Elections with Opinion Polls: Information Acquisition and Aggregation*

Andrei Gomberg† Tetsuya Hoshino‡

September 26, 2024

Abstract

We study common-value elections with opinion polls. A poll contains information about the likelihood of a close election. Rationally inattentive voters acquire information about alternatives and polls at a cost proportional to expected entropy reduction. Every election has a unique symmetric equilibrium with information acquisition, and the probability of making the correct choice is independent of the electorate size and voters' prior. However, elections become closer as the electorate size grows. Additionally, elections with polls are more likely to make the correct choice than those without. Our findings offer novel implications for the regression discontinuity design in close elections.

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

JEL codes: D72, D82, D83

*We would like to thank Yu Awaya, Gadi Barlevy, John Duffy, John Duggan, Ruben Enikolopov, Srihari Govindan, Emilio Gutierrez, Chiaki Hara, Horacio Larreguy, Shaun McRae, César Martinelli, Tetsuya Matsubayashi, Jaakko Meriläinen, Romans Pancs, Mauricio Romero, Yuliy Sannikov, Katsutoshi Wakai, Xinyang Wang, and participants of various seminars and conferences for helpful comments. All remaining errors are our own.

†ITAM. Email: gomberg@itam.mx

‡ITAM. Email: tetsuyahoshino137@gmail.com

1 Introduction

The literature on information aggregation through voting presents a fundamental tension between two perspectives. Condorcet's (1785) Jury Theorem offers an optimistic view: when voters share common interests but possess dispersed information about alternatives, majority voting can effectively choose the commonly preferred—or correct—alternative. However, this optimism is challenged by Downs's (1957) **rational ignorance hypothesis**, which posits that voters must acquire information at a cost. Consequently, voters may rationally choose to remain uninformed, especially in large electorates where an individual vote is unlikely to be pivotal. This gives rise to a free-rider problem: while it would be socially beneficial for voters to be well-informed, it is not individually rational for them to bear the costs of information acquisition when the benefits are shared by all. As a result, there may be an under-provision of political information. This insight casts doubt on whether a large electorate will choose the correct alternative and has spurred research on voters' information acquisition and pivotality (e.g., Persico, 2004; Martinelli, 2006, 2007; Koriyama and Szentes, 2009; Oliveros, 2013; Triossi, 2013).

We argue that the problem of rational ignorance may be alleviated by public **opinion polls**, hereafter referred to as polls, that provide information about voting intentions. Polls facilitate voters' access to information about their potential pivotality. If a poll indicates a close election, individual voters might perceive a higher probability of being pivotal, which could encourage them to acquire more information. Conversely, a poll suggesting a one-sided election may discourage information acquisition. This mechanism implies that polls could facilitate voters' coordination in acquiring information, thereby alleviating the free-rider problem.

Our goal is to examine the role of polls, focusing on their impact on voters' information acquisition and voting behavior, as well as on the probability that the electorate chooses the correct alternative. We seek to elucidate how polls may influence information aggregation and mitigate the free-rider problem in large electorates.

Model Consider a common-value election with simple majority rule. There are two alternatives, 1 and 0. There are N voters, with N being odd, each voting for either alternative. Each voter receives a payoff of 1 if the majority vote chooses the "correct" alternative and a payoff of 0 otherwise. The correct alternative is modeled by a state θ, a random variable that takes values 1 or 0 according to a common prior μ.

Voters may acquire costly information about the state and about a poll that indicates a probable vote share among the two alternatives. We model information acquisition using the framework of **rational inattention** (Sims, 2003). In this framework, voters can acquire any costly signals that may correlate with the state and the poll. This feature allows them to pay attention to the state if the poll indicates a close race, but ignore it otherwise. Following Sims (2003), we assume that the cost of information is proportional to the expected reduction in uncertainty, measured in terms of entropy. Each voter maximizes her expected payoff from the collective decision, minus the private cost of information.

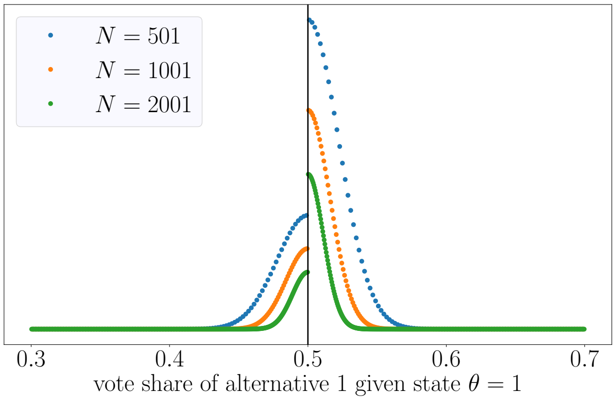


Figure 1: with opinion polls

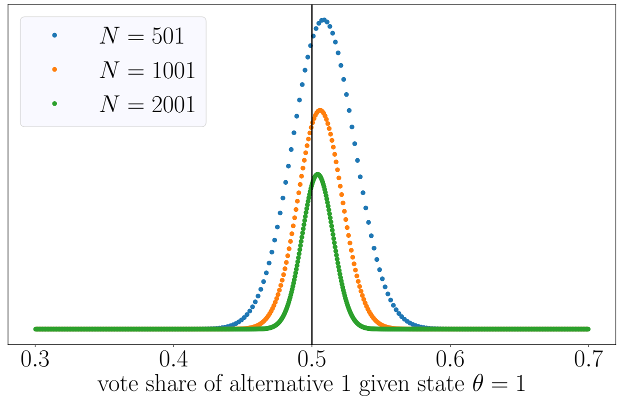


Figure 2: without opinion polls

Note: the distributions of the equilibrium vote shares in elections with the symmetric prior

In equilibrium, the following triadic relationship must hold: (i) voters' behavior is individually optimal under their beliefs; (ii) their beliefs are consistent with their information about the state and the poll; and (iii) the poll is consistent with the voters' actual behavior. An equilibrium is defined as a fixed point of this interactive system. This equilibrium concept is built on Denti (2023) and is interpreted by Hébert and La'O (2023) as a hybrid of a Bayesian Nash equilibrium and a rational-expectations equilibrium where agents learn from endogenous aggregate behavior while choosing their strategies, as in Grossman and Stiglitz (1980). In our context, this rational-expectations interpretation translates into voters learning about average voting behavior from the poll, while casting their votes. In addition, the equilibrium concept admits an interpretation as an outcome of a dynamic electoral process.

There are two types of equilibrium: an **informative equilibrium**, where voters acquire information; and an uninformative equilibrium, where they acquire no information. We show that an informative equilibrium exists unless either state occurs with an extremely high prior probability, and it is unique if it exists. In the Introduction, we focus on the informative equilibrium and refer to it simply as the equilibrium.

Main Results We consider equilibrium voting behavior in an election with a poll. Figure 1 plots the distributions of the equilibrium vote share for alternative 1, given state $\theta = 1$. Here we assume that the prior is symmetric, i.e., each state is equally likely. Since alternative 1 is chosen if the vote share exceeds 0.5 and alternative 0 is chosen otherwise, the majority vote is correct on the right side of 0.5 but incorrect on the left. It is notable that the distribution "jumps" at the winning threshold of 0.5. This jump is in line with our intuition that voters can learn, from the poll, how close the election could be, and that the poll encourages information acquisition when the election is close.

What is the probability of choosing the correct alternative? Our first main result (Theorem 1) is that *in any election with an opinion poll, the probability of correct choice is independent of the* *number of voters N and a prior μ.* In other words, in the presence of polls, Condorcet's optimism is perfectly balanced by Downs's pessimism of rational ignorance, keeping the probability of correct choice constant as the electorate size grows.

How is the equilibrium vote share distributed when the electorate becomes large? Our second main result (Theorem 2) is that *in any election with an opinion poll, as the number of voters N tends to infinity, the equilibrium vote share converges in probability to* 0.5 *even if a prior μ is asymmetric*. In other words, a large election is likely to be close even if either alternative has a higher prior probability of being correct. This result is illustrated in Figure 1, where the distributions concentrate on the winning threshold of 0.5 as N increases.

Both theorems are still true under supermajority rule, including unanimity rule. That is, the probability of correct choice is independent of the winning threshold as well as of the number of voters N and a prior μ (Theorem $1'$); moreover, the equilibrium vote share converges in probability to the winning threshold as N increases even if μ is asymmetric (Theorem $2'$).

Comparison between Elections with and without Opinion Polls Do polls help the electorate make the correct decision? To address this question, we compare an election with a poll to an otherwise identical election without a poll. Without a poll, voters only learn about a state, as in the earlier studies of the rational ignorance hypothesis. Figure 2 plots the equilibrium vote-share distributions when polls are unavailable. There is no jump at the winning threshold of 0.5, but instead the distributions are shifted in the direction of the correct alternative.

We demonstrate that *the probability of correct choice is strictly higher in any election with an opinion poll than in the identical election without an opinion poll, when the number of voters N is large and the prior μ is nearly symmetric, $\mu(1) \approx \frac{1}{2}$* (Proposition 1). This result highlights the role of polls in the acquisition and aggregation of information.

Regression Discontinuity Design in Close Elections Our findings offer novel implications for the **regression discontinuity (RD) design** in close elections. This is a common empirical strategy to identify treatment effects in electoral environments by, roughly speaking, comparing a candidate who barely wins an election with a candidate who barely loses. This approach has been widely used in the literature.

Despite the wide applications, the validity of the RD design in close elections has often been questioned. Its validity relies on the assumption that those candidates whose vote shares are immediately above the winning threshold are not systematically different from those whose vote shares are immediately below the threshold (Hahn, Todd, and Van der Klaauw, 2001). A violation of this assumption is known as sorting in the literature. Existing evidence is mixed. Caughey and Sekhon (2011) show that winners and losers in close U.S. House of Representatives elections significantly differ on pretreatment covariates, such as finance and incumbency; Snyder (2005) finds that incumbents won a disproportionate share of close U.S. House elections. In contrast, Eggers, Fowler, Hainmueller, Hall, and Snyder (2015) do not detect sorting in many close elections, yet confirm Snyder's (2005) finding for the U.S. House elections. These conflicting findings highlight the need to better understand what could lead to sorting, if it exists.

