a slight abuse of notation, we denote a symmetric equilibrium by the strategy Q_N^*. A symmetric equilibrium Q_N^* is called **uninformative** if all voters acquire no information; otherwise, it is called **informative**.

Given a symmetric equilibrium Q_N^*, we define the **equilibrium vote share** as a random variable $\bar{b}_N^* = \sum_{i=1}^N b_i^*/N$, where $b_1^*, \ldots, b_N^*$ are i.i.d. random variables distributed according to $Q_N^*(\cdot \mid \theta)$ given state θ. The conditional probability of $\bar{b}_N^* = k/N$ (i.e., k votes for alternative 1 and $N-k$ votes for alternative 0) given state θ is denoted by

$$\Pr\Big(\bar{b}_N^* = \frac{k}{N} \mid \theta\Big) = \sum_{b: \bar{b}_N = k/N} \prod_{i=1}^N Q_N^*(b_i \mid \theta).$$

When all voters $j \neq i$ choose the same strategy Q_N^*, voter i's gross payoff (excluding information costs) from choosing a strategy Q_i is

$$\sum_\theta \mu(\theta) \sum_{k=0}^{2n} \binom{2n}{k} (Q_N^*(1 \mid \theta))^k (Q_N^*(0 \mid \theta))^{2n-k} \sum_{b_i} Q_i(b_i \mid \theta) u(\bar{b}_N, \theta),$$

where $\bar{b}_N = (k + b_i)/N$ is the vote share of alternative 1 when among the other $N-1 = 2n$ voters, k vote for alternative 1 and $2n-k$ vote for alternative 0. Voter i's action b_i matters only if she is pivotal: if $k = n$, her gross payoff is $\mathbb{1}\{b_i = \theta\}$; otherwise, it does not depend on b_i. When all voters $j \neq i$ play the same strategy Q_N^*, voter i's pivot probability at state θ is

$$\Pi_N(\theta) \equiv \binom{2n}{n} (Q_N^*(1 \mid \theta))^n (Q_N^*(0 \mid \theta))^n.$$

Voter i's gross payoff simplifies to $\sum_\theta \mu(\theta) \Pi_N(\theta) Q_i(\theta \mid \theta)$ plus a constant. Her problem is, therefore, equivalent to

$$\max_{Q_i} \quad \sum_{\theta} \mu(\theta)\Pi_N(\theta)Q_i(\theta \mid \theta) - \lambda\mathbb{I}(\theta; b_i). \tag{8}$$

As shown by Matějka and McKay (2015), the solution to this type of problem follows a biased logit distribution. Based on their results, we obtain the following lemma.

Lemma 4. *In every election $\mathcal{Q}_N$, every symmetric equilibrium Q_N^* has some $q_N^* \in [0,1]$ such that*

$$Q_N^*(1 \mid 1) = \frac{q_N^* \mathrm{e}^{\Pi_N(1)/\lambda}}{q_N^* \mathrm{e}^{\Pi_N(1)/\lambda} + 1 - q_N^*}, \quad Q_N^*(1 \mid 0) = \frac{q_N^*}{q_N^* + (1-q_N^*)\mathrm{e}^{\Pi_N(0)/\lambda}}, \tag{9}$$

where q_N^* equals the marginal probability of each voter choosing action 1.[17]

The following properties hold:

1. *Q_N^* is an uninformative equilibrium strategy if and only if $q_N^* \in \{0,1\}$.*
2. *Q_N^* is an informative equilibrium strategy if and only if $q_N^* \in (0,1)$.*

Large Elections As the electorate size N grows, each voter's pivot probability vanishes, leading them to acquire less information. Indeed, we have $\Pi_N(\theta) \to 0$ as $N \to \infty$ (as shown in the proof of Lemma 5), which implies that $|Q_N^*(1 \mid 1) - q_N^*| \to 0$ and $|Q_N^*(1 \mid 0) - q_N^*| \to 0$ (Lemma 4). For any subsequence of $\{q_N^*\}_N$ with the limit q_∞^*, the equilibrium vote share $\bar{b}_N^*$ then converges in probability to a constant. Formally, we have the following lemma:

Lemma 5. *For every election $\mathcal{Q}_N$, let Q_N^* be a symmetric equilibrium and q_N^* be the equilibrium marginal probability of each voter choosing action 1. For any subsequence, still denoted $\{Q_N^*\}_N$, such that the corresponding $\{q_N^*\}_N$ converges to q_∞^*, and for any $\varepsilon > 0$,*

$$\lim_{N\to\infty} \Pr\Big(\Big|\bar{b}_N^* - q_\infty^*\Big| < \varepsilon\Big) = 1.$$

By Lemma 5, if $q_\infty^* > \frac{1}{2}$ then the probability of choosing alternative 1 converges to 1 and thus the probability of correct choice approaches $\mu(1)$; similarly, if $q_\infty^* < \frac{1}{2}$ then the probability of correct choice approaches $\mu(0)$. The only nontrivial case is $q_\infty^* = \frac{1}{2}$. The next lemma characterizes the limit probability of correct choice in this case.

Lemma 6. *For every election $\mathcal{Q}_N$, let Q_N^* be a symmetric equilibrium and q_N^* be the equilibrium marginal probability of each voter choosing action 1. For any subsequence, still denoted $\{Q_N^*\}_N$, such that the corresponding $\{q_N^*\}_N$ converges to q_∞^*, the following properties hold:*

1. *If $q_\infty^* > \frac{1}{2}$ then the equilibrium probability of correct choice converges to $\mu(1)$, while if $q_\infty^* < \frac{1}{2}$ then the equilibrium probability of correct choice converges to $\mu(0)$.*
2. *If $q_\infty^* = \frac{1}{2}$ then the equilibrium probability of correct choice converges to $\mu(1)$, $\mu(0)$, or*

$$\mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)$$

[17] That is, $q_N^* = \mu(1)Q_N^*(1 \mid 1) + \mu(0)Q_N^*(1 \mid 0)$.

where $(t^1, t^0) \in \mathbb{R}^2_{++}$ *is a solution to the system of equations*

$$\lambda\mu(1)(t^1 + t^0) = \phi(t^0), \qquad \lambda\mu(0)(t^1 + t^0) = \phi(t^1).$$

Here, Φ *and* ϕ *denote the standard normal cdf and pdf, respectively.*

In the proof of Lemma 6, we cannot directly apply a standard central limit theorem since voters' equilibrium strategies Q^*_N vary with the electorate size N. Instead, we use the Berry–Esseen theorem, which bounds the discrepancy between the actual distribution and its normal approximation at any finite N, enabling a controlled normal approximation of $\bar{b}^*_N$.

3.2 Comparison between Elections with and without Opinion Polls

Now we compare the probability of correct choice in elections with and without polls. We show that for any unit cost of information λ and any nearly symmetric prior $\mu(1) \approx \frac{1}{2}$, the presence of a poll strictly increases the probability of correct choice in large electorates.

Proposition 1. *For any unit cost of information* $\lambda > 0$*, there exists an* $\varepsilon > 0$ *such that for any prior* μ *with* $|\mu(1) - \frac{1}{2}| < \varepsilon$*, there exists an* $\bar{N} \in \mathbb{N}$ *such that for all* $N > \bar{N}$*, the probability of correct choice in the informative equilibrium* P^*_N *of the election* $\mathcal{P}_N$ *is strictly greater than that in any symmetric equilibrium* Q^*_N *of the election* $\mathcal{Q}_N$*.*

Figure 6 illustrates Proposition 1 under the prior $\mu(1) = 0.5$. This figure plots the limit probabilities of correct choice when the electorate size N grows: the blue graph plots the informative equilibrium probability of correct choice with a poll, $\mathrm{e}^{1/\lambda}/(1 + \mathrm{e}^{1/\lambda})$ (Theorem 1); and the orange graph plots the (highest) limit equilibrium probability of correct choice without a poll (Lemma 6).

These probabilities illustrate how information accessibility shapes collective accuracy. When $\lambda \to 0$ (i.e., learning is essentially free), the probability of correct choice converges to 1 regardless of whether polls are present; when $\lambda \to \infty$ (i.e., learning is prohibitively expensive), the probability of correct choice falls to $\frac{1}{2}$ in both settings. Hence, polls have no impact in these two extreme cases. However, for any finite $\lambda > 0$, the presence of polls strictly increases the probability of making the correct choice. In this sense, the effect of information accessibility on collective accuracy is non-monotonic.

4 Regression Discontinuity Design in Close Elections

Our findings provide new insights into the **regression discontinuity (RD) design** in close elections. The RD design in electoral contexts identifies treatment effects based on the assumption that candidates who barely win or lose are comparable in all respects except the election outcome. Yet our results show that when polls are available, better candidates are systematically more likely to win close elections. This insight offers a new perspective on the mixed evidence concerning RD validity in the existing literature.

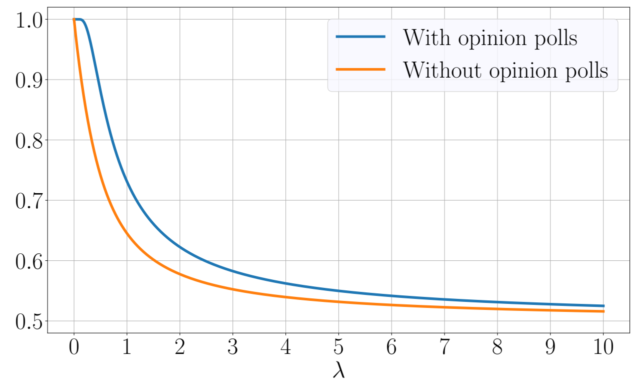

Figure 6: Proposition 1 under the prior $\mu(1) = 0.5$

Regression Discontinuity Design The RD design, initially developed by Thistlethwaite and Campbell (1960), is an empirical strategy for identifying treatment effects without random assignment of subjects to treatments. In electoral settings, the RD design is applied to close elections in which narrowly winning and narrowly losing candidates are presumed to form a quasi-experimental comparison. The validity of the RD design relies on a continuity assumption: namely, candidates whose vote shares fall just above the winning threshold (i.e., who barely win) should be similar in unobservable traits to those whose shares lie just below the threshold (i.e., who barely lose) (Hahn, Todd, and Van der Klaauw, 2001).