What could cause sorting? One possibility is post-election manipulation of vote counts (Snyder, 2005). Another is that well-organized campaigns might be able to exert precisely measured effort before the election to secure victory in close races (Caughey and Sekhon, 2011). These mechanisms, however, hinge on availability of voting information about the elections, which is reminiscent of polling. Indeed, Eggers et al. (2015) note that close U.S. House elections, in which sorting is repeated detected, are more frequently polled than most other elections for which RD designs have been employed.

We propose a novel mechanism to explain how polling could result in sorting in close elections, based exclusively on individual voter decisions. As shown in Figures 1 and 2, the vote-share distribution has a jump at the winning threshold if and only if an election is polled. This finding suggests the following interpretation: *even when alternatives are ex-ante identical, the electorate may coordinate, through polls, to make the (ex-post) correct decision by a narrow margin.* This polling-based mechanism is compatible with existing studies and helps reconcile the mixed evidence in the literature.

Related Literature Our study bridges the literature on rational ignorance and that on rational inattention. Downs's (1957) rational ignorance hypothesis argues that voters acquire information only when the benefits outweigh the costs. It suggests that individual voters may rationally choose to remain uninformed about politics in large elections, where the probability of being pivotal is low.[1] This hypothesis, by endogenizing voters' information acquisition, can be seen as a precursor to the modern approach of rational inattention. Existing studies of rational ignorance assume a "rigid" information structure, making parametric assumptions about the signal structures available to voters. Martinelli (2006) assumes that voters choose a parameter of signal precision about the correct alternative. He examines whether, in large electorates, the deterioration in signal quality, which stems from reduced incentives for information acquisition due to lower pivot probabilities, can outweigh the benefit of having more individual signals. Building on this model of information acquisition, Oliveros (2013) incorporates heterogeneous preferences, while Triossi (2013) introduces heterogeneous costs of information. Other related studies further restrict voters' decisions to 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). Experimental studies implement similar binary-choice information acquisition models (e.g., Bhattacharya, Duffy, and Kim, 2017; Elbittar, Gomberg, Martinelli, and Palfrey, 2020).

Our study is not the first to consider voters learning about their probability of being pivotal.

[1] In elections with exogenous information, Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1996) identify an incentive for voters to rationally disregard their private information. They argue that the value of a voter's pivotality interacts with the "swing voters' curse," which occurs when uninformed voters might prefer to abstain to avoid potentially canceling out the votes of more informed voters who share their preferences. This may lead voters to disregard their private information, because being pivotal itself signals how others are likely voting, which can be more informative than a voter's private signal.

Ekmekci and Lauermann (2022) assume that voters receive information about the electorate size, which indirectly informs them about their probability of being pivotal. They show that availability of such information may result in inefficient information aggregation. Due to the exogenous nature of information in their setting, they do not explore its implications for voters' information acquisition, which is our focus. In our model, voters are sure about the electorate size but may learn about a probable vote share through polls, which influences their probability of being pivotal.

The literature has examined the impact of communication between voters on information aggregation through voting. Assuming that voters have exogenous information, Feddersen and Pesendorfer (1998) show that without any communication between voters, unanimity rule results in voters disregarding a part of information to account for the "swing voter's curse." In response to this conclusion, Coughlan (2000) observes that a single round of cheap talk communication could allow for full information sharing. Gerardi and Yariv (2007) consider more general cheap-talk protocols, showing that a wide class of voting rules is equivalent with respect to equilibrium outcomes following the cheap talk. While these earlier studies explore the impact of information-sharing technology, we consider voters' acquisition of information about polls and show that an analogous equivalence result still holds (Theorems $1'$ and $2'$): if polls are available, the equilibrium probability of correct choice is independent of voting rules (i.e., simple majority, supermajority, or unanimity rule); in addition, a large election tends to be close regardless of these voting rules. This suggests that the underlying equivalence stems from voters learning about others, rather than from the specific modeling of information sharing.[2]

Although the rational inattention framework has not been used to examine the implications of the rational ignorance hypothesis, it has found other applications in political economy. Matějka and Tabellini (2021) examine a spatial model of electoral competition in which voters are rationally inattentive, and two candidates select their policies. Yuksel (2022) explores political polarization among rationally inattentive voters, and Li and Hu (2023) investigate politicians' accountability with rationally inattentive voters.[3]

Layout The remainder of this paper is organized as follows. Section 2 studies elections with opinion polls, providing main results. Section 3 compares elections with opinion polls to those without in terms of the probabilities of correct choice. Section 4 discusses the implications of our findings for the regression discontinuity design in close elections. Section 5 concludes by discussing extensions and limitations of our analysis. The proofs are found in Appendix A.

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

[2] Experimental studies examine the effect of polls. Sinclair and Plott (2012) find that voters update their beliefs when they are, through polls, informed about others' voting intentions and show that the number of voting errors decreases as more polls become available. Agranov, Goeree, Romero, and Yariv (2018) conjecture that voters are interested in the likelihood of being pivotal and observe that the availability of polls affects their decisions.

[3] These papers assume that there is a continuum of rationally inattentive agents. While this modeling may have certain advantages, it would not provide an appropriate framework for our question, which seeks to explore the balance between the Condorcet-type information aggregation and the free-riding incentives of the rational ignorance hypothesis. This is because an individual voter in a continuum would have an infinitesimally small impact, would never be pivotal, and thus would never acquire costly information.

2 Elections with Opinion Polls

In this section, we study elections with opinion polls. We first introduce an election model and an equilibrium concept and then examine equilibria in finite-voter elections. Lastly, we study large elections when the number of voters tend 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 $a \in A = \{0, 1\}$. Each voter i votes for alternative a_i, which we call action $a_i \in A$. As usual, let $a = (a_1, \ldots, a_N)$ be an action profile and a_{-i} be an action profile of all voters but i. No abstention is allowed. Given an action profile a, the **vote share (of alternative 1)** is defined as

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

while the vote share of alternative 0 is $1 - \bar{a}_N$. Majority vote chooses alternative 1 if $\bar{a}_N > \frac{1}{2}$ and alternative 0 otherwise. That is, the chosen alternative is $\mathbb{1}\{\bar{a}_N > \frac{1}{2}\}$, where $\mathbb{1}$ is the indicator function. There is no tie, since N is odd.

All voters have common interests and want to choose the **correct alternative**, which is initially unknown. The correct alternative is modeled by a state θ, a random variable with a common prior μ over a set $\Theta = \{0, 1\}$. The prior μ is said to be symmetric if $\mu(1) = \frac{1}{2}$ and asymmetric otherwise. Each voter's payoff is 1 if the chosen alternative is correct and 0 otherwise. Formally, we define each voter's payoff function $u : [0, 1] \times \Theta \to \{0, 1\}$ by

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

Information Acquisition Voters acquire information at cost. They have two variables to learn about: the state θ and an **opinion poll** that indicates how the votes in the rest of the electorate are likely to split. Specifically, voter i may acquire information about the vote share $\bar{a}_{-i} = \frac{1}{N-1} \sum_{j \neq i} a_j$.

We model voters' information acquisition, using the rational-attention framework (Sims, 2003). In this framework, voters flexibly choose what information to acquire, not only how much to acquire. This flexibility enables voters, for example, to learn more about θ when the election is likely to be close, and less so when it is not.

We now define voter i's strategies. She chooses 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 $(\bar{a}_{-i}, \theta)$ and then takes an action based on a signal realization. We can restrict, without loss of generality, voter i to "direct" signal structures where the signal space is $S_i = A$ and a signal realization s_i is interpreted as a recommendation to take action s_i.[4] Consequently, she chooses a conditional action distribution $P_i(\cdot \mid \bar{a}_{-i}, \theta) \in \Delta(A)$ for each $(\bar{a}_{-i}, \theta)$. Her strategy is the system of conditional action distributions P_i.

The cost of information is linear in mutual information (e.g., Sims, 2003; Matějka and McKay, 2015). Mutual information measures the reduction of voter i's uncertainty about $(\bar{a}_{-i}, \theta)$ due to information acquisition. This uncertainty is measured in terms of entropy. Suppose that she has a prior $\mu_i(\bar{a}_{-i}, \theta)$, which is endogenously formed in equilibrium (defined below). Then, the mutual information of P_i under μ_i is

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

where $\mathbb{H}$ is the entropy function. That is, $\mathbb{H}(\bar{a}_{-i}, \theta)$ 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.[5] That is, $\mathbb{I}(\bar{a}_{-i}, \theta; a_i)$ is the expected reduction of entropy by observing a_i. 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 the **election with a poll** by $\mathcal{P}_N$, omitting the prior μ and the unit cost of information λ.

2.2 Equilibrium

Voter i's strategy P_i is said to be optimal under her belief μ_i if it maximizes her expected payoff minus her 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, μ_i).

We define an equilibrium by the following triadic relationship: voters' strategies are optimal under their beliefs; their beliefs are consistent with their information about the state and a poll; and the poll is consistent with the voters' strategies. Formally, the equilibrium is the joint distribution over action profiles and states, which is induced by voters' strategies and beliefs.

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

1. The marginal distribution of θ is 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,

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

[4] By a standard argument from the literature (e.g., Matějka and McKay, 2015), every pair of a signal structure and a mapping from signal realizations to actions admits a (weakly) cheaper direct signal structure that induces the same conditional action distribution given each $(\bar{a}_{-i}, \theta)$. Also, voter i does not randomize among signal structures.

[5] The entropy $\mathbb{H}(Y)$ of a discrete random variable Y 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. The conditional entropy $\mathbb{H}(Z \mid Y)$ of a discrete random variable Z given 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, Chapter 2) for a comprehensive treatment.

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

We focus on a **symmetric equilibrium**, in which the strategies and beliefs are identical across all voters. Even in a symmetric equilibrium, voters may receive different signal realizations and thus vote for different alternatives.