Many studies apply the RD design to close elections. Lee (2001, 2008) and Butler (2009) estimate incumbency advantage in subsequent U.S. congressional elections; Ferreira and Gyourko (2009, 2014) examine how party affiliation or politicians' gender influence municipal fiscal policy; Firpo, Ponczek, and Sanfelice (2015) analyzes Brazil's federal budgeting; and Dell (2015) investigates the effects of law enforcement on drug-trafficking networks and drug-related violence in Mexico.

Despite its widespread adoption, the validity of the RD design in close elections remains debated. This validity hinges on the continuity assumption, which breaks down if certain types of candidates systematically win close races, a phenomenon called sorting. As discussed in Section 1, existing evidence on sorting is mixed. As de la Cuesta and Imai (2016) note, "the literature is remarkably divided on the question of whether sorting exists in the close election context."

Novel Mechanism for Sorting As noted in Section 1, Eggers, Fowler, Hainmueller, Hall, and Snyder (2015) suggest that polling could generate sorting in close elections. Our mechanism supports their insight. Figures 1 and 2 illustrate the equilibrium vote-share distributions. In elections *with* polls (Figure 1), a discontinuity emerges at the winning threshold of 0.5, because in close elections, voters perceive a higher pivot probability and aggressively learn more to vote for the correct candidate. This heightened coordination near 0.5 increases the probability that the correct candidate wins. By contrast, in elections *without* polls (Figure 2), voters remain ignorant of their potential pivotality and no such discontinuity arises.[18]

Our results suggest that when polls are available, the superior candidate wins with a significantly higher probability, regardless of how narrow the margin might be (Theorems 1 and 2). Existing analysis of sorting relies on the fact that although the identity of the superior candidate may be unobservable, it may correlate with observable characteristics (e.g., incumbency status or financial resources). Our findings suggest that sorting may arise in elections with polls even when such correlation is absent, as long as voters acquire information unavailable to the researchers. Consequently, even if winners and losers in close elections appear balanced on observable covariates, the RD design may warrant careful consideration when voters have access to polls. Our results do not invalidate the RD design in close elections; instead, we provide a complementary explanation for the mixed empirical findings in the literature.

5 Concluding Remarks

We analyzed how opinion polls affect information acquisition and aggregation in voting. To do so, we developed a model that captures the interaction between voting and polling. We conclude by discussing several extensions and limitations of our analysis.

5.1 Role of Entropy-Based Costs of Information

We focused on entropy-based costs of information, arguably the most standard specification in the literature (e.g., Sims, 2003; Matějka and McKay, 2015). Our formal arguments relied on this functional form. For example, Lemma 1 (that gives the biased logit characterization of equilibrium) is specific to the entropy-based costs. In addition, we used the properties of the entropy-based costs in some steps of the proofs of Theorems 1 and 2.

Nevertheless, our intuition for these theorems extends beyond the entropy-based costs of information. The key intuition is that information sharing through opinion polls can encourage voters to acquire more information when elections are likely to be close (i.e., when the voters could be pivotal) and can mitigate free-riding incentives. This intuition does not depend on any specific cost structure and thus appears to be robust. An important direction for future research would be to understand how these results depend on this specification of information costs.

5.2 Information Acquisition as Public Good Provision

An analogy to public-good provision is useful to better understand Theorem 1 (that with opinion polls, the probability that majority vote makes the correct choice is independent of the number of voters). In the voting context, information, once shared, is non-rivalrous and non-excludable, which makes information acquisition analogous to the voluntary provision of public goods. In standard

[18] Our criticism does not apply to studies exploiting actual randomization in elections with exact ties (Hyytinen, Meriläinen, Saarimaa, Toivanen, and Tukiainen, 2018; De Magalhães, Hangartner, Hirvonen, Meriläinen, Ruiz, and Tukiainen, 2025).

public-good provision games (e.g., Bergstrom, Blume, and Varian, 1986), the equilibrium quantity of the public good is determined by equating individual marginal benefits to marginal costs; under quasi-linear utilities, this equilibrium quantity is independent of the number of individuals. In our model, as in Downs's hypothesis of rational ignorance, the equilibrium provision of information is set by individual marginal benefits, discounted by their pivot probabilities, equaling marginal costs. Furthermore, the additively separable utilities in rational inattention resemble quasi-linear utilities. These parallels illuminate how Theorem 1 connects to the classic insight on public-good provision.

5.3 Elections with Abstention

We restricted voters to choosing between two alternatives, with no abstention allowed. This assumption facilitated comparisons with existing studies, highlighting the role of opinion polls. Nevertheless, it is worth noting that the public-good intuition from the previous subsection naturally extends to elections in which voters are allowed to abstain. From the perspective that views informed voting as a public-good provision, if some voters opt to abstain (thus reducing their information provision to the public), this loss would be compensated for by the rest of the electorate. Based on this intuition, we conjecture that if abstention is allowed, the equilibrium probability of making the correct choice would be independent of the electorate size, and this probability would remain constant across equilibria, although equilibrium uniqueness might no longer hold.

5.4 Opinion Polls Based on Random Sampling

In our model, the opinion poll represents a complete census aggregating all voters' intentions. In practice, polls are typically based on random samples of the electorate. While sampling error would introduce additional noise into polling information, sufficiently large samples should approximate the same information as a complete census. We thus expect our main results to hold under uniform random sampling. Going forward, new questions of polling design would arise: not only determining the optimal sample size, but also what to ask and whom to ask (given that voters' opinions are correlated with observable covariates). Such questions, which remain largely unexplored in the literature, would connect our work to research on information design (e.g., Bergemann and Morris, 2019).

A Appendix

A.1 Lemma 1

Our base game is a common-interest game and thus a potential game (Monderer and Shapley, 1996). Indeed, the payoff function u is a potential. Lemma A follows directly from Denti's (2023) Corollary 1 applied to our setting.

Lemma A. *In an election $\mathcal{P}_N$, every symmetric equilibrium P_N^* has some $p_N^* \in [0,1]$ such that for each $a = (a_1, \dots, a_N)$ and each θ,*

$$P_N^*(a_1, \dots, a_N \mid \theta) = \frac{\exp\left(\frac{u(\bar{a}_N, \theta)}{\lambda}\right) \prod_{i:a_i=1} p_N^* \prod_{i:a_i=0} (1 - p_N^*)}{\sum_{a' \in \{0,1\}^N} \exp\left(\frac{u(\bar{a}'_N, \theta)}{\lambda}\right) \prod_{i:a'_i=1} p_N^* \prod_{i:a'_i=0} (1 - p_N^*)}, \tag{10}$$

where $(p_N^*, \dots, p_N^*)$ is a symmetric pure-strategy Nash equilibrium of the normal-form game such that all players $i = 1, \dots, N$ have the same action space $[0,1]$ and the same payoff function

$$U(p'_1, \dots, p'_N) = \sum_\theta \mu(\theta) \ln\left(\sum_{a' \in \{0,1\}^N} \exp\left(\frac{u(\bar{a}'_N, \theta)}{\lambda}\right) \prod_{i:a'_i=1} p'_i \prod_{i:a'_i=0} (1 - p'_i)\right).$$

Using Lemma A, we show Lemma 1. Consider the denominator on the right-hand side of (10). Reorganizing the sum over all action profiles a' by grouping them according to the number of ones in $a' = (a'_1, \dots, a'_N)$, or equivalently partitioning the sum according to $\bar{a}'_N = \frac{k}{N}$ for $k = 0, 1, \dots, N$, we can rewrite the denominator as

$$\begin{aligned}\sum_{k=0}^{N} \sum_{a':\bar{a}'_N=k/N} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k} &= \sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k} \\ &= Z_N(p_N^*, \theta),\end{aligned}$$

where the first equality follows from the number of action profiles a' with exactly k ones being $\binom{N}{k}$ and the second one follows from (2). Then, we obtain (1) as follows:

$$\Pr\left(\bar{a}_N^* = \frac{k}{N} \mid \theta\right) = \binom{N}{k} \cdot \frac{1}{Z_N(p_N^*, \theta)} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k}.$$

where the number of actions profiles a with its average $\bar{a}_N = k/N$ is $\binom{N}{k}$ and the equilibrium P_N^* is symmetric (so that each of such action profiles has equal probability).

To prove Lemma 1, it suffices to show that any symmetric pure-strategy Nash equilibrium $(p_N^*, \dots, p_N^*)$ of the normal-form game of Lemma A is either $p_N^* \in \{0,1\}$ or p_N^* being a solution to (3). At a symmetric Nash equilibrium, it must be that $p_N^* \in \operatorname{argmax}_{p'_i} U(p'_i, p_N^*, \dots, p_N^*)$. This is immediate if $p_N^* \in \{0,1\}$. In the interior case of $p_N^* \in (0,1)$, we must have the first-order condition, $\frac{\partial U}{\partial p'_i}(p'_i, p_N^*, \dots, p_N^*)|_{p'_i = p_N^*} = 0$. The first-order condition is sufficient since $U(p'_i, p_N^*, \dots, p_N^*)$ is strictly concave in p'_i (Caplin, Dean, and Leahy, 2019, p. 1066). We write the first-order condition as

$$\sum_{\theta} \mu(\theta) \cdot \frac{\sum_{k=0}^{2n} \binom{2n}{k} \left[\exp\left(\frac{u(\frac{k+1}{N}, \theta)}{\lambda} \right) - \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda} \right) \right] (p_N^*)^k (1-p_N^*)^{2n-k}}{\underbrace{\sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda} \right) (p_N^*)^k (1-p_N^*)^{N-k}}_{= Z_N(p_N^*, \theta) \text{ by (2)}}} = 0. \tag{11}$$

In the numerator, if $k \neq n$, the square bracket is zero since $u(\frac{k+1}{N}, \theta) = u(\frac{k}{N}, \theta)$, while if $k = n$, the square bracket is $\mathsf{e}^{1/\lambda} - 1$ when $\theta = 1$ and $1 - \mathsf{e}^{1/\lambda}$ when $\theta = 0$. By substituting them into (11),

$$\frac{\mu(1)}{Z_N(p_N^*, 1)} \binom{2n}{n} (p_N^*)^n (1-p_N^*)^n (\mathsf{e}^{1/\lambda} - 1) + \frac{\mu(0)}{Z_N(p_N^*, 0)} \binom{2n}{n} (p_N^*)^n (1-p_N^*)^n (1 - \mathsf{e}^{1/\lambda}) = 0.$$

Rearranging the terms, we obtain (3). For $p_N^* \in (0,1)$, if $(p_N^*, \dots, p_N^*)$ is a Nash equilibrium then p_N^* is a solution to (3).

A.2 Lemma 2

We define the function $W_N : [0,1] \times \Theta \to \mathbb{R}$ as

$$\begin{aligned} W_N(p, 1) &= \sum_{k=n+1}^{N} \binom{N}{k} p^k (1-p)^{N-k}, \\ W_N(p, 0) &= \sum_{k=0}^{n} \binom{N}{k} p^k (1-p)^{N-k}. \end{aligned} \tag{12}$$

Note that $W_N(p,1) + W_N(p,0) = 1$ by the binomial theorem. Note that $W_N(p,1)$ is strictly increasing in p and $W_N(p,0)$ is strictly decreasing in p, as can be shown by differentiation. Then, we rewrite Z_N, as defined in (2), as

$$\begin{aligned} Z_N(p, 1) &= W_N(p, 0) + \mathsf{e}^{1/\lambda} W_N(p, 1), \\ Z_N(p, 0) &= \mathsf{e}^{1/\lambda} W_N(p, 0) + W_N(p, 1). \end{aligned} \tag{13}$$

Note that $Z_N(p,1)$ is continuous and strictly increasing in p and $Z_N(p,0)$ is continuous and strictly decreasing in p.[19] Hence, $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p.

By Lemma 1, it suffices to show that (3) has a unique solution if Condition 1 is satisfied and no solution otherwise. Note that $W_N(1,1) = W_N(0,0) = 1$ and $W_N(0,1) = W_N(1,0) = 0$. By (13),

$$\frac{Z_N(0,1)}{Z_N(0,0)} = \mathsf{e}^{-1/\lambda}, \quad \frac{Z_N(1,1)}{Z_N(1,0)} = \mathsf{e}^{1/\lambda}.$$

[19]To see that $Z_N(p,1)$ is strictly increasing in p, we note that $Z_N(p,1) = 1 + (\mathsf{e}^{1/\lambda} - 1)W_N(p,1)$ since $W_N(p,0) + W_N(p,1) = 1$ and recall that $W_N(p,1)$ is strictly increasing in p. Similarly, $Z_N(p,0)$ is strictly decreasing in p.

If Condition 1 is satisfied, then since $\frac{Z_N(0,1)}{Z_N(0,0)} < \frac{\mu(1)}{\mu(0)} < \frac{Z_N(1,1)}{Z_N(1,0)}$, (3) has a unique solution $p_N^* \in (0,1)$, where $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p. If Condition 1 is not satisfied, then either $\frac{Z_N(0,1)}{Z_N(0,0)} = \mathsf{e}^{-1/\lambda} \geq \frac{\mu(1)}{\mu(0)}$ or $\frac{Z_N(1,1)}{Z_N(1,0)} = \mathsf{e}^{1/\lambda} \leq \frac{\mu(1)}{\mu(0)}$, in both of which cases (3) has no solution in $(0,1)$.

A.3 Theorem 1

The conditional probability of choosing alternative 1 given state $\theta = 1$ is

$$\Pr\left(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1\right) = \sum_{k=n+1}^{N} \Pr\left(\bar{a}_N^* = \frac{k}{N} \mid \theta = 1\right).$$

By Lemma 1, this right-hand side equals

$$\frac{1}{Z_N(p_N^*, 1)} \sum_{k=n+1}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, 1)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k}.$$

Since $u(\frac{k}{N}, 1) = 1$ for all $k = n+1, \ldots, N$, by (12) and (13), we have

$$\Pr\left(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1\right) = \frac{\mathsf{e}^{1/\lambda} W_N(p_N^*, 1)}{W_N(p_N^*, 0) + \mathsf{e}^{1/\lambda} W_N(p_N^*, 1)}.$$

By (13), we rewrite (3) as

$$\frac{W_N(p_N^*, 0) + \mathsf{e}^{1/\lambda} W_N(p_N^*, 1)}{\mathsf{e}^{1/\lambda} W_N(p_N^*, 0) + W_N(p_N^*, 1)} = \frac{\mu(1)}{\mu(0)},$$

or equivalently, $\frac{W_N(p_N^*,1)}{W_N(p_N^*,0)} = \frac{\mathsf{e}^{1/\lambda}\mu(1) - \mu(0)}{\mathsf{e}^{1/\lambda}\mu(0) - \mu(1)}$. By substitution,

$$\Pr\left(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1\right) = \frac{\mathsf{e}^{1/\lambda} - \frac{\mu(0)}{\mu(1)}}{\mathsf{e}^{1/\lambda} - \mathsf{e}^{-1/\lambda}}.$$

Similarly, the conditional probability of choosing alternative 0 given state $\theta = 0$ is

$$\Pr\left(\bar{a}_N^* < \frac{1}{2} \mid \theta = 0\right) = \frac{\mathsf{e}^{1/\lambda} - \frac{\mu(1)}{\mu(0)}}{\mathsf{e}^{1/\lambda} - \mathsf{e}^{-1/\lambda}}.$$

Hence, we obtain that

$$\Pr(u(\bar{a}_N^*, \theta) = 1) = \mu(1) \Pr\left(\bar{a}_N^* > \frac{1}{2} \mid \theta = 1\right) + \mu(0) \Pr\left(\bar{a}_N^* < \frac{1}{2} \mid \theta = 0\right) = \frac{\mathsf{e}^{1/\lambda}}{1 + \mathsf{e}^{1/\lambda}}.$$

We show that $\mathsf{e}^{1/\lambda}/(1+\mathsf{e}^{1/\lambda}) > \max\{\mu(1), \mu(0)\}$ if the informative equilibrium exists. Since the existence is equivalent to Condition 1 (Lemma 2), it suffices to prove that Condition 1 implies the inequality. By simple algebra, we rewrite Condition 1 as $1/(1+\mathsf{e}^{1/\lambda}) < \mu(1) < \mathsf{e}^{1/\lambda}/(1+\mathsf{e}^{1/\lambda})$, or equivalently, $1/(1+\mathsf{e}^{1/\lambda}) < \mu(0) < \mathsf{e}^{1/\lambda}/(1+\mathsf{e}^{1/\lambda})$, which implies the desired inequality.

A.4 Lemma 3

To prove Lemma 3, it suffices to show that for any small $\varepsilon > 0$, if N is sufficiently large,

$$\frac{Z_N(\frac{1}{2}-\varepsilon,1)}{Z_N(\frac{1}{2}-\varepsilon,0)} < \frac{\mu(1)}{\mu(0)} < \frac{Z_N(\frac{1}{2}+\varepsilon,1)}{Z_N(\frac{1}{2}+\varepsilon,0)}. \tag{14}$$

To see this sufficiency, note that since $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p (as shown in the proof of Lemma 2), if (14) is true then $p_N^* \in (\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon)$, where p_N^* is a solution to (3).

We show auxiliary inequalities. For any small $\delta > 0$, there is an N_δ such that for any $N > N_\delta$,

$$\begin{aligned} &W_N(\tfrac{1}{2}+\varepsilon,1) > 1-\delta, \quad W_N(\tfrac{1}{2}+\varepsilon,0) < \delta, \\ &W_N(\tfrac{1}{2}-\varepsilon,0) > 1-\delta, \quad W_N(\tfrac{1}{2}-\varepsilon,1) < \delta, \end{aligned} \tag{15}$$

where W_N is defined in (12). To see these inequalities, let $w_1, \ldots, w_N$ be i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\frac{1}{2}+\varepsilon$ and $\frac{1}{2}-\varepsilon$ respectively. Then, $W_N(\frac{1}{2}+\varepsilon,1)$ and $W_N(\frac{1}{2}+\varepsilon,0)$ are the probabilities that the sample average $\sum_{i=1}^N w_i/N$ is, respectively, strictly greater than $\frac{1}{2}$ and strictly less than $\frac{1}{2}$. By the law of large numbers, there is an N'_δ such that for any $N > N'_\delta$, we have $W_N(\frac{1}{2}+\varepsilon,1) > 1-\delta$ and $W_N(\frac{1}{2}+\varepsilon,0) < \delta$. To see the other two inequalities, let $w'_1, \ldots, w'_N$ be i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\frac{1}{2}-\varepsilon$ and $\frac{1}{2}+\varepsilon$ respectively. By the same argument, there is an N''_δ such that for any $N > N''_\delta$, we have $W_N(\frac{1}{2}-\varepsilon,0) > 1-\delta$ and $W_N(\frac{1}{2}-\varepsilon,1) < \delta$. Lastly, let $N_\delta = \max\{N'_\delta, N''_\delta\}$.

We show another inequality. Under Condition 1, there is a small $\delta > 0$ such that

$$\frac{1+\mathsf{e}^{1/\lambda}\delta}{\mathsf{e}^{1/\lambda}(1-\delta)} < \frac{\mu(1)}{\mu(0)} < \frac{\mathsf{e}^{1/\lambda}(1-\delta)}{\mathsf{e}^{1/\lambda}\delta+1}. \tag{16}$$

To see this inequality, note that $\mathsf{e}^{-1/\lambda} < \frac{\mu(1)}{\mu(0)} < \mathsf{e}^{1/\lambda}$ (Condition 1) and that for a small enough δ, the left- and right-hand sides of (16) are arbitrarily close to $\mathsf{e}^{-1/\lambda}$ and $\mathsf{e}^{1/\lambda}$, respectively.