This equilibrium concept, built on Denti (2023), is interpreted by Hébert and La'O (2023) as a hybrid of a Bayesian Nash equilibrium and a rational-expectations equilibrium. It is a Bayesian Nash equilibrium in the sense that agents behave optimally under uncertainty. It is a rational-expectations equilibrium in the sense that agents learn from endogenous aggregate behavior while simultaneously choosing their strategies as in Grossman and Stiglitz (1980). Agents' behavior is optimal under their beliefs, while their beliefs are consistent with their endogenous behavior. The equilibrium is then a fixed point of this interactive system. In our context, an opinion poll facilitates the rational-expectations interpretation. Voters make decisions while learning about a state and the poll; their voting behavior is consistent with the poll, because otherwise, the poll would fail to reflect voting behavior.

This equilibrium concept, though fully static, admits an interpretation as an outcome of a dynamic electoral process. This is in line with the stochastic best-response dynamics interpretation (the 2020 working paper version of Denti, 2023; Hoshino, 2018). Consider a sequential procedure where, at each period, one voter is randomly selected and receives an opportunity to acquire information and myopically revise her decision. Any revisions would be reflected in the subsequent poll, and the revision process would then repeat. This revision protocol can be viewed as a model of an environment in which polling is frequent and accessible to all voters. The myopia assumption would be sensible in large elections since individual votes have negligible impacts. This protocol results in a Markov chain that converges to a unique stationary distribution. The stationary distribution coincides with the equilibrium behavior of Definition 1.

An **equilibrium vote share** of a symmetric equilibrium P_N^* is defined as $\bar{a}_N = \frac{1}{N}\sum_{i=1}^N a_i$, where action profile $a = (a_1, \ldots, a_N)$ and state θ are distributed according to P_N^*. The conditional probability that $\bar{a}_N$ is in an interval T given state θ is $\Pr(\bar{a}_N \in T \mid \theta) = \sum_{a:\bar{a}_N \in T} P_N^*(a \mid \theta)$, where Pr denotes the probability. The conditional probabilities of choosing alternatives 1 and 0 given state θ are $\Pr(\bar{a}_N > \frac{1}{2} \mid \theta)$ and $\Pr(\bar{a}_N < \frac{1}{2} \mid \theta)$, respectively. The unconditional **probability of correct choice** is defined as

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

A symmetric equilibrium P_N^* is said to be **uninformative** if everyone votes without acquiring any information.[6] A symmetric equilibrium P_N^* is said to be **informative** if it is not uninformative (i.e., if all voters acquire some information).

[6] If $N \geq 3$, there exist two uninformative equilibria in which all vote for alternatives 1 or 0, respectively. This is because when the other $N-1$ voters vote for either alternative, the remaining voter acquires no information.

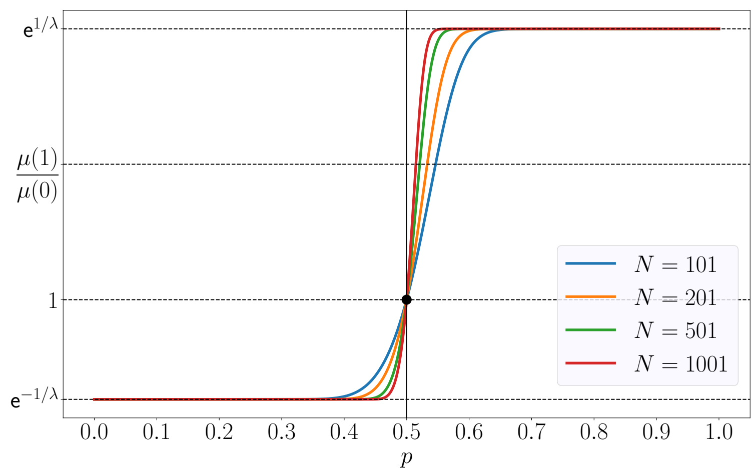


Figure 3: the graphs of $\frac{Z_N(p,1)}{Z_N(p,0)}$

It is well known that the optimal behavior under entropy-based costs is given by a biased-logit distribution (Matějka and McKay, 2015; Caplin, Dean, and Leahy, 2019; Denti, 2023). We demonstrate that our model has at most one informative equilibrium. To show this, we characterize equilibrium biased-logit distributions.

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

$$\Pr\left(\bar{a}_N = \frac{k}{N} \mid \theta\right) = \frac{1}{Z_N(p_N^*, \theta)} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k}, \tag{2}$$

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

$$Z_N(p, \theta) = \sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) p^k (1-p)^{N-k}, \tag{3}$$

and $p_N^* \in [0,1]$ is the unconditional probability of each individual voting for alternative 1.

One of the following holds:

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 $p_N^* \in (0,1)$ is a solution to equation*

$$\frac{Z_N(p,1)}{Z_N(p,0)} = \frac{\mu(1)}{\mu(0)}. \tag{4}$$

Informative Equilibrium By Lemma 1, an informative equilibrium exists if and only if (4) has a solution $p_N^* \in (0,1)$. Figure 3 plots $\frac{Z_N(p,1)}{Z_N(p,0)}$ as a function of p. Note that it is continuous and strictly increasing in p and ranges from $e^{-1/\lambda}$ to $e^{1/\lambda}$.[7] This implies that there is a unique informative equilibrium if $\frac{\mu(1)}{\mu(0)}$ is between $e^{-1/\lambda}$ and $e^{1/\lambda}$. This observation is formalized below.

[7] It is shown in the proof of Lemma 2.

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

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

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

How does the informative equilibrium vote-share distribution (2) look? Figure 1 in Section 1 plots the vote-share distributions at state $\theta = 1$ for the case of the symmetric prior μ, given the unit cost of information $\lambda = 1$. Figure 4 here shows the same distributions under the identical conditions, but for the case of asymmetric priors μ, $\mu(1) = 0.6, 0.4$.[8] In these figures, the horizontal axis indicates a vote share x, and the vertical axis indicates the conditional probability $\Pr(\bar{a}_N = x \mid \theta = 1)$.

The informative equilibrium vote-share distribution "jumps" at the winning threshold of 0.5, which increases the probability of correct choice. The reason for this jump is in line with our intuition that an opinion poll indicating a close election encourages voters to acquire more information and vote for the correct alternative.

2.3 Equilibrium Probability of Correct Choice

The informative equilibrium vote-share distribution changes as the number of voters N varies, but we show that the probability of correct choice does not.

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

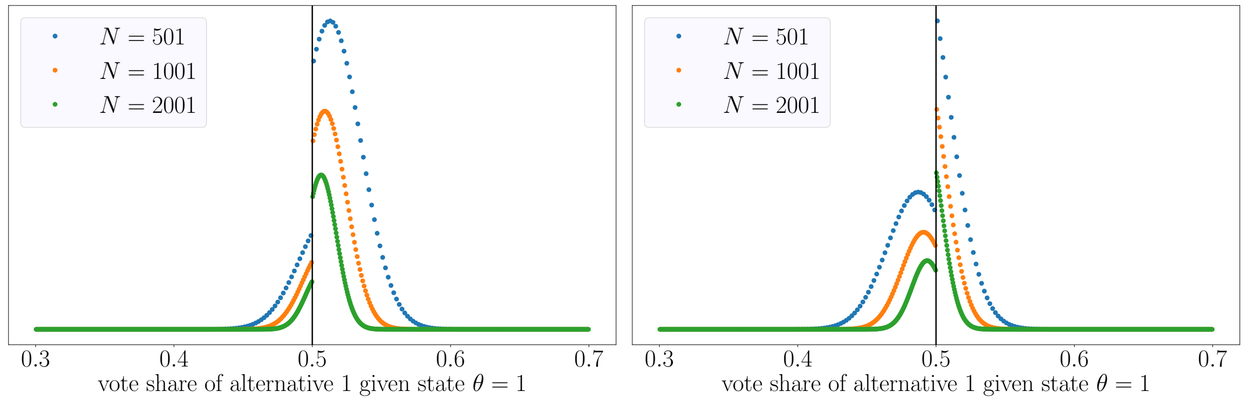
1. *In an uninformative equilibrium, with everyone voting for alternative $a = 0, 1$, the probability of correct choice is $\mu(a)$.*
2. *In an informative equilibrium, the probability of correct choice is*

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

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

Theorem 1 states that in the informative equilibrium, the probability of correct choice (5) is independent of the number of voters N and the prior μ. It demonstrates the perfect balance between Condorcet's (1785) optimism and Downs's (1957) pessimism. Increasing N influences the quality of the group decision in two ways: the positive effect of aggregating information from more voters, and the negative effect of each voter acquiring less information. Condorcet's Jury Theorem assumes that voters' information is exogenous and independent; thus, the positive effect translates into the law of large numbers and the negative effect is absent. Downs's rational ignorance hypothesis,

[8] Condition 1 is satisfied: $\frac{\mu(1)}{\mu(0)} = \frac{3}{2}, \frac{2}{3}$ if $\mu(1) = 0.6, 0.4$, respectively; $\mathsf{e}^{-1} \approx 0.3679$ and $\mathsf{e} \approx 2.7183$.



(a) the prior $\mu(1) = 0.6$ (b) the prior $\mu(1) = 0.4$

Figure 4: the distribution of the informative equilibrium vote share

which posits that voters' information is endogenous (and is independent in existing studies) stresses the negative effect. In contrast, we assume that voters' information is endogenous and correlated through polls, demonstrating that the two opposing effects precisely cancel out each other.

We sketch the proof of Theorem 1. The conditional probability of choosing alternative 1 at state $\theta = 1$ is $\Pr(\bar{a}_N > \frac{1}{2} \mid \theta = 1) = \sum_{k=n+1}^{N} \Pr(\bar{a}_N = \frac{k}{N} \mid \theta = 1)$, and the conditional probability of choosing alternative 0 at state $\theta = 0$ is $\Pr(\bar{a}_N < \frac{1}{2} \mid \theta = 0) = \sum_{k=0}^{n} \Pr(\bar{a}_N = \frac{k}{N} \mid \theta = 0)$. We express these probabilities using (2) and (4) and substitute them into (1), obtaining the theorem.

2.4 Large Elections

What happens to the equilibrium vote share $\bar{a}_N$ when the number of voters N increases? We examine the limiting probability $\lim_{N\to\infty} \Pr(\bar{a}_N \in T)$ for any interval $T \subset [0, 1]$. This analysis is not trivial. We cannot use basic tools such as the law of large numbers, because the behavior of any two voters is correlated and each voter's individual behavior changes as N changes. The equilibrium behavior is neither independently nor identically distributed.