Now we prove (14). For any $N > N_\delta$,

$$\frac{Z_N(\frac{1}{2}+\varepsilon,1)}{Z_N(\frac{1}{2}+\varepsilon,0)} = \frac{W_N(\frac{1}{2}+\varepsilon,0)+\mathsf{e}^{1/\lambda}W_N(\frac{1}{2}+\varepsilon,1)}{\mathsf{e}^{1/\lambda}W_N(\frac{1}{2}+\varepsilon,0)+W_N(\frac{1}{2}+\varepsilon,1)} > \frac{\mathsf{e}^{1/\lambda}(1-\delta)}{\mathsf{e}^{1/\lambda}\delta+1} > \frac{\mu(1)}{\mu(0)},$$

where we use (13) for the equality, (15) for the first inequality, and (16) for the second one. Similarly,

$$\frac{Z_N(\frac{1}{2}-\varepsilon,1)}{Z_N(\frac{1}{2}-\varepsilon,0)} = \frac{W_N(\frac{1}{2}-\varepsilon,0)+\mathsf{e}^{1/\lambda}W_N(\frac{1}{2}-\varepsilon,1)}{\mathsf{e}^{1/\lambda}W_N(\frac{1}{2}-\varepsilon,0)+W_N(\frac{1}{2}-\varepsilon,1)} < \frac{1+\mathsf{e}^{1/\lambda}\delta}{\mathsf{e}^{1/\lambda}(1-\delta)} < \frac{\mu(1)}{\mu(0)}.$$

Hence, we have (14), which completes the proof.

A.5 Theorem 2

Fix any $\theta \in \Theta$ and any $\varepsilon > 0$. By Lemma 1,

$$\Pr\left(\left|\bar{a}_N^* - \frac{1}{2}\right| \geq \varepsilon \mid \theta\right) = \frac{1}{Z_N(p_N^*, \theta)} \underbrace{\sum_{k:|\frac{k}{N}-\frac{1}{2}|\geq\varepsilon} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1-p_N^*)^{N-k}}_{(17^*)}, \tag{17}$$

where the sum runs over all $k = 0, 1, \ldots, N$ such that $|\frac{k}{N} - \frac{1}{2}| \geq \varepsilon$. Since $u(\frac{k}{N}, \theta) \leq 1$ for all k,

$$(17^*) \leq \mathrm{e}^{1/\lambda} \sum_{k:|\frac{k}{N}-\frac{1}{2}|\geq\varepsilon} \binom{N}{k} (p_N^*)^k (1-p_N^*)^{N-k}.$$

Since $u(\frac{k}{N}, \theta) \geq 0$ for all k, (2) gives a lower bound

$$Z_N(p_N^*, \theta) \geq \sum_{k=0}^{N} \binom{N}{k} (p_N^*)^k (1-p_N^*)^{N-k} = 1,$$

where we use the binomial theorem. By evaluating the right-hand side of (17) with these bounds,

$$\Pr\left(\left|\bar{a}_N^* - \frac{1}{2}\right| \geq \varepsilon \mid \theta\right) \leq \mathrm{e}^{1/\lambda} \underbrace{\sum_{k:|\frac{k}{N}-\frac{1}{2}|\geq\varepsilon} \binom{N}{k} (p_N^*)^k (1-p_N^*)^{N-k}}_{(18^*)}. \tag{18}$$

Next, we show that $(18^*) \to 0$ as $N \to \infty$. Using a random variable $B_N \sim \text{Binomial}(N, p_N^*)$, we rewrite $(18^*) = \Pr(|B_N/N - 1/2| > \varepsilon)$. By Lemma 3, for any $\varepsilon > 0$, there exists an N' such that $|p_N^* - 1/2| < \varepsilon/2$ for all $N \geq N'$. For such N, the triangle inequality gives

$$\left|\frac{B_N}{N} - \frac{1}{2}\right| \leq \left|\frac{B_N}{N} - p_N^*\right| + \left|p_N^* - \frac{1}{2}\right| < \left|\frac{B_N}{N} - p_N^*\right| + \frac{\varepsilon}{2}.$$

Hence,

$$(18^*) = \Pr\left(\left|\frac{B_N}{N} - \frac{1}{2}\right| > \varepsilon\right) \leq \Pr\left(\left|\frac{B_N}{N} - p_N^*\right| > \frac{\varepsilon}{2}\right).$$

Since B_N/N has the mean p_N^* and the variance $p_N^*(1-p_N^*)/N$, with $p_N^*(1-p_N^*) \leq 1/4$, Chebyshev's inequality gives

$$\Pr\left(\left|\frac{B_N}{N} - p_N^*\right| > \frac{\varepsilon}{2}\right) \leq \frac{1}{\varepsilon^2 N} \xrightarrow{N\to\infty} 0.$$

Hence, $(18^*) \to 0$ as desired. By (18), $\Pr(|\bar{a}_N^* - \frac{1}{2}| \geq \varepsilon \mid \theta) \to 0$. Since this holds for each θ, we have $\Pr(|\bar{a}_N^* - \frac{1}{2}| < \varepsilon) \to 1$.

A.6 Lemma 4

This lemma is immediate from the biased logit formula, shown by Matějka and McKay (2015, Theorem 1) and Caplin, Dean, and Leahy (2019, Proposition 1).

A.7 Lemma 5

As $N = 2n + 1 \to \infty$,

$$\Pi_N(\theta) = \binom{2n}{n}(Q_N^*(1 \mid \theta))^n (Q_N^*(0 \mid \theta))^n \leq \binom{2n}{n}\frac{1}{2^{2n}} \to 0.$$

By Lemma 4, $|Q_N^*(1 \mid 1) - q_N^*| \to 0$ and $|Q_N^*(1 \mid 0) - q_N^*| \to 0$.

Consider the case of $\theta = 1$, as the case of $\theta = 0$ is analogous. Fix any $\varepsilon > 0$. There exists an N_1 such that for all $N > N_1$, we have $|Q_N^*(1 \mid 1) - q_N^*| < \varepsilon/3$. Since $q_N^* \to q_\infty^*$ by assumption, there exists an N_2 such that for all $N > N_2$, we have $|q_N^* - q_\infty^*| < \varepsilon/3$. For any $N > \max\{N_1, N_2\}$,

$$|Q_N^*(1 \mid 1) - q_\infty^*| \leq |Q_N^*(1 \mid 1) - q_N^*| + |q_N^* - q_\infty^*| < \frac{2\varepsilon}{3}.$$

The equilibrium actions $b_1^*, b_2^*, \ldots$ are conditionally independent given state $\theta = 1$. By the law of large numbers, for any $\delta > 0$, there exists an N_3 such that for all $N > N_3$,

$$\Pr\left(|\bar{b}_N^* - Q_N^*(1 \mid 1)| < \frac{\varepsilon}{3} \mid \theta = 1\right) > 1 - \delta.$$

Since $|Q_N^*(1 \mid 1) - q_\infty^*| < 2\varepsilon/3$, we have $\Pr(|\bar{b}_N^* - q_\infty^*| < \varepsilon \mid \theta = 1) > 1 - \delta$ for all $N > \max\{N_1, N_2, N_3\}$.

A.8 Lemma 6

We focus on the case of $q_\infty^* = \frac{1}{2}$ since we have discussed the case of $q_\infty^* \neq \frac{1}{2}$ in the main text. For each N, we have $t_N^1, t_N^0 \in [-\frac{1}{2}, \frac{1}{2}]$ such that $Q_N^*(1 \mid 1) = \frac{1}{2} + t_N^1$ and $Q_N^*(1 \mid 0) = \frac{1}{2} - t_N^0$. We also have $t_N \in [-\frac{1}{2}, \frac{1}{2}]$ such that $q_N^* = \frac{1}{2} + t_N$. Since q_N^* is the marginal probability of choosing action 1 (Lemma 4), we have $q_N^* = \mu(1)Q_N^*(1 \mid 1) + \mu(0)Q_N^*(1 \mid 0)$. That is,

$$t_N = \mu(1)t_N^1 - \mu(0)t_N^0. \tag{19}$$

Note that $t_N^1 \to 0$, $t_N^0 \to 0$, and $t_N \to 0$ as $N \to \infty$ since $q_N^* \to \frac{1}{2}$ by assumption.

Step 1 The probability of correct choice is

$$\mu(1)\Pr\left(\bar{b}_N^* > \frac{1}{2} \mid \theta = 1\right) + \mu(0)\Pr\left(\bar{b}_N^* < \frac{1}{2} \mid \theta = 0\right). \tag{20}$$

Given state $\theta = 1$, the equilibrium actions $b_1^*, \ldots, b_N^*$ are i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\frac{1}{2}+t_N^1$ and $\frac{1}{2}-t_N^1$ respectively. We have the mean $\mu_N^1 \equiv \frac{1}{2}+t_N^1$ and the variance $(\sigma_N^1)^2 \equiv \frac{1}{4}-(t_N^1)^2$. Hence,

$$\Pr\left(\bar{b}_N^* > \frac{1}{2} \mid \theta = 1\right) = 1 - \Pr\left(\frac{\bar{b}_N^* - \mu_N^1}{\sigma_N^1/\sqrt{N}} \leq -\frac{\sqrt{N}t_N^1}{\sigma_N^1} \mid \theta = 1\right).$$

By Berry–Esseen theorem (Durrett, 2010, Theorem 3.4.9),

$$\left|\Pr\left(\frac{\bar{b}_N^* - \mu_N^1}{\sigma_N^1/\sqrt{N}} \leq -\frac{\sqrt{N}t_N^1}{\sigma_N^1} \mid \theta = 1\right) - \Phi\left(-\frac{\sqrt{N}t_N^1}{\sigma_N^1}\right)\right| \leq \frac{3\mathbb{E}[|b_i^* - \mu_N^1|^3 \mid \theta = 1]}{(\sigma_N^1)^3\sqrt{N}},$$

where Φ is the standard normal cdf. The right-hand side vanishes as $N \to \infty$ since $\sigma_N^1 \to \frac{1}{2}$ and $\mathbb{E}[|b_i^* - \mu_N^1|^3 \mid \theta = 1] \leq \frac{1}{8}$. Thus,

$$\lim_{N\to\infty}\left|\Pr\left(\frac{\bar{b}_N^* - \mu_N^1}{\sigma_N^1/\sqrt{N}} \leq -\frac{\sqrt{N}t_N^1}{\sigma_N^1} \mid \theta = 1\right) - \Phi\left(-\frac{\sqrt{N}t_N^1}{\sigma_N^1}\right)\right| = 0.$$

Hence,

$$\lim_{N\to\infty}\left|\Pr\left(\bar{b}_N^* > \frac{1}{2} \mid \theta = 1\right) - \Phi\left(\frac{\sqrt{N}t_N^1}{\sigma_N^1}\right)\right| = 0. \tag{21}$$

Given state $\theta = 0$, the equilibrium actions $b_1^*, \ldots, b_N^*$ are i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\frac{1}{2}-t_N^0$ and $\frac{1}{2}+t_N^0$ respectively. We have the mean $\mu_N^0 = \frac{1}{2}-t_N^0$ and the variance $(\sigma_N^0)^2 = \frac{1}{4}-(t_N^0)^2$. By the same argument as above,

$$\lim_{N\to\infty}\left|\Pr\left(\bar{b}_N^* < \frac{1}{2} \mid \theta = 0\right) - \Phi\left(\frac{\sqrt{N}t_N^0}{\sigma_N^0}\right)\right| = 0. \tag{22}$$

Step 2 Assume that $t_N^1 \neq 0$ and $t_N^0 \neq 0$. (The other cases are trivial, and we will discuss them in footnote 21). We rewrite (9) as

$$\lambda\left(\ln\frac{Q_N^*(1\mid 1)}{Q_N^*(0\mid 1)} - \ln\frac{q_N^*}{1-q_N^*}\right) = \binom{2n}{n}(Q_N^*(1\mid 1))^n(Q_N^*(0\mid 1))^n,$$
$$\lambda\left(\ln\frac{Q_N^*(0\mid 0)}{Q_N^*(1\mid 0)} - \ln\frac{1-q_N^*}{q_N^*}\right) = \binom{2n}{n}(Q_N^*(1\mid 0))^n(Q_N^*(0\mid 0))^n,$$

where we have $q_N^* \in (0,1)$ for large N, since $q_\infty^* = \frac{1}{2}$ by assumption. Recall that $Q_N^*(1\mid 1) = \frac{1}{2}+t_N^1$, $Q_N^*(1\mid 0) = \frac{1}{2}-t_N^0$, and $q_N^* = \frac{1}{2}+t_N$. Using the function $f : (-\frac{1}{2},\frac{1}{2}) \ni t \mapsto \ln((\frac{1}{2}+t)/(\frac{1}{2}-t)) \in \mathbb{R}$, we rewrite the above equations as

$$\lambda\Big(f(t_N^1) - f(t_N)\Big) = \binom{2n}{n}\left(\frac{1}{2}+t_N^1\right)^n\left(\frac{1}{2}-t_N^1\right)^n,$$

$$\lambda\Big(f(t_N^0)+f(t_N)\Big)=\binom{2n}{n}\left(\frac{1}{2}+t_N^0\right)^n\left(\frac{1}{2}-t_N^0\right)^n.$$

By the mean value theorem, there exists a τ_N^θ between 0 and t_N^θ such that $f(t_N^\theta)=t_N^\theta f'(\tau_N^\theta)$. Similarly, there exists a τ_N between 0 and t_N such that $f(t_N)=t_N f'(\tau_N)$. Since $f'(t)=(\frac{1}{4}-t^2)^{-1}$, we rearrange the terms to obtain that

$$\frac{\lambda}{2}\left(\frac{\sqrt{N}t_N^1}{\frac{1}{4}-(\tau_N^1)^2}-\frac{\sqrt{N}t_N}{\frac{1}{4}-(\tau_N)^2}\right)=\frac{1}{2}\sqrt{\frac{N}{n}}\binom{2n}{n}\frac{\sqrt{n}}{2^{2n}}\cdot\left(1-\frac{(2\sqrt{N}t_N^1)^2}{N}\right)^n, \tag{23}$$

$$\frac{\lambda}{2}\left(\frac{\sqrt{N}t_N^0}{\frac{1}{4}-(\tau_N^0)^2}+\frac{\sqrt{N}t_N}{\frac{1}{4}-(\tau_N)^2}\right)=\frac{1}{2}\sqrt{\frac{N}{n}}\binom{2n}{n}\frac{\sqrt{n}}{2^{2n}}\cdot\left(1-\frac{(2\sqrt{N}t_N^0)^2}{N}\right)^n. \tag{24}$$

Note that $\tau_N^\theta\to 0$ and $\tau_N\to 0$ as $N\to\infty$ since $t_N^\theta\to 0$ and $t_N\to 0$.

Step 3 Consider any subsequence of $\{N\}$ along which $\sqrt{N}t_N^1\to T^1\in\mathbb{R}$ and $\sqrt{N}t_N^0\to T^0\in\mathbb{R}$. (We examine other subsequences to Step 4.) By (19), $\sqrt{N}t_N\to T\in\mathbb{R}$ along the subsequence:

$$T=\mu(1)T^1-\mu(0)T^0. \tag{25}$$

Letting $N\to\infty$ along the subsequence, we derive, from (23) and (24), that

$$\begin{aligned}2\lambda(T^1-T)&=\phi(2T^1),\\ 2\lambda(T^0+T)&=\phi(2T^0),\end{aligned} \tag{26}$$

where ϕ is the standard normal pdf. Note that

$$\begin{aligned}&\text{the left-hand side of (23)}\to 2\lambda(T^1-T),\\ &\text{the left-hand side of (24)}\to 2\lambda(T^0+T).\end{aligned}$$

By Stirling's formula,

$$\lim_{N\to\infty}\frac{1}{2}\sqrt{\frac{N}{n}}\binom{2n}{n}\frac{\sqrt{n}}{2^{2n}}=\frac{1}{\sqrt{2\pi}},$$

where $N=2n+1$. For each $\theta=0,1$, we have[20]

$$\lim_{N\to\infty}\left(1-\frac{(2\sqrt{N}t_N^\theta)^2}{N}\right)^n=\exp\left(-\frac{(2T^\theta)^2}{2}\right).$$

Hence,

$$\text{the right-hand side of (23)}\to\phi(2T^1),$$

[20] If $\lim_{m\to\infty}c_m=c$ then $\lim_{m\to\infty}(1+c_m/m)^m=\mathrm{e}^c$ (Durrett, 2010, Theorem 3.4.2).

$$\text{the right-hand side of (24)} \to \phi(2T^0),$$

from which we obtain (26).

Substituting (25) into (26), we have

$$\begin{aligned}\lambda\mu(0)(2T^1+2T^0) &= \phi(2T^1),\\ \lambda\mu(1)(2T^1+2T^0) &= \phi(2T^0).\end{aligned} \tag{27}$$

By (21) and (22),

$$\begin{aligned}\lim_{N\to\infty} \Pr\left(\bar{b}_N^* > \frac{1}{2} \mid \theta = 1\right) &= \lim_{N\to\infty} \Phi\left(\frac{\sqrt{N}t_N^1}{\sigma_N^1}\right) = \Phi(2T^1),\\ \lim_{N\to\infty} \Pr\left(\bar{b}_N^* < \frac{1}{2} \mid \theta = 0\right) &= \lim_{N\to\infty} \Phi\left(\frac{\sqrt{N}t_N^0}{\sigma_N^0}\right) = \Phi(2T^0),\end{aligned}$$

where Φ is continuous, and $\sqrt{N}t_N^\theta \to T^\theta$ and $\sigma_N^\theta \to \frac{1}{2}$ for each $\theta = 0, 1$. Therefore, as $N \to \infty$, the probability of correct choice (20) converges to

$$\mu(1)\Phi(2T^1) + \mu(0)\Phi(2T^0).$$

Letting $t^1 = 2T^1$ and $t^0 = 2T^0$ here and in (27), we have the desired result.[21]

Step 4 Consider any subsequence along which at least one of $\{\sqrt{N}t_N^1\}$ and $\{\sqrt{N}t_N^0\}$ diverges. We have three cases to consider. First, suppose that one of $\{\sqrt{N}t_N^1\}$ and $\{\sqrt{N}t_N^0\}$ converges and the other diverges to $\pm\infty$. Without loss of generality, let $\sqrt{N}t_N^1 \to T_1 \in \mathbb{R}$ and $\sqrt{N}t_N^0 \to \pm\infty$. By (19), $\sqrt{N}t_N \to \mp\infty$. The left-hand side of (23) diverges to $\pm\infty$, but the right-hand side converges to a finite value. This is a contradiction. Second, suppose that both $\{\sqrt{N}t_N^1\}$ and $\{\sqrt{N}t_N^0\}$ diverge to $+\infty$. Then, $\{\sqrt{N}t_N\}$ converges or diverges. If it converges, the left-hand sides of (23) and (24) diverge to $+\infty$ but their right-hand sides converge to finite values; otherwise, the left-hand side of either (23) or (24) diverges but the right-hand sides of both (23) and (24) converge. This is a contradiction. Third, suppose that one of $\{\sqrt{N}t_N^1\}$ and $\{\sqrt{N}t_N^0\}$ diverges to $+\infty$ and the other diverges to $-\infty$. If $\sqrt{N}t_N^1 \to +\infty$ and $\sqrt{N}t_N^0 \to -\infty$ then $\Phi(\sqrt{N}t_N^1/\sigma_N^1) \to 1$ and $\Phi(\sqrt{N}t_N^0/\sigma_N^0) \to 0$ in (21) and (22), which implies that $\Pr(u(\bar{b}_N^*, \theta) = 1) \to \mu(1)$ in (20). Similarly, if $\sqrt{N}t_N^1 \to -\infty$ and $\sqrt{N}t_N^0 \to +\infty$ then $\Pr(u(\bar{b}_N^*, \theta) = 1) \to \mu(0)$.

[21] Steps 2 and 3 assume that $t_N^1 \neq 0$ and $t_N^0 \neq 0$. If $t_N^1 = 0$ for all sufficiently large N then (21) implies $\lim_N \Pr(\bar{b}_N^* > 1/2 \mid \theta = 1) = 1/2$; this case is covered in our result since $t^1 = 2\lim_N \sqrt{N}t_N^1 = 0$. If $t_N^0 = 0$ for all sufficiently large N then (22) implies $\lim_N \Pr(\bar{b}_N^* < 1/2 \mid \theta = 0) = 1/2$; this case is also covered in our result. Moreover, consider the case where $t_N^\theta = 0$ for infinitely many N but for any $\bar{N}$, there exists an $N > \bar{N}$ such that $t_N^\theta \neq 0$. Then, we can use the same argument as in the proof by taking a subsequence along which for all N, $t_N^\theta \neq 0$.