We show that as the number of voters N increases, the unconditional probability p_N^* that each voter chooses alternative 1 approaches $\frac{1}{2}$ even if the prior μ is asymmetric. That is, even if either alternative has a higher prior of being correct, p_N^* is close to $\frac{1}{2}$ for any large N. In a sense, voters are "neutralized" from their asymmetric prior. By Lemma 1, p_N^* solves (4) in an informative equilibrium; namely, p_N^* is determined by the intersection of $\frac{Z_N(p,1)}{Z_N(p,0)}$ and $\frac{\mu(1)}{\mu(0)}$. Figure 3 plots $\frac{Z_N(p,1)}{Z_N(p,0)}$ as a function of p and the likelihood ratio $\frac{\mu(1)}{\mu(0)} \neq 1$ associated with an asymmetric prior μ. This figure illustrates that p_N^* approaches $\frac{1}{2}$ as N increases.

Lemma 3. *For each N, let P_N^* be the informative equilibrium of any election $\mathcal{P}_N$ that satisfies Condition 1. Then,*

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

Next we demonstrate that a large election has a high probability of being close. This result is illustrated in Figures 1 and 4: as the number of voters N increases, the informative equilibrium vote shares converge in probability to $\frac{1}{2}$, even if the prior μ is asymmetric.

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

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

Here is an intuition for Theorem 2. Given a large electorate, each voter has a low probability of being pivotal and thus acquires almost no information. Hence, the conditional probability of each voter choosing alternative 1 given any state θ must be close to the unconditional probability p_N^*, which is close to $\frac{1}{2}$ by Lemma 3. This, in turn, suggests that the vote share should be close to $\frac{1}{2}$ conditional on any θ. This illustration may be intuitive but is informal because voters' behavior is correlated in a complicated way, and we cannot rely on the law of large numbers, for example. To overcome this difficulty, we approximate the probability $\Pr(|\bar{a}_N - \frac{1}{2}| < \epsilon)$ with a more tractable form and then evaluate the approximation error.[9]

3 Comparison between Elections with and without Opinion Polls

We examine the role of polls by comparing elections with polls and those without. Then, we show that polls help a large electorate make the correct decision when a prior is nearly symmetric.

3.1 Elections without Opinion Polls

Base Environment The base environment is the same as in Section 2, but we use different notation for the sake of comparison between elections with and without polls. For the elections without polls, we denote voter i's action by $b_i \in A$. Let $b = (b_1, \ldots, b_N)$ be an action profile and b_{-i} be an action profile of all voters but i. Given an action profile b, we denote the vote share of alternative 1 by $\bar{b}_N = \frac{1}{N}\sum_{i=1}^{N} b_i \in [0, 1]$.

Information Acquisition We model information acquisition in the same way as in Section 2, except that now voters have no access to polling information. They only learn about state θ. We assume, without loss of generality, that voter i choose a conditional action distributions, which we denote by $Q_i(b_i \mid \theta)$. Her strategy is the system of conditional action distributions Q_i. This environment

[9] Hoshino and Ui (2024) study a class of games in which rationally inattentive players strategically interact. They examine the asymptotic equilibrium behavior as the number of players tends to infinity, but their model and analysis are different from ours. In our model, each voter's decision has a negligible impact on her own payoff and thus acquires negligible information when there are many voters. Hoshino and Ui allow each player's action to affect her payoff regardless of the number of players. Their model includes macro-finance settings, such as Keynesian beauty contests. This difference is crucial, requiring different analytical methods: we rely on Sanov's theorem, but this theorem is inapplicable to their class of games; Lemma 3 and Theorem 2 are specific to our setting and do not follow from their method.

differs from the one with a poll in that all voters' choice of action is conditionally independent given θ.

The cost of information is modeled in the same way as in Section 2. Since voters acquire information only about θ, mutual information measures the reduction of the entropy about θ due to information acquisition. The mutual information of Q_i under μ is $\mathbb{I}(\theta; b_i) = \mathbb{H}(\theta) - \mathbb{H}(\theta \mid b_i)$, where $\mathbb{H}(\theta)$ is the entropy of θ, and $\mathbb{H}(\theta \mid b_i)$ is the conditional entropy of θ given b_i, where the distribution of (b_i, θ) is induced by 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 the **election without a poll** by $\mathcal{Q}_N$, omitting the prior μ and the unit cost of information λ.

Equilibrium Voter i's strategy Q_i is said to be optimal given the others' strategies $Q_{-i} = (Q_j)_{j \neq i}$ if it maximizes her expected payoff minus her 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}, μ).

We define an **equilibrium** of the election without a poll $\mathcal{Q}_N$ as a strategy profile $(Q_1^*, \ldots, Q_N^*)$ such that each voter i's strategy Q_i^* is optimal given the others' strategies Q_{-i}^*. As in Section 2, we focus on a **symmetric equilibrium**, in which the equilibrium strategies are identical across all voters, $Q_1^* = \cdots = Q_N^*$. We denote by Q_N^* a symmetric equilibrium by abuse of notation.[10] A symmetric equilibrium Q_N^* is said to be **uninformative** if all voters vote without acquiring any information, and it is said to be **informative** if it is not uninformative.

An **equilibrium vote share** of a symmetric equilibrium Q_N^* is defined as a random variable $\bar{b}_N = \frac{1}{N}\sum_{i=1}^N b_i$, where $b_1, \ldots, b_N$ are i.i.d. random variables with the distribution $Q_N^*(\cdot \mid \theta)$ conditional on state θ. The conditional probability that $\bar{b}_N$ is in an interval $T \subset [0, 1]$ given state θ is $\Pr(\bar{b}_N \in T \mid \theta) = \sum_{b:\bar{b}_N \in T} \prod_{i=1}^N Q_N^*(b_i \mid \theta)$.

We study a symmetric equilibrium. Suppose that all voters $j \neq i$ choose the same strategy Q_N^*. Then, voter i's gross payoff (excluding information costs) when playing 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 when among $N - 1 = 2n$ voters, k vote for alternative 1 and $2n - k$ vote for alternative 0. Voter i's vote b_i affects her gross payoff if and only if her vote is pivotal: if $k = n$ then her gross payoff is $\mathbb{1}\{b_i = \theta\}$; if $k \neq n$ then her gross payoff is independent of b_i. The probability of being pivotal at state θ is $\Pi_N(\theta) \equiv \binom{2n}{n}(Q_N^*(1 \mid \theta))^n (Q_N^*(0 \mid \theta))^n$ when voters $-i$ play Q_N^*. Then, voter i's gross payoff is $\sum_\theta \mu(\theta)(\Pi_N(\theta) \cdot Q_i(\theta \mid \theta) + 0 \cdot Q_i(1 - \theta \mid \theta))$ plus

[10] In a symmetric equilibrium, voters may receive different signal realizations, voting for different alternatives.

a constant, and her problem is equivalent to

$$\max_{Q_i} \quad \sum_{\theta} \mu(\theta)\Pi_N(\theta)Q_i(\theta \mid \theta) - \lambda \mathbb{I}(\theta; b_i). \tag{6}$$

This problem is what Matějka and McKay (2015) and Caplin et al. (2019) have studied. They have shown that the optimal strategy is a biased-logit distribution. The following lemma is immediate from their results.

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

$$Q_N^*(1 \mid 1) = \frac{q_N^* \mathrm{e}^{\Pi_N(1)/\lambda}}{q_N^* \mathrm{e}^{\Pi_N(1)/\lambda} + 1 - q_N^*}, \quad Q_N^*(1 \mid 0) = \frac{q_N^*}{q_N^* + (1 - q_N^*)\mathrm{e}^{\Pi_N(0)/\lambda}}, \tag{7}$$

where $Q_N^*(0 \mid \theta) = 1 - Q_N^*(1 \mid \theta)$ and $\Pi_N(\theta) = \binom{2n}{n}(Q_N^*(1 \mid \theta))^n(Q_N^*(0 \mid \theta))^n$ for each $\theta = 0, 1$, and $q_N^* \in [0,1]$ is the unconditional probability of each voter voting for alternative 1.

One of the following holds:

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

Large Elections We study a large election, where the number of voters N tends to infinity. In line with the rational ignorance hypothesis, the probability of a single vote being pivotal vanishes and each voter acquires less information. Indeed, since $\lim_{N\to\infty} \Pi_N(\theta) = 0$, it follows, from Lemma 4, that

$$\lim_{N\to\infty} |Q_N^*(1 \mid 1) - q_N^*| = 0, \qquad \lim_{N\to\infty} |Q_N^*(1 \mid 0) - q_N^*| = 0.$$

This implies that for any subsequence of $\{q_N^*\}_N$ (which we denote by $\{q_N^*\}_N$ by slight abuse of notation) that has the limit $q_\infty^* \equiv \lim_{N\to\infty} q_N^*$, the equilibrium vote share $\bar{b}_N$ converges in probability to a constant. Formally, we have the following lemma:

Lemma 5. *For any subsequence of a sequence of symmetric equilibria $\{Q_N^*\}_N$, still denoted $\{Q_N^*\}_N$, along which the limit $q_\infty^* = \lim_{N\to\infty} q_N^*$ exists, and for any $\epsilon > 0$,*

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

From Lemma 5, it follows that if $q_\infty^* > \frac{1}{2}$ then the limit probability of correct choice is $\mu(1)$ since the electorate chooses alternative 1; similarly, if $q_\infty^* < \frac{1}{2}$ then the limit probability of correct choice is $\mu(0)$. The only nontrivial case is, therefore, the case of $q_\infty^* = \frac{1}{2}$. The limit probability of correct choice in this case is given in the following lemma.

Lemma 6. *For any subsequence of a sequence of symmetric equilibria $\{Q_N^*\}_N$, still denoted $\{Q_N^*\}_N$, along which the limit $q_\infty^* = \lim_{N\to\infty} q_N^*$ exists, one of the following holds:*

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

$$\lim_{N\to\infty} \Pr(u(\bar{b}_N, \theta) = 1) = \mu(1)\Phi(t^1) + \mu(0)\Phi(t^0), \tag{8}$$

where (t^1, t^0) is a solution to equations

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

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

The proof of Lemma 6 is somewhat subtle. The central limit theorem is inapplicable because voters' behavior, described by their equilibrium strategies Q^*_N, varies as the number of voters N changes. Instead, we rely on Berry–Esseen theorem, which quantifies a normal approximation by providing a bound on the difference between the actual distribution and the normal distribution for any finite N.