A.9 Proposition 1

In an election with a poll, we have a unique informative equilibrium (Lemma 2) since for any λ, Condition 1 holds for any small enough $\varepsilon > 0$ and any μ such that $|\mu(1) - \frac{1}{2}| < \varepsilon$; moreover, the probability of correct choice is $\mathsf{e}^{1/\lambda}/(1 + \mathsf{e}^{1/\lambda})$ at the informative equilibrium (Theorem 1). In an election without a poll, the probability of correct choice at any (informative or uninformative) equilibrium converges to one of $\mu(1)$, $\mu(0)$, or $\mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)$ as $N \to \infty$ (Lemma 6).

It suffices to show that for any small $\varepsilon > 0$ and for any μ and λ such that $|\mu(1) - \frac{1}{2}| < \varepsilon$,

$$\frac{\mathsf{e}^{1/\lambda}}{1 + \mathsf{e}^{1/\lambda}} > \max\Big\{\mu(1), \mu(0), \mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)\Big\}, \tag{28}$$

where

$$\lambda\mu(0)(t^1 + t^0) = \phi(t^1) \quad \text{and} \quad \lambda\mu(1)(t^1 + t^0) = \phi(t^0).$$

Since t^1 and t^0 are continuous in μ, so is $\mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)$. Hence, it suffices to prove (28) for the prior $\mu(1) = \frac{1}{2}$.

Lemma B. *For each N, any election $\mathcal{Q}_N$ with the prior $\mu(1) = \frac{1}{2}$ has a symmetric equilibrium Q^*_N such that $q^*_N = \frac{1}{2}$. For the sequence of these equilibria $\{Q^*_N\}_N$, the equilibrium probability of correct choice converges to either $\frac{1}{2}$ or $\Phi(t) > \frac{1}{2}$, where $t > 0$ is a unique solution to equation $\lambda t = \phi(t)$.*

Proof. Under $\mu(1) = \frac{1}{2}$, we have a symmetric equilibrium $Q^*_N(1 \mid 1) = Q^*_N(0 \mid 0)$ by Lemma 4. Then, $q^*_N = \mu(1)Q^*_N(1 \mid 1) + \mu(0)Q^*_N(1 \mid 0) = \frac{1}{2}$.

The latter half of this lemma follows from Lemma 6. We use the same notation as in the proof of Lemma 6. By symmetry, $Q^*_N(1 \mid 1) = Q^*_N(0 \mid 0) > \frac{1}{2}$ and thus $t^1_N = t^0_N > 0$. Hence, $\lim_N \sqrt{N}t^1_N = \lim_N \sqrt{N}t^0_N$. We have $t \equiv t^1 = t^2$ in Lemma 6 because $t^1 = 2\lim_N \sqrt{N}t^1_N$ and $t^2 = 2\lim_N \sqrt{N}t^0_N$ (as defined at the end of Step 3 in the proof). Substituting it into Lemma 6, we have $\mu(1)\Phi(t^1) + \mu(0)\Phi(t^0) = \Phi(t)$, where $\lambda t = \phi(t)$. ■

By Lemma B, we rewrite (28) under $\mu(1) = \frac{1}{2}$ as

$$\frac{\mathsf{e}^{1/\lambda}}{1 + \mathsf{e}^{1/\lambda}} > \Phi(t),$$

where $t > 0$ is a unique solution to $\lambda t = \phi(t)$. That is, t is defined by the implicit function of λ. Since this relation is bijective, by reparametrizing $1/\lambda = t/\phi(t) = \sqrt{2\pi}t\mathsf{e}^{t^2/2}$ and rearranging the terms, we rewrite this inequality as

$$\sqrt{2\pi}t\mathsf{e}^{t^2/2} - \ln\frac{\Phi(t)}{1 - \Phi(t)} > 0. \tag{29}$$

By the fundamental theorem of calculus,

$$\ln \frac{\Phi(t)}{1-\Phi(t)} = \int_0^t \left[\frac{\phi(s)}{\Phi(s)} + \frac{\phi(s)}{1-\Phi(s)}\right] \mathrm{d}s.$$

where $\ln(\Phi(0)/(1-\Phi(0)) = 0$ for $\Phi(0) = \frac{1}{2}$. For all $s > 0$, we have $\phi(s) < \phi(0) = 1/\sqrt{2\pi}$ and $\Phi(s) > \Phi(0) = \frac{1}{2}$, and thus $\phi(s)/\Phi(s) < \sqrt{2/\pi}$. The inverse Mills ratio $\phi(s)/(1-\Phi(s))$ is bounded above by $s + \sqrt{2/\pi}$.[22] By evaluating the integral with these bounds, we have, for all $t > 0$,

$$\ln \frac{\Phi(t)}{1-\Phi(t)} < \frac{t^2}{2} + 2\sqrt{\frac{2}{\pi}}t.$$

To prove (29), it suffices to show that $[\sqrt{2\pi}\mathrm{e}^{t^2/2} - (t/2 + 2\sqrt{2/\pi})]t > 0$ for all $t > 0$. Since $\mathrm{e}^{t^2/2} > 1 + t^2/2$ for all $t > 0$, we only need to show that $\sqrt{2\pi}(1 + t^2/2) - (t/2 + 2\sqrt{2/\pi}) > 0$, which is easily verified since the left-hand side is quadratic.

References

M. Agranov, J. K. Goeree, J. Romero, and L. Yariv. What makes voters turn out: The effects of polls and beliefs. *Journal of European Economic Association*, 16(3):825–856, 2018.

D. Austen-Smith and J. S. Banks. Information aggregation, rationality, and the Condorcet jury theorem. *American Political Science Review*, 90(1):34–45, 1996.

D. Bergemann and S. Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.

T. Bergstrom, L. Blume, and H. Varian. On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49, 1986.

S. Bhattacharya, J. Duffy, and S. Kim. Voting with endogenous information acquisition: Experimental evidence. *Games and Economic Behavior*, 102:316–338, 2017.

D. M. Butler. A regression discontinuity design analysis of the incumbency advantage and tenure in the us house. *Electoral Studies*, 28(1):123–128, 2009.

A. Caplin, M. Dean, and J. Leahy. Rational inattention, optimal consideration sets, and stochastic choice. *Review of Economic Studies*, 86(3):1061–1094, 2019.

D. Caughey and J. S. Sekhon. Elections and the regression discontinuity design: Lessons from close US House races, 1942–2008. *Political Analysis*, 19(4):385–408, 2011.

M. d. Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. L'Imprimerie Royale, 1785.

P. J. Coughlan. In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *American Political Science Review*, 94(2):375–393, 2000.

[22] To see this bound, we write $M(s) = \phi(s)/(1-\Phi(s))$ for the inverse Mills ratio. Then, $M(0) = \sqrt{2/\pi}$ and $M'(s) \leq 1$ for all $s \geq 0$. By the mean value theorem, for all $s > 0$, there exists a $c_s \in (0, s)$ such that $M(s) = M'(c_s)s + M(0)$. Hence, $M(s) < s + \sqrt{2/\pi}$.

T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley-Interscience, 2006.

B. de la Cuesta and K. Imai. Misunderstandings about the regression discontinuity design in the study of close elections. *Annual Review of Political Science*, 19:375–396, 2016.

L. De Magalhães, D. Hangartner, S. Hirvonen, J. Meriläinen, N. A. Ruiz, and J. Tukiainen. When can we trust regression discontinuity design estimates from close elections? evidence from experimental benchmarks. *Political Analysis*, pages 1–8, 2025.

H. de Oliveira. Axiomatic foundations for entropic costs of attention. 2019.

M. Dell. Trafficking networks and the Mexican drug war. *American Economic Review*, 105(6): 1738–1779, 2015.

T. Denti. Unrestricted information acquisition. *Theoretical Economics*, 18(3):1101–1140, 2023.

A. Downs. An economic theory of democracy. 1957.

R. Durrett. *Probability: Theory and examples (the 4th edition).* Cambridge University Press, 2010.

A. C. Eggers, A. Fowler, J. Hainmueller, A. B. Hall, and J. M. Snyder. On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science*, 59(1):259–274, 2015.

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

A. Elbittar, A. Gomberg, C. Martinelli, and T. R. Palfrey. Ignorance and bias in collective decisions. *Journal of Economic Behavior & Organization*, 174:332–359, 2020.

T. Feddersen and W. Pesendorfer. Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review*, 92(1):23–35, 1998.

T. J. Feddersen and W. Pesendorfer. The swing voter's curse. *American Economic Review*, 86(3): 408–424, 1996.

F. Ferreira and J. Gyourko. Do political parties matter? Evidence from US cities. *Quarterly Journal of Economics*, 124(1):399–422, 2009.

F. Ferreira and J. Gyourko. Does gender matter for political leadership? The case of US mayors. *Journal of Public Economics*, 112:24–39, 2014.

S. Firpo, V. Ponczek, and V. Sanfelice. The relationship between federal budget amendments and local electoral power. *Journal of Development Economics*, 116:186–198, 2015.

D. Gerardi and L. Yariv. Deliberative voting. *Journal of Economic Theory*, 134(1):317–338, 2007.

D. Gerardi and L. Yariv. Information acquisition in committees. *Games and Economic Behavior*, 62(2):436–459, 2008.

S. J. Grossman and J. E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, 1980.

J. Hahn, P. Todd, and W. Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209, 2001.

B. M. Hébert and J. La'O. Information acquisition, efficiency, and non-fundamental volatility. *Journal of Political Economy*, 131:2666–2723, 2023.

T. Hoshino. Rational inattention in games. *Working Paper*, 2018.

T. Hoshino and T. Ui. Rational inattention and endogenous volatility: A large deviation approach. *Working Paper*, 2024.

A. Hyytinen, J. Meriläinen, T. Saarimaa, O. Toivanen, and J. Tukiainen. When does regression discontinuity design work? evidence from random election outcomes. *Quantitative Economics*, 9 (2):1019–1051, 2018.

Y. Koriyama and B. Szentes. A resurrection of the Condorcet Jury Theorem. *Theoretical Economics*, 4(2):227–252, 2009.

D. S. Lee. The electoral advantage to incumbency and voters' valuation of politicians' experience: A regression discontinuity analysis of elections to the U.S. House. *National Bureau of Economic Research*, 2001.

D. S. Lee. Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, 142(2):675–697, 2008.

A. Li and L. Hu. Electoral accountability and selection with personalized information aggregation. *Games and Economic Behavior*, 140:296–315, 2023.