3.2 Comparison between Elections with and without Opinion Polls

We compare the probability of correct choice of elections with and without polls. Our result is that for any unit cost of information λ and any nearly symmetric prior $\mu(1) \approx \frac{1}{2}$, polls increase the probability of correct choice when there are sufficiently many voters.

Proposition 1. *For any unit cost of information $\lambda > 0$, there exists an $\epsilon > 0$ such that for any prior μ with $|\mu(1) - \frac{1}{2}| < \epsilon$, there exists an $\bar{N} \in \mathbb{N}$ such that for any $N > \bar{N}$, the probability of correct choice in the informative equilibrium P^*_N of the election with opinion polls $\mathcal{P}_N$ is strictly greater than that in any symmetric equilibrium Q^*_N of the election without opinion polls $\mathcal{Q}_N$.*

Proposition 1 is illustrated in Figure 5, which plots the probabilities of correct choice as a function of a unit cost λ in the elections with and without polls in the case of the symmetric prior $\mu(1) = \frac{1}{2}$. The blue graph is the informative equilibrium probability of correct choice $\Pr(u(\bar{a}_N, \theta) = 1) = \frac{\mathrm{e}^{1/\lambda}}{1+\mathrm{e}^{1/\lambda}}$ in the election with polls (Theorem 1), which is independent of N. The orange graph is the limit probability of correct choice $\lim_{N\to\infty} \Pr(u(\bar{b}_N, \theta) = 1)$ in the election without polls (Lemma 6). The key to Proposition 1 is to show that the blue graph is above the orange graph. Once we establish it, since $\lim_{N\to\infty} \Pr(u(\bar{b}_N, \theta) = 1)$ is continuous in μ, we have the desired result for any nearly symmetric prior $\mu(1) \approx \frac{1}{2}$.

Lastly, we note how the probability of correct choice varies as the unit cost λ vanishes. Information accessibility helps the electorate make the correct decision. As in Figure 5, the probability of correct choice tends to 1 as λ vanishes regardless of whether the election is polled or not, but polls increase the probability of correct choice for any $\lambda > 0$.

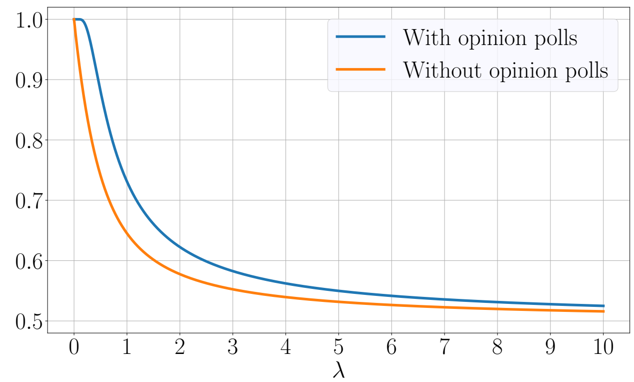


Figure 5: Proposition 1 under the symmetric prior μ

4 Regression Discontinuity Design in Close Elections

Our findings offer novel implications for the **regression discontinuity (RD) design** in close elections. The RD design in such contexts identifies electoral treatment effects based on the assumption that narrowly winning and narrowly losing candidates are similar in all respects, except for the election outcome. However, our results demonstrate that better candidates may systematically win close elections when polls are available. This offers a new perspective to interpret mixed evidence about the RD validity discussed in the existing literature.

Regression Discontinuity Design The RD design, initially proposed by Thistlethwaite and Campbell (1960), is an empirical strategy for identifying treatment effects without random experimental assignments of subjects to treatments. In electoral settings, the RD design is applied to close elections because the comparison between narrowly winning and narrowly losing candidates is considered an ideal quasi-experiment. The key to the RD design is the continuity assumption, which posits that narrowly winning candidates, whose vote shares are just above the winning threshold, and narrowly losing candidates, whose vote shares are just below the threshold, have similar distributions of unobservable characteristics (Hahn et al., 2001).

Numerous studies utilize the RD design in close elections. Lee (2001, 2008) employs the RD design to estimate the incumbency advantage in future U.S. congressional elections. Ferreira and Gyourko (2009, 2014) use the RD design to explore the impact of political parties and politicians' gender on municipal fiscal policies in the U.S. Firpo, Ponczek, and Sanfelice (2015) apply the RD design to study the federal budget process in Brazil. Dell (2015) leverages the RD design to investigate the interaction between drug trafficking networks and drug-related violence in Mexico.

Despite their wide application, the validity of the RD design in close elections is often questioned. This validity hinges on the continuity assumption, but this assumption would be violated when certain types of candidates systematically win close races, resulting in the so-called sorting. As discussed in Section 1, the existing literature provides mixed evidence about the existence of sorting. As de la Cuesta and Imai (2016) summarize in their survey, "the literature is remarkably divided on the question of whether sorting exists in the close election context."

Novel Mechanism for Sorting As discussed in Section 1, Eggers et al. (2015) point out polls as a possible reason for sorting in close elections. The novel mechanism that we propose is line with their insight. The proposed mechanism is illustrated in Figures 1 and 2, which plot the equilibrium vote-share distributions under the symmetric prior given state $\theta = 1$. Figure 1 shows that in elections with polls, there is a jump at the winning threshold of 0.5. This jump occurs because a voter who learns that she might be pivotal through the poll has a strong incentive to acquire information and vote for the correct candidate. This behavior near the threshold increases the probability of the correct candidate winning. In contrast, Figure 2 shows that in elections without polls, there is no such jump because a voter never learns about her pivotality.

Thus, our results suggest that in the presence of polls, the better candidate should systematically win with a substantially higher probability, regardless of how close the election is (Theorems 1 and 2). The identity of the better candidate may not be observable, but it may be correlated with observable variables, such as incumbency or financial resources. Our results indicate that even if winners and losers in close elections appear balanced on observable characteristics, the RD design should be approached with caution when voters have access to polls. Our findings do not invalidate the RD design in close elections but rather provide a complementary perspective that helps reconcile the mixed evidence in the existing literature.

5 Concluding Remarks

We have demonstrated the impact of opinion polls on information acquisition and aggregation in common-value elections. To this end, we have developed a model that captures interaction between voters' decisions and opinion polls. We conclude by discussing extensions and limitations of our analysis.

5.1 Supermajority Rule

We have so far assumed simple majority rule, but our main results, Theorems 1 and 2, continue to hold even if a supermajority vote (including unanimity) is required to overturn a default choice.

We consider the same model as before, except that now alternative 1 is chosen if and only if the vote share $\bar{a}_N$ is at least a given threshold $\alpha \in (\frac{1}{2}, 1]$.[11] This model includes the unanimity rule as the special case of $\alpha = 1$. As in the main model, each voter's payoff is 1 if the chosen alternative is correct and 0 otherwise. We denote this supermajority election with a poll by $\tilde{\mathcal{P}}_N$, omitting the prior μ, the unit cost of information λ, and the winning threshold α. We use the same equilibrium concept as Definition 1.

We extend our main results to the case of the supermajority rule. We relegate the proofs to Online Appendix C because they are mostly similar to those of the original theorems although we

[11] This setting is consistent with the case of the simple majority rule, under which there is no tie since N is odd.

need more intricate analysis and more involved notation. Lemmas 1 and 2 extend to any election $\tilde{\mathcal{P}}_N$ with any winning threshold α, as shown in Online Appendix C. In particular, an informative equilibrium exists if and only if Condition 1 is satisfied, and it is unique whenever it exists.

We then extend Theorem 1 to any election $\tilde{\mathcal{P}}_N$ with a winning threshold $\alpha \in (\frac{1}{2}, 1]$. That is, the probability of correct choice at the informative equilibrium is $\frac{e^{1/\lambda}}{1+e^{1/\lambda}}$, which is independent of α as well as of the number of voters N and the prior μ with Condition 1.

Theorem 1′. *Theorem 1 holds as is, in any election $\tilde{\mathcal{P}}_N$.*

Next, we extend Theorem 2 to any election $\tilde{\mathcal{P}}_N$ with a winning threshold $\alpha \in (\frac{1}{2}, 1]$. That is, as the number of voters grows, the informative equilibrium vote share tends to be closer to α.

Theorem 2′. *For each N, let $\bar{a}_N$ be the vote share in the informative equilibrium P_N^* of any election $\tilde{\mathcal{P}}_N$ with Condition 1. For each $\epsilon > 0$,*

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

5.2 Role of Entropy-Based Costs of Information

We focus on the entropy-based costs of information. This is arguably the most standard specification in the literature (Sims, 2003; Matějka and McKay, 2015). We exploit this functional form in our proofs. The biased-logit characterization of equilibria is specific to the entropy-based costs, and so is Theorem 1. The proof of Theorem 2 also relies on the entropy-based specification. Recall that the proof idea is that we approximate the probability $\Pr(|\bar{a}_N - \frac{1}{2}| < \epsilon)$ that the equilibrium vote share $\bar{a}_N$ is near the winning threshold of $\frac{1}{2}$, using a more tractable form, and then evaluate the approximation error. Our error evaluation technique relies on a property of the entropy-based costs. To which extent any of these results may generalize for other specifications of information costs remains an open question.

Nevertheless, much of the intuition for our results is independent of the entropy specification. While it is an open question whether the perfect balance between Condorcet's optimism and Downs's pessimism (Theorem 1) would remain true under different cost functions, the idea that information sharing through polls helps alleviate the free rider problem appears to be generally true. The full role that the entropy costs play here remains an important and potentially intriguing avenue for future research.

5.3 Information as Public Goods

A striking implication of our model is that the probability of majority vote making the correct choice in the presence of polls is independent of the number of voters (Theorem 1). To understand the result and the potential difficulty with generalizing it, an analogy with voluntary provision of a public good might be useful. Information once shared is both non-rivalrous and non-excludable, and information acquisition in our environment can be viewed as a voluntary contribution to its provision. As in the standard voluntary provision game (e.g., Bergstrom, Blume, and Varian, 1986) the amount of the public good provided is determined by equating individual marginal benefits and marginal costs of information, and the additively separable structure of information costs is reminiscent of the quasi-linear utilities that obviate income effects. In the absence of income effects, the equilibrium amount of a voluntarily provided public good would likewise be independent of the number of agents.

A major difference from the classic voluntary provision game is that in our voting model, the benefit from acquiring information is discounted by the probability of being pivotal. If polling information were free, voters would only acquire information when they are pivotal, which completes the public-good analogy. Since information about polls is costly in our model, it may in general not be clear how equilibrium costs of information about the polls and the state interact. This makes us suspect that the full neutrality we observe here may depend on the choice of the information cost specification.

5.4 Opinion Polls Based on Random Sampling

In this paper, an opinion poll refers to a complete census that aggregates all voters' intentions. In practice, an opinion poll is based on a subset consisting of randomly sampled voters. Sampling error would introduce additional noise into voters' information, but as long as a sufficient number of voters are randomly sampled, voters should be able to access essentially the same information. Hence, we believe that in such a model, most of our results would still obtain. Explicitly modeling sampling, however, may open additional questions about the interaction between information acquisition and the design of opinion polls, such as considerations of sampling size, content and dissemination of polls, or how costly learning about them should be. These questions, which have not, to the best of our knowledge, been addressed in the literature, would link our work to the 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 the potential. Applying Denti's (2023) Corollary 1 to our setting, we have the following lemma:

Lemma A. *Every symmetric equilibrium P_N^* of an election with opinion polls $\mathcal{P}_N$ is such that for* *some* $p_N^* \in [0,1]$ *and for each* $a = (a_1, \ldots, a_N)$ *and each* θ*,*

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

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

$$U(p_1, \ldots, p_N) = \sum_\theta \mu(\theta) \ln \left(\sum_{a' \in \{0,1\}^N} \exp\left(\frac{u(\bar{a}'_N, \theta)}{\lambda}\right) \prod_{i=1}^N \left(p_i \mathbb{1}\{a'_i = 1\} + (1 - p_i) \mathbb{1}\{a'_i = 0\}\right) \right).$$

The biased logit formula of (2) is immediate from this lemma because $\Pr(\bar{a}_N = \frac{k}{N} \mid \theta) = \binom{N}{k} P_N^*(a_1, \ldots, a_N \mid \theta)$, and the denominator of $P_N^*(a_1, \ldots, a_N \mid \theta)$ is equal to $Z_N(p, \theta)$ by the definition of (3).

To prove Lemma 1, it suffices to show that any symmetric pure-strategy Nash equilibrium $(p_N^*, \ldots, p_N^*)$ of the normal-form game of Lemma A is either $p_N^* \in \{0,1\}$ or a solution to (4). At the symmetric Nash equilibrium, it must be that $p_N^* \in \operatorname{argmax}_{p_i} U(p_i, p_N^*, \ldots, p_N^*)$. Since this is immediate if $p_N^* \in \{0,1\}$, let $p_N^* \in (0,1)$. In this interior case, $p_i = p_N^*$ satisfies the first-order condition, $\frac{\partial U}{\partial p_i}(p_N^*, p_N^*, \ldots, p_N^*) = 0$. It is also sufficient since $U(p_i, p_N^*, \ldots, p_N^*)$ is strictly concave in p_i (Caplin et al., 2019, p. 1066). We write the first-order condition as

$$\sum_\theta \mu(\theta) \cdot \frac{\displaystyle\sum_{k=0}^{2n} \binom{2n}{k} \left[\exp\left(\frac{u(\frac{k+1}{N}, \theta)}{\lambda}\right) - \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right)\right] (p_N^*)^k (1 - p_N^*)^{2n-k}}{\displaystyle\sum_{k=0}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N}, \theta)}{\lambda}\right) (p_N^*)^k (1 - p_N^*)^{N-k}} = 0. \tag{9}$$

The denominator is $Z_N(p_N^*, \theta)$ by the definition of (3). In the numerator, if $k \neq n$ then the square bracket is zero since $u(\frac{k+1}{N}, \theta) = u(\frac{k}{N}, \theta)$, while if $k = n$ then the square bracket is $\mathsf{e}^{1/\lambda} - 1$ when $\theta = 1$ and $1 - \mathsf{e}^{1/\lambda}$ when $\theta = 0$. Substituting them into (9), we have

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

which is equivalent to (4). Hence, $(p_N^*, \ldots, p_N^*)$ is a Nash equilibrium if and only if p_N^* is a solution to (4).

A.2 Lemma 2

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

$$\begin{aligned} W_N(p,1) &= \sum_{k=n+1}^{N} \binom{N}{k} p^k (1-p)^{N-k}, \\ W_N(p,0) &= \sum_{k=0}^{n} \binom{N}{k} p^k (1-p)^{N-k}. \end{aligned} \tag{10}$$

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 taking derivatives. Then, we rewrite Z_N, as defined in (3), as

$$\begin{aligned} Z_N(p,1) &= W_N(p,0) + \mathsf{e}^{1/\lambda} W_N(p,1), \\ Z_N(p,0) &= \mathsf{e}^{1/\lambda} W_N(p,0) + W_N(p,1). \end{aligned} \tag{11}$$

Note that $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p, because $Z_N(p,1)$ is strictly increasing in p and $Z_N(p,0)$ strictly decreasing in p.[12]

Step 2 By Lemma 1, it suffices to show that (4) has a unique solution if and only if Condition 1 holds. Note that $W_N(1,1) = W_N(0,0) = 1$ and $W_N(0,1) = W_N(1,0) = 0$ by substitution. By (11),

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

Hence, if Condition 1 is satisfied, $\frac{Z_N(0,1)}{Z_N(0,0)} < \frac{\mu(1)}{\mu(0)} < \frac{Z_N(1,1)}{Z_N(1,0)}$. Since $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p, (4) has a unique solution $p_N^* \in (0,1)$. If Condition 1 is not satisfied, we have 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 either case, (4) has no solution in $(0,1)$.

A.3 Theorem 1

By (2), the probability of choosing alternative 1 at state $\theta = 1$ is

$$\begin{aligned} \Pr\left(\bar{a}_N > \frac{1}{2} \mid \theta = 1\right) &= \sum_{k=n+1}^{N} \Pr\left(\bar{a}_N = \frac{k}{N} \mid \theta = 1\right) \\ &= \frac{1}{Z_N(p_N^*,1)} \sum_{k=n+1}^{N} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N},1)}{\lambda}\right) (p_N^*)^k (1-p_N^*)^{N-k}. \end{aligned}$$

[12]To see that $Z_N(p,1)$ is strictly increasing in p, 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 that $W_N(p,1)$ is strictly increasing in p. Similarly, we see that $Z_N(p,0)$ is strictly decreasing in p.

Since $u(\frac{k}{N}, 1) = 1$ for all $k = n+1, \ldots, N$, we have, by (10) and (11),

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

By (11), we rewrite (4) as

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

Equivalently, $\frac{W_N(p_N^*,1)}{W_N(p_N^*,0)} = \frac{\mathrm{e}^{1/\lambda}\mu(1) - \mu(0)}{\mathrm{e}^{1/\lambda}\mu(0) - \mu(1)}$. Hence,

$$\Pr\Big(\bar{a}_N > \frac{1}{2} \mid \theta = 1\Big) = \frac{\mathrm{e}^{1/\lambda} - \frac{\mu(0)}{\mu(1)}}{\mathrm{e}^{1/\lambda} - \mathrm{e}^{-1/\lambda}}$$

Similarly,

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

By substitution, we obtain the desired result:

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

Lastly, we show that $\frac{\mathrm{e}^{1/\lambda}}{1+\mathrm{e}^{1/\lambda}} > \max\{\mu(1), \mu(0)\}$ if the informative equilibrium exists. Since the existence is equivalent to Condition 1 (Lemma 2), it suffices to derive the inequality under Condition 1. This is verified by simple algebra. Indeed, $\mathrm{e}^{-1/\lambda} < \frac{\mu(1)}{\mu(0)} < \mathrm{e}^{1/\lambda}$ (Condition 1) is equivalent to $\frac{\mathrm{e}^{-1/\lambda}}{1+\mathrm{e}^{-1/\lambda}} < \mu(1) < \frac{\mathrm{e}^{1/\lambda}}{1+\mathrm{e}^{1/\lambda}}$, which is equivalent to the desired inequality.

A.4 Lemma 3

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

$$\frac{Z_N(\frac{1}{2} - \epsilon, 1)}{Z_N(\frac{1}{2} - \epsilon, 0)} < \frac{\mu(1)}{\mu(0)} < \frac{Z_N(\frac{1}{2} + \epsilon, 1)}{Z_N(\frac{1}{2} + \epsilon, 0)}. \tag{12}$$

To see that (12) is sufficient, we note that $\frac{Z_N(p,1)}{Z_N(p,0)}$ is continuous and strictly increasing in p. If (12) is true then $p_N^* \in (\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon)$, where p_N^* is a solution to (4).

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

$$\begin{aligned} W_N(\tfrac{1}{2} + \epsilon, 1) > 1 - \delta, \quad W_N(\tfrac{1}{2} + \epsilon, 0) < \delta, \\ W_N(\tfrac{1}{2} - \epsilon, 0) > 1 - \delta, \quad W_N(\tfrac{1}{2} - \epsilon, 1) < \delta, \end{aligned} \tag{13}$$

where W_N is defined in (10). 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 $\frac{1}{2}+\epsilon$ and $\frac{1}{2}-\epsilon$ respectively. Then, $W_N(\frac{1}{2}+\epsilon,1)$ and $W_N(\frac{1}{2}+\epsilon,0)$ are the probabilities that the sample average $\frac{1}{N}\sum_{i=1}^N w_i$ 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 $W_N(\frac{1}{2}+\epsilon,1) > 1-\delta$ and $W_N(\frac{1}{2}+\epsilon,0) < \delta$. 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 $\frac{1}{2}-\epsilon$ and $\frac{1}{2}+\epsilon$ respectively. By the same argument, there is an N''_δ such that for any $N > N''_\delta$, we have $W_N(\frac{1}{2}-\epsilon,0) > 1-\delta$ and $W_N(\frac{1}{2}-\epsilon,1) < \delta$. Lastly, let $N_\delta = \max\{N'_\delta, N''_\delta\}$.