C. Martinelli. Would rational voters acquire costly information? *Journal of Economic Theory*, 129: 225–51, 2006.

C. Martinelli. Rational ignorance and voting behavior. *International Journal of Game Theory*, 35 (3):315–335, 2007.

F. Matějka and A. McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98, 2015.

F. Matějka and G. Tabellini. Electoral competition with rationally inattentive voters. *Journal of European Economic Association*, 19(3):1899–1935, 2021.

D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14(1):124–143, 1996.

K. Mukhopadhaya. Jury size and the free rider problem. *Journal of Law, Economics, and Organization*, 19(1):24–44, 2003.

S. Oliveros. Abstention, ideology and information acquisition. *Journal of Economic Theory*, 148 (3):871–902, 2013.

N. Persico. Committee design with endogenous information. *Review of Economic Studies*, 71(1): 165–191, 2004.

C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.

B. Sinclair and C. R. Plott. From uninformed to informed choices: Voters, pre-election polls and updating. *Electoral Studies*, 31(1):83–95, 2012.

J. Snyder. Detecting manipulation in U.S. House elections. *Working Paper*, 2005.

C. R. Taylor and H. Yildirim. Public information and electoral bias. *Games and Economic Behavior*, 68(1):353–375, 2010.

D. L. Thistlethwaite and D. T. Campbell. Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, 51(6):309, 1960.

M. Triossi. Costly information acquisition. is it better to toss a coin? *Games and Economic Behavior*, 82:169–191, 2013.

S. Yuksel. Specialized learning and political polarization. *International Economic Review*, 63(1): 457–474, 2022.

Online Appendix

In this online appendix, we give omitted proofs for Theorems 1′ and 2′ and related lemmas.

B Proofs of Theorems 1′ and 2′

We extend all results in Section 2 (Lemmas 1 to 3 and Theorems 1 and 2) to elections under supermajority rule. The proofs are analogous to the original proofs, but not the same.

Preliminaries Let u_α be the payoff function under supermajority rule. Formally, we define each voter's payoff function $u_\alpha : [0,1] \times \Theta \to \{0,1\}$ by, given a threshold $\alpha \in (\frac{1}{2}, 1]$,

$$u_\alpha(\bar{a}_N, \theta) = \begin{cases} 1 & \text{if } \mathbb{1}\{\bar{a}_N \geq \alpha\} = \theta \\ 0 & \text{if } \mathbb{1}\{\bar{a}_N \geq \alpha\} \neq \theta, \end{cases}$$

where the chosen alternative is denoted by $\mathbb{1}\{\bar{a}_N \geq \alpha\}$.

Let n_α be the integer such that alternative 1 is chosen if and only if it receives at least $n_\alpha + 1$ votes. That is, $n_\alpha = k$ for the unique integer k such that $\frac{k}{N} < \alpha \leq \frac{k+1}{N}$.[23] In particular, $n_1 = N - 1$ for the unanimity rule ($\alpha = 1$).

B.1 Lemma 1′

Lemma 1′. *In any election $\mathcal{P}_{N,\alpha}$, every symmetric equilibrium $P^*_{N,\alpha}$ has some $p^*_{N,\alpha} \in [0,1]$ such that for each θ and each $k = 0, 1, \ldots, N$, the equilibrium vote share $\bar{a}_N$ satisfies*

$$\Pr\left(\bar{a}^*_N = \frac{k}{N} \mid \theta\right) = \frac{1}{Z_{N,\alpha}(p^*_{N,\alpha}, \theta)} \binom{N}{k} \exp\left(\frac{u_\alpha(\frac{k}{N}, \theta)}{\lambda}\right) \left(p^*_{N,\alpha}\right)^k \left(1 - p^*_{N,\alpha}\right)^{N-k}, \tag{30}$$

where $Z_{N,\alpha} : [0,1] \times \Theta \to \mathbb{R}$ is the function defined by

$$Z_{N,\alpha}(p, \theta) = \sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u_\alpha(\frac{k}{N}, \theta)}{\lambda}\right) p^k (1-p)^{N-k}, \tag{31}$$

and $p^*_{N,\alpha}$ is the marginal probability of each voter choosing action 1.

The following properties hold:

1. *$P^*_{N,\alpha}$ is an uninformative equilibrium if and only if $p^*_{N,\alpha} \in \{0,1\}$.*
2. *$P^*_{N,\alpha}$ is an informative equilibrium if and only if $p^*_{N,\alpha} \in (0,1)$ is a solution to equation*

$$\frac{Z_{N,\alpha}(p, 1)}{Z_{N,\alpha}(p, 0)} = \frac{\mu(1)}{\mu(0)}. \tag{32}$$

[23] Under simple majority rule, we have $n_{1/2} = n$, where there are $N = 2n + 1$ voters.

Proof. The proof of Lemma 1 goes through up to deriving the first-order condition (11) except u is replaced by u_α.[24] Here is the modified first-order condition:

$$\sum_\theta \mu(\theta) \cdot \frac{\sum_{k=0}^{2n} \binom{2n}{k} \left[\exp\left(\frac{u_\alpha(\frac{k+1}{N}, \theta)}{\lambda} \right) - \exp\left(\frac{u_\alpha(\frac{k}{N}, \theta)}{\lambda} \right) \right] \left(p^*_{N,\alpha}\right)^k \left(1 - p^*_{N,\alpha}\right)^{2n-k}}{\underbrace{\sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u_\alpha(\frac{k}{N}, \theta)}{\lambda} \right) \left(p^*_{N,\alpha}\right)^k \left(1 - p^*_{N,\alpha}\right)^{N-k}}_{= Z_{N,\alpha}(p^*_{N,\alpha}, \theta) \text{ by (31)}}} = 0.$$

In the numerator, if $k \neq n_\alpha$, the square bracket is zero, while if $k = n_\alpha$ then the square bracket is $\mathsf{e}^{1/\lambda} - 1$ when $\theta = 1$ and $1 - \mathsf{e}^{1/\lambda}$ when $\theta = 0$. By substituting them into the above equation,

$$\begin{aligned} &\frac{\mu(1)}{Z_{N,\alpha}(p^*_{N,\alpha}, 1)} \binom{2n}{n_\alpha} (p^*_{N,\alpha})^{n_\alpha} (1 - p^*_{N,\alpha})^{2n - n_\alpha} (\mathsf{e}^{1/\lambda} - 1) \\ &+ \frac{\mu(0)}{Z_{N,\alpha}(p^*_{N,\alpha}, 0)} \binom{2n}{n_\alpha} (p^*_{N,\alpha})^{n_\alpha} (1 - p^*_{N,\alpha})^{2n - n_\alpha} (1 - \mathsf{e}^{1/\lambda}) = 0. \end{aligned}$$

Rearranging the terms, we obtain (32). Hence, $(p^*_{N,\alpha}, \ldots, p^*_{N,\alpha})$ is a Nash equilibrium if and only if $p^*_{N,\alpha}$ is a solution to (32). ∎

B.2 Lemma 2′

Lemma 2′. *An election $\mathcal{P}_{N,\alpha}$ has an informative equilibrium if and only if it satisfies Condition 1. The informative equilibrium is unique whenever it exists.*

Proof. The proof of Lemma 2 goes through with small modifications. We define the function $W_{N,\alpha} : [0,1] \times \Theta \to \mathbb{R}$ by

$$\begin{aligned} W_{N,\alpha}(p, 1) &= \sum_{k=n_\alpha+1}^{N} \binom{N}{k} p^k (1-p)^{N-k}, \\ W_{N,\alpha}(p, 0) &= \sum_{k=0}^{n_\alpha} \binom{N}{k} p^k (1-p)^{N-k}. \end{aligned} \tag{33}$$

Note that $W_{N,\alpha}(p, 1) + W_{N,\alpha}(p, 0) = 1$ by the binomial theorem. Note that $W_{N,\alpha}(p, 1)$ is strictly increasing in p and $W_{N,\alpha}(p, 0)$ is strictly decreasing in p, as can be shown by differentiation. Then, we rewrite $Z_{N,\alpha}$, as defined in (31), as

$$\begin{aligned} Z_{N,\alpha}(p, 1) &= W_{N,\alpha}(p, 0) + \mathsf{e}^{1/\lambda} W_{N,\alpha}(p, 1), \\ Z_{N,\alpha}(p, 0) &= \mathsf{e}^{1/\lambda} W_{N,\alpha}(p, 0) + W_{N,\alpha}(p, 1). \end{aligned} \tag{34}$$

Note that $\frac{Z_{N,\alpha}(p,1)}{Z_{N,\alpha}(p,0)}$ is continuous and strictly increasing in p, because $Z_{N,\alpha}(p, 1)$ is strictly increasing

[24] Lemma A remains true except u is replaced by u_α.

in p and $Z_{N,\alpha}(p, 0)$ strictly decreasing in p.

The remaining argument is the same in the original proof except W_N and Z_N are replaced by $W_{N,\alpha}$ and $Z_{N,\alpha}$, respectively. ■

B.3 Proof of Theorem 1′

This proof is analogous to the proof of Theorem 1. We imitate the same argument but replace the winning threshold $\frac{1}{2}$ with α, the integer n with n_α, the function u with u_α, and W_N with $W_{N,\alpha}$. Then,

$$\Pr(\bar{a}_N^* \geq \alpha \mid \theta = 1) = \frac{\mathsf{e}^{1/\lambda} - \frac{\mu(0)}{\mu(1)}}{\mathsf{e}^{1/\lambda} - \mathsf{e}^{-1/\lambda}},$$
$$\Pr(\bar{a}_N^* < \alpha \mid \theta = 0) = \frac{\mathsf{e}^{1/\lambda} - \frac{\mu(1)}{\mu(0)}}{\mathsf{e}^{1/\lambda} - \mathsf{e}^{-1/\lambda}}.$$

Hence,

$$\Pr(u_\alpha(\bar{a}_N^*, \theta) = 1) = \mu(1) \Pr(\bar{a}_N^* \geq \alpha \mid \theta = 1) + \mu(0) \Pr(\bar{a}_N^* < \alpha \mid \theta = 1) = \frac{\mathsf{e}^{1/\lambda}}{1 + \mathsf{e}^{1/\lambda}}.$$

The proof that $\mathsf{e}^{1/\lambda}/(1+\mathsf{e}^{1/\lambda}) > \max\{\mu(1), \mu(0)\}$ when the informative equilibrium exists is exactly the same as in the original proof.

B.4 Lemma 3′

Lemma 3′. *For any election $\mathcal{P}_{N,\alpha}$ that satisfies Condition 1, let $P_{N,\alpha}^*$ be the informative equilibrium and $p_{N,\alpha}^*$ be the marginal probability of each voter choosing action 1. Then,*

$$\lim_{N \to \infty} p_{N,\alpha}^* = \alpha.$$

Proof. First, we consider a winning threshold $\alpha \in (\frac{1}{2}, 1)$. It suffices to show that for any small $\varepsilon > 0$ such that $0 < \alpha - \varepsilon < \alpha + \varepsilon < 1$, if we have a sufficiently large N then

$$\frac{Z_{N,\alpha}(\alpha - \varepsilon, 1)}{Z_{N,\alpha}(\alpha - \varepsilon, 0)} < \frac{\mu(1)}{\mu(0)} < \frac{Z_{N,\alpha}(\alpha + \varepsilon, 1)}{Z_{N,\alpha}(\alpha + \varepsilon, 0)}. \tag{35}$$

To see this sufficiency, note that since $\frac{Z_{N,\alpha}(p,1)}{Z_{N,\alpha}(p,0)}$ is continuous and strictly increasing in p, if (35) is true then $p_{N,\alpha}^* \in (\alpha - \varepsilon, \alpha + \varepsilon)$, where $p_{N,\alpha}^*$ is a solution to (32).

We show auxiliary inequalities. For any small $\delta > 0$, there is an N_δ such that for any $N > N_\delta$,

$$\begin{aligned} W_{N,\alpha}(\alpha + \varepsilon, 1) &> 1 - \delta, \quad W_{N,\alpha}(\alpha + \varepsilon, 0) < \delta, \\ W_{N,\alpha}(\alpha - \varepsilon, 0) &> 1 - \delta, \quad W_{N,\alpha}(\alpha - \varepsilon, 1) < \delta, \end{aligned} \tag{36}$$

where $W_{N,\alpha}$ is defined in (33). To see these inequalities, let $w_1, \ldots, w_N$ be i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\alpha+\varepsilon$ and $1-\alpha-\varepsilon$ respectively. Then, $W_{N,\alpha}(\alpha+\varepsilon,1)$ and $W_{N,\alpha}(\alpha+\varepsilon,0)$ are the probabilities that the sample average $\frac{1}{N}\sum_{i=1}^{N} w_i$ is, respectively, strictly greater than α and strictly less than α. By the law of large numbers, there is an N'_δ such that for any $N > N'_\delta$, we have $W_{N,\alpha}(\alpha+\varepsilon,1) > 1-\delta$ and $W_{N,\alpha}(\alpha+\varepsilon,0) < \delta$. To see the other two inequalities, let $w'_1,\dots,w'_N$ be i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\alpha-\varepsilon$ and $1-\alpha+\varepsilon$ respectively. By the same argument, there is an N''_δ such that for any $N > N''_\delta$, we have $W_N(\alpha-\varepsilon,0) > 1-\delta$ and $W_N(\alpha-\varepsilon,1) < \delta$. Lastly, let $N_\delta = \max\{N'_\delta, N''_\delta\}$.

Now we prove (35). This step is the same as in the original proof except the functions W_N and Z_N are replaced by $W_{N,\alpha}$ and $Z_{N,\alpha}$ respectively and the winning threshold $\frac{1}{2}$ is replaced by α.

Second, we consider the winning threshold $\alpha = 1$. It suffices to show that for any small $\varepsilon > 0$, if N is sufficiently large,

$$\frac{Z_{N,1}(1-\varepsilon,1)}{Z_{N,1}(1-\varepsilon,0)} < \frac{\mu(1)}{\mu(0)}. \tag{37}$$

To see this sufficiency, note that since $\frac{Z_{N,1}(p,1)}{Z_{N,1}(p,0)}$ is continuous and strictly increasing in p, if (37) is true then $p^*_{N,1} > 1-\varepsilon$, where $p^*_{N,1}$ is a solution to (32).

We show auxiliary inequalities. For any $\delta > 0$, there is an N_δ such that for any $N > N_\delta$,

$$W_{N,1}(1-\varepsilon,1) < \delta, \quad W_{N,1}(1-\varepsilon,0) > 1-\delta, \tag{38}$$

where $W_{N,1}$ is defined in (33). To see these inequalities, let $w_1,\dots,w_N$ be i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $1-\varepsilon$ and ε respectively. Then, $W_{N,1}(1-\varepsilon,1)$ and $W_{N,1}(1-\varepsilon,0)$ are the probabilities that the sample average $\frac{1}{N}\sum_{i=1}^{N} w_i$ is, respectively, equal to 1 and strictly less than 1. By the law of large numbers, there is an N_δ such that for any $N > N_\delta$, we have $W_{N,1}(1-\varepsilon,1) < \delta$ and $W_{N,1}(1-\varepsilon,0) > 1-\delta$.

We show another inequality. Under Condition 1, there is a small $\delta > 0$ such that

$$\frac{1+\mathsf{e}^{1/\lambda}\delta}{\mathsf{e}^{1/\lambda}(1-\delta)} < \frac{\mu(1)}{\mu(0)}. \tag{39}$$

To see this inequality, note that $\mathsf{e}^{-1/\lambda} < \frac{\mu(1)}{\mu(0)}$ (Condition 1) and that for a small enough δ, the left-hand side of (39) is arbitrarily close to $\mathsf{e}^{-1/\lambda}$.

Now we prove (37). For any $N > N_\delta$,

$$\frac{Z_{N,1}(1-\varepsilon,1)}{Z_{N,1}(1-\varepsilon,0)} = \frac{W_{N,1}(1-\varepsilon,0)+\mathsf{e}^{1/\lambda}W_{N,1}(1-\varepsilon,1)}{\mathsf{e}^{1/\lambda}W_{N,1}(1-\varepsilon,0)+W_{N,1}(1-\varepsilon,1)} < \frac{1+\mathsf{e}^{1/\lambda}\delta}{\mathsf{e}^{1/\lambda}(1-\delta)} < \frac{\mu(1)}{\mu(0)},$$

where we use (34) for the equality, (38) for the first inequality, and (39) for the second one. Hence, we have (37), which completes the proof. ■

B.5 Proof of Theorem 2′

Fix any $\theta \in \Theta$ and any $\varepsilon > 0$. By Lemma 1′,

$$\Pr(|\bar{a}_N^* - \alpha| \geq \varepsilon \mid \theta) = \frac{1}{Z_{N,\alpha}(p_N^*, \theta)} \underbrace{\sum_{k:|\frac{k}{N}-\alpha|\geq\varepsilon} \binom{N}{k} \exp\left(\frac{u_\alpha(\frac{k}{N}, \theta)}{\lambda}\right) \left(p_{N,\alpha}^*\right)^k \left(1 - p_{N,\alpha}^*\right)^{N-k}}_{(40^*)}, \quad (40)$$

where the sum runs over all $k = 0, 1, \ldots, N$ such that $|\frac{k}{N} - \alpha| \geq \varepsilon$. Since $u_\alpha(\frac{k}{N}, \theta) \leq 1$ for all k,

$$(40^*) \leq \mathrm{e}^{1/\lambda} \sum_{k:|\frac{k}{N}-\alpha|\geq\varepsilon} \binom{N}{k} \left(p_{N,\alpha}^*\right)^k \left(1 - p_{N,\alpha}^*\right)^{N-k}.$$

Since $u_\alpha(\frac{k}{N}, \theta) \geq 0$ for all k, (31) gives a lower bound

$$Z_{N,\alpha}(p_{N,\alpha}^*, \theta) \geq \sum_{k=0}^{N} \binom{N}{k} (p_N^*)^k (1 - p_N^*)^{N-k} = 1,$$

where we use the binomial theorem. By evaluating the right-hand side of (40) with these bounds,

$$\Pr(|\bar{a}_N^* - \alpha| \geq \varepsilon \mid \theta) \leq \mathrm{e}^{1/\lambda} \underbrace{\sum_{k:|\frac{k}{N}-\alpha|\geq\varepsilon} \binom{N}{k} \left(p_{N,\alpha}^*\right)^k \left(1 - p_{N,\alpha}^*\right)^{N-k}}_{(41^*)}. \quad (41)$$

Next, we show that $(41^*) \to 0$ as $N \to \infty$. Using a random variable $B_{N,\alpha} \sim \mathrm{Binomial}(N, p_{N,\alpha}^*)$, we rewrite $(41^*) = \Pr(|B_{N,\alpha}/N - \alpha| > \varepsilon)$. By Lemma 3′, for any $\varepsilon > 0$, there exists an N' such that $|p_{N,\alpha}^* - \alpha| < \varepsilon/2$ for all $N \geq N'$. For such N, the triangle inequality gives

$$\left|\frac{B_{N,\alpha}}{N} - \alpha\right| \leq \left|\frac{B_{N,\alpha}}{N} - p_{N,\alpha}^*\right| + \left|p_{N,\alpha}^* - \alpha\right| < \left|\frac{B_{N,\alpha}}{N} - p_{N,\alpha}^*\right| + \frac{\varepsilon}{2}.$$

Hence,

$$\Pr\left(\left|\frac{B_{N,\alpha}}{N} - \alpha\right| > \varepsilon\right) \leq \Pr\left(\left|\frac{B_{N,\alpha}}{N} - p_{N,\alpha}^*\right| > \frac{\varepsilon}{2}\right).$$

Since $B_{N,\alpha}/N$ has the mean $p_{N,\alpha}^*$ and the variance $p_{N,\alpha}^*(1 - p_{N,\alpha}^*)/N$, with $p_{N,\alpha}^*(1 - p_{N,\alpha}^*) \leq 1/4$, Chebyshev's inequality gives

$$\Pr\left(\left|\frac{B_{N,\alpha}}{N} - p_{N,\alpha}^*\right| > \frac{\varepsilon}{2}\right) \leq \frac{1}{\varepsilon^2 N} \xrightarrow{N\to\infty} 0.$$

Hence, $(41^*) \to 0$ as desired. By (41), $\Pr(|\bar{a}_N^* - \alpha| \geq \varepsilon \mid \theta) \to 0$. Since this holds for each θ, we have $\Pr(|\bar{a}_N^* - \alpha| < \varepsilon) \to 1$.