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

$$\frac{1+\mathrm{e}^{1/\lambda}\delta}{\mathrm{e}^{1/\lambda}(1-\delta)} < \frac{\mu(1)}{\mu(0)} < \frac{\mathrm{e}^{1/\lambda}(1-\delta)}{\mathrm{e}^{1/\lambda}\delta+1}. \tag{14}$$

To prove this inequality, note that for a small enough δ, we have the LHS and RHS arbitrarily close to $\mathrm{e}^{-1/\lambda}$ and $\mathrm{e}^{1/\lambda}$, respectively. Since $\mathrm{e}^{-1/\lambda} < \frac{\mu(1)}{\mu(0)} < \mathrm{e}^{1/\lambda}$ (Condition 1), we obtain (14).

Now we prove (12). For any $N > N_\delta$,

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

where the equality is by (11), the first inequality by (13), and the second inequality by (14). Also,

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

where the equality is by (11), the first inequality by (13), and the second inequality by (14). Thus, we have (12), which completes the proof.

A.5 Theorem 2

Fix any θ and any l,h such that $0 \le l < h \le 1$. By Lemma 1,

$$\Pr(\bar{a}_N \in [l,h] \mid \theta) = \frac{\mathcal{Z}_N([l,h],\theta)}{\mathcal{Z}_N([0,1],\theta)},$$

where for any interval $T \subset [0,1]$,

$$\mathcal{Z}_N(T,\theta) \equiv \sum_{k:\frac{k}{N}\in T} \binom{N}{k} \exp\left(\frac{u(\frac{k}{N},\theta)}{\lambda}\right)(p_N^*)^k(1-p_N^*)^{N-k}.$$

Here, $\sum_{k:\frac{k}{N}\in T}$ runs over all $k = 0,\dots,N$ such that $\frac{k}{N} \in T$. Hence,

$$\frac{1}{N}\ln \Pr(\bar{a}_N \in [l,h] \mid \theta) = \frac{1}{N}\ln \mathcal{Z}_N([l,h],\theta) - \frac{1}{N}\ln \mathcal{Z}_N([0,1],\theta). \tag{15}$$

Step 1 Fix any $\delta > 0$. There exists an N_1 such that for any $N \geq N_1$ and any k,

$$-N\delta < \frac{u(\frac{k}{N}, \theta)}{\lambda} < N\delta, \tag{16}$$

which is immediate for $u(\frac{k}{N}, \theta) = 0, 1$. By Lemma 3, there exists an N_2 such that for any $N \geq N_2$, $|\ln p_N^* - \ln \frac{1}{2}| < \delta$ and $|\ln(1 - p_N^*) - \ln \frac{1}{2}| < \delta$. By the triangle inequality,

$$\left| k \ln p_N^* + (N-k)\ln(1-p_N^*) - N \ln \frac{1}{2} \right| \leq k \left| \ln p_N^* - \ln \frac{1}{2} \right| + (N-k) \left| \ln(1-p_N^*) - \ln \frac{1}{2} \right| < N\delta.$$

It follows that for any $N \geq N_2$ and any k,

$$\mathsf{e}^{-N\delta} 2^{-N} < (p_N^*)^k (1-p_N^*)^{N-k} < \mathsf{e}^{N\delta} 2^{-N}. \tag{17}$$

Evaluating $\mathcal{Z}_N([l,h], \theta)$ with (16) and (17), we obtain that for any $N \geq \max\{N_1, N_2\}$,

$$\left| \frac{1}{N} \ln \mathcal{Z}_N([l,h], \theta) - \frac{1}{N} \ln \sum_{k: \frac{k}{N} \in [l,h]} \binom{N}{k} 2^{-N} \right| < 2\delta. \tag{18}$$

Step 2 Note that $\sum_{k: \frac{k}{N} \in [l,h]} \binom{N}{k} 2^{-N}$ is the probability that the sample average of N i.i.d. symmetric Bernoulli random variables, which take values 1 and 0 with equal probabilities $\frac{1}{2}$, is in $[l,h]$. By Sanov's theorem (Cover and Thomas, 2006, Theorem 11.4.1), there exists an N_3 such that for any $N \geq N_3$,

$$\left| \frac{1}{N} \ln \sum_{k: \frac{k}{N} \in [l,h]} \binom{N}{k} 2^{-N} + \min_{t \in [l,h]} \left\{ D_{\mathsf{KL}} \Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2}) \Big) \right\} \right| < \delta, \tag{19}$$

where $\mathsf{B}(t)$ is the Bernoulli distribution that assigns to values 1 and 0 probabilities t and $1-t$, and $D_{\mathsf{KL}}(\mathsf{B}(t) \parallel \mathsf{B}(\frac{1}{2}))$ is the Kullback–Leibler divergence of $\mathsf{B}(t)$ from $\mathsf{B}(\frac{1}{2})$, which is defined as

$$D_{\mathsf{KL}} \Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2}) \Big) \equiv t \ln \left(\frac{t}{1/2} \right) + (1-t) \ln \left(\frac{1-t}{1/2} \right),$$

where $0 \ln(0) = 0$ by convention. In the rest of the proof, we will use the non-negativity property: $D_{\mathsf{KL}}(\mathsf{B}(t) \parallel \mathsf{B}(\frac{1}{2})) \geq 0$ for all $t \in [0,1]$, with equality if and only if $t = \frac{1}{2}$ (Cover and Thomas, 2006, Theorem 2.6.3). The minimum exists in (19) since $D_{\mathsf{KL}}(\mathsf{B}(t) \parallel \mathsf{B}(\frac{1}{2}))$ is continuous in t and $[l,h]$ is compact.

Step 3 By the triangle inequality with (18) and (19), for any $N \geq \max\{N_1, N_2, N_3\}$,

$$\left| \frac{1}{N} \ln \mathcal{Z}_N([l,h], \theta) + \min_{t \in [l,h]} \left\{ D_{\mathsf{KL}} \Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2}) \Big) \right\} \right| < 3\delta.$$

Since the choice of $\delta > 0$ is arbitrary,

$$\lim_{N\to\infty} \frac{1}{N} \ln \mathcal{Z}_N([l,h],\theta) = -\min_{t\in[l,h]} \Big\{ D_{\mathsf{KL}}\Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2})\Big)\Big\}.$$

In particular,

$$\lim_{N\to\infty} \frac{1}{N} \ln \mathcal{Z}_N([0,1],\theta) = -\min_{t\in[0,1]} \Big\{ D_{\mathsf{KL}}\Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2})\Big)\Big\} = 0,$$

where we use the non-negativity of the Kullback–Leibler divergence.

By (15), we conclude that for any θ and any l, h such that $0 \le l < h \le 1$,

$$\lim_{N\to\infty} \frac{1}{N} \ln \Pr(\bar{a}_N \in [l,h] \mid \theta) = -\min_{t\in[l,h]} \Big\{ D_{\mathsf{KL}}\Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2})\Big)\Big\}. \tag{20}$$

Step 4 Fix any $\epsilon > 0$. By (20),

$$\lim_{N\to\infty} \frac{1}{N} \ln \Pr\Big(\bar{a}_N \in \Big[0, \frac{1}{2} - \epsilon\Big] \mid \theta\Big) = -\min_{t\in[0,\frac{1}{2}-\epsilon]} \Big\{ D_{\mathsf{KL}}\Big(\mathsf{B}(t) \parallel \mathsf{B}(\tfrac{1}{2})\Big)\Big\} < 0,$$

where the inequality is by the non-negativity of the Kullback–Leibler divergence. Hence,

$$\lim_{N\to\infty} \Pr\Big(\bar{a}_N \in \Big[0, \frac{1}{2} - \epsilon\Big] \mid \theta\Big) = 0.$$

Similarly,

$$\lim_{N\to\infty} \Pr\Big(\bar{a}_N \in \Big[\frac{1}{2} + \epsilon, 1\Big] \mid \theta\Big) = 0.$$

It then follows that

$$\lim_{N\to\infty} \Pr\Big(\Big|\bar{a}_N - \frac{1}{2}\Big| < \epsilon \mid \theta\Big) = 1.$$

Since this is true for each θ, we have $\lim_{N\to\infty} \Pr(|\bar{a}_N - \frac{1}{2}| < \epsilon) = 1$.

A.6 Lemma 4

This lemma is immediate from the biased-logit formula, which is derived by Matějka and McKay (2015, Theorem 1) and Caplin et al. (2019, Proposition 1).

A.7 Lemma 5

Note that as $N = 2n + 1 \to \infty$,

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

By Lemma 4, $\lim_{N\to\infty} |Q^*_N(1 \mid 1) - q^*_N| = 0$ and $\lim_{N\to\infty} |Q^*_N(1 \mid 0) - q^*_N| = 0$.

Consider the case of state $\theta = 1$, as the case of state $\theta = 0$ is analogous. Fix any $\epsilon > 0$. There exists an N_1 such that for each $N > N_1$, $|Q^*_N(1 \mid 1) - q^*_N| < \frac{\epsilon}{3}$. Since $\lim_{N\to\infty} q^*_N = q^*_\infty$ by assumption, there exists an N_2 such that for each $N > N_2$, $|q^*_N - q^*_\infty| < \frac{\epsilon}{3}$. For any $N > \max\{N_1, N_2\}$,

$$|Q^*_N(1 \mid 1) - q^*_\infty| \leq |Q^*_N(1 \mid 1) - q^*_N| + |q^*_N - q^*_\infty| < \frac{2\epsilon}{3}.$$

Voters' actions $b_1, b_2, \dots$ 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 any $N > N_3$,

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

Since $|Q^*_N(1 \mid 1) - q^*_\infty| < \frac{2\epsilon}{3}$, we have $\Pr(|\bar{b}_N - q^*_\infty| < \epsilon \mid \theta = 1) > 1 - \delta$ for any $N > \max\{N_1, N_2, N_3\}$.

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, there exist $t^1_N, t^0_N \in [-\frac{1}{2}, \frac{1}{2}]$ such that $Q^*_N(1 \mid 1) = \frac{1}{2} + t^1_N$ and $Q^*_N(1 \mid 0) = \frac{1}{2} - t^0_N$. For each N, there exists $t_N \in [-\frac{1}{2}, \frac{1}{2}]$ such that $q^*_N = \frac{1}{2} + t_N$. Since q^*_N is the unconditional probability of voting for alternative 1 (Lemma 4), it must be that $q^*_N = \mu(1)Q^*_N(1 \mid 1) + \mu(0)Q^*_N(1 \mid 0)$. Hence,

$$t_N = \mu(1)t^1_N - \mu(0)t^0_N. \tag{21}$$

Since $q^*_N \to \frac{1}{2}$ as $N \to \infty$, it follows that $t^1_N \to 0$, $t^0_N \to 0$, and $t_N \to 0$.

Step 1 Consider the unconditional probability of correct choice. Note that

$$\Pr(u(\bar{b}_N, \theta) = 1) = \mu(1)\Pr\left(\bar{b}_N > \frac{1}{2} \mid \theta = 1\right) + \mu(0)\Pr\left(\bar{b}_N < \frac{1}{2} \mid \theta = 0\right). \tag{22}$$

We consider $\Pr(\bar{b}_N > \frac{1}{2} \mid \theta = 1)$ in (22). Given state $\theta = 1$, the equilibrium actions $b_1, \dots, b_N$ are i.i.d. Bernoulli random variables that take values 1 and 0 with probabilities $\frac{1}{2} + t^1_N$ and $\frac{1}{2} - t^1_N$ respectively. We have the mean $\mu^1_N \equiv \frac{1}{2} + t^1_N$ and the variance $(\sigma^1_N)^2 \equiv \frac{1}{4} - (t^1_N)^2$. Hence,

$$\Pr\left(\bar{b}_N > \frac{1}{2} \mid \theta = 1\right) = 1 - \Pr\left(\frac{\bar{b}_N - \mu^1_N}{\sigma^1_N/\sqrt{N}} \leq -\frac{\sqrt{N}t^1_N}{\sigma^1_N} \mid \theta = 1\right).$$

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

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

where Φ is the standard normal cdf. Since $\mathbb{E}[|b_i - \mu_N^1|^3 \mid \theta = 1] \leq \frac{1}{8}$, the RHS vanishes as $N \to \infty$. 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,

$$\lim_{N\to\infty}\left|\Pr\left(\bar{b}_N > \frac{1}{2} \mid \theta = 1\right) - \Phi\left(\frac{\sqrt{N}t_N^1}{\sigma_N^1}\right)\right| = 0. \tag{23}$$

Next, we consider $\Pr(\bar{b}_N < \frac{1}{2} \mid \theta = 0)$ in (22). 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,

$$\lim_{N\to\infty}\left|\Pr\left(\bar{b}_N < \frac{1}{2} \mid \theta = 0\right) - \Phi\left(\frac{\sqrt{N}t_N^0}{\sigma_N^0}\right)\right| = 0. \tag{24}$$

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 15.) Rewrite the biased-logit formula (7) as follows:[13]

$$\begin{aligned}
\lambda \ln \frac{Q_N^*(1 \mid 1)}{Q_N^*(0 \mid 1)} - \lambda \ln \frac{q_N^*}{1 - q_N^*} &= \binom{2n}{n}(Q_N^*(1 \mid 1))^n (Q_N^*(0 \mid 1))^n, \\
\lambda \ln \frac{Q_N^*(0 \mid 0)}{Q_N^*(1 \mid 0)} - \lambda \ln \frac{1 - q_N^*}{q_N^*} &= \binom{2n}{n}(Q_N^*(1 \mid 0))^n (Q_N^*(0 \mid 0))^n,
\end{aligned}$$

where we have $q_N^* \in (0, 1)$ for any large N because $\lim_{N\to\infty} q_N^* = \frac{1}{2}$ 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 $f : (-\frac{1}{2}, \frac{1}{2}) \to \mathbb{R}$ defined by $f(t) = \ln\left(\frac{1/2+t}{1/2-t}\right)$, we rewrite the above equations as

$$\begin{aligned}
\lambda f(t_N^1) - \lambda f(t_N) &= \binom{2n}{n}\left(\frac{1}{2} + t_N^1\right)^n \left(\frac{1}{2} - t_N^1\right)^n, \\
\lambda f(t_N^0) + \lambda f(t_N) &= \binom{2n}{n}\left(\frac{1}{2} + t_N^0\right)^n \left(\frac{1}{2} - t_N^0\right)^n.
\end{aligned}$$

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

[13] Here is how we derive these equations. We only consider the case of state $\theta = 1$, as the case of state $\theta = 0$ is analogous. Since $Q_N^*(0 \mid 1) = 1 - Q_N^*(1 \mid 1)$, it follows that $\frac{Q_N^*(1|1)}{Q_N^*(0|1)} = \frac{q_N^*}{1-q_N^*} \mathrm{e}^{\Pi_N(1)/\lambda}$. Taking the logarithm on both sides, we obtain the desired equations.

we rearrange the terms to obtain that

$$\frac{\lambda T_N^1}{\frac{1}{4}-(\tau_N^1)^2}-\frac{\lambda T_N}{\frac{1}{4}-(\tau_N)^2}=\binom{2n}{n}\frac{\sqrt{n}}{2^{2n}}\sqrt{\frac{N}{n}}\left(1-\frac{(2T_N^1)^2}{N}\right)^n, \tag{25}$$

$$\frac{\lambda T_N^0}{\frac{1}{4}-(\tau_N^0)^2}+\frac{\lambda T_N}{\frac{1}{4}-(\tau_N)^2}=\binom{2n}{n}\frac{\sqrt{n}}{2^{2n}}\sqrt{\frac{N}{n}}\left(1-\frac{(2T_N^0)^2}{N}\right)^n, \tag{26}$$

where we denote $T_N^\theta \equiv \sqrt{N} t_N^\theta$ and $T_N \equiv \sqrt{N} t_N$ for brevity.

As $N \to \infty$, we have $\tau_N^\theta \to 0$ and $\tau_N \to 0$ because $t_N^\theta \to 0$ and $t_N \to 0$.

Step 3 We take the limit of (25) and (26). Consider any subsequence of the sequence $\{N\}$ such that $\lim_{N\to\infty} T_N^1 = T^1 \in \mathbb{R}$ and $\lim_{N\to\infty} T_N^0 = T^0 \in \mathbb{R}$. We consider other subsequences in Step 4. By (21), the subsequence $\{T_N\}$ converges to T such that

$$T = \mu(1)T^1 - \mu(0)T^0. \tag{27}$$

Take the limit of (25) and (26) along the subsequence. Then,

$$\begin{aligned}\lim_{N\to\infty} \text{LHS of (25)} &= 4\lambda(T^1 - T),\\ \lim_{N\to\infty} \text{LHS of (26)} &= 4\lambda(T^0 + T).\end{aligned} \tag{28}$$

By Stirling's formula,

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

For $N = 2n+1$, $\lim_{N\to\infty}\sqrt{N/n} = \sqrt{2}$ and[14]

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

Hence,

$$\begin{aligned}\lim_{N\to\infty} \text{RHS of (25)} &= 2\phi(2T^1),\\ \lim_{N\to\infty} \text{RHS of (26)} &= 2\phi(2T^0),\end{aligned} \tag{29}$$

where ϕ is the standard normal pdf. By (27), (28), and (29),

$$\begin{aligned}\lambda\mu(0)(2T^1 + 2T^0) &= \phi(2T^1),\\ \lambda\mu(1)(2T^1 + 2T^0) &= \phi(2T^0).\end{aligned} \tag{30}$$

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

By (23) and (24),

$$\lim_{N\to\infty} \Pr\Big(\bar{b}_N > \frac{1}{2} \mid \theta = 1\Big) = \lim_{N\to\infty} \Phi\left(\frac{T_N^1}{\sigma_N^1}\right) = \Phi(2T^1),$$
$$\lim_{N\to\infty} \Pr\Big(\bar{b}_N < \frac{1}{2} \mid \theta = 0\Big) = \lim_{N\to\infty} \Phi\left(\frac{T_N^0}{\sigma_N^0}\right) = \Phi(2T^0),$$

where Φ is continuous, and $T_N^\theta \to T^\theta$ and $\sigma_N^\theta \to \frac{1}{2}$ for each $\theta = 0, 1$. Now we take the limit $N \to \infty$ in (22). Then,

$$\lim_{N\to\infty} \Pr(u(\bar{b}_N, \theta) = 1) = \mu(1)\Phi(2T^1) + \mu(0)\Phi(2T^0).$$

Letting $t^1 = 2T^1$ and $t^0 = 2T^0$ in this equation and equations (30), we have the desired result.[15]

Step 4 Consider any subsequence of the sequence $\{N\}$ along which either $\{T_N^1\}$ or $\{T_N^0\}$ or both diverge. There are three cases to consider. First, suppose that along the subsecquence, one of $\{T_N^1\}$ and $\{T_N^0\}$ converges (to a finite value) and the other one diverges to $\pm\infty$. Assume, without loss, that $T_N^1 \to T_1 \in \mathbb{R}$ but $T_N^0 \to \pm\infty$. By (21), $\{T_N\} \to \mp\infty$. Then, the LHS of (25) diverges to $\pm\infty$, but the RHS converges to finite values, a contradiction. Second, suppose that along the subsecquence, both $\{T_N^1\}$ and $\{T_N^0\}$ diverge to $+\infty$. By (21), $\{T_N\}$ may converge or diverge. If $\{T_N\}$ converges then on both (25) and (26), the LHS diverge to $+\infty$ but the RHS converge to finite values, a contradiction. If $\{T_N\}$ diverges then the LHS on either (25) or (26) diverges but the RHS on both converge, a contradiction. Third, suppose that along the subsecquence, one of $\{T_N^1\}$ and $\{T_N^0\}$ diverges to $+\infty$ and the other diverges to $-\infty$. If $T_N^1 \to +\infty$ and $T_N^0 \to -\infty$ then $\Phi(T_N^1/\sigma_N^1) \to 1$ and $\Phi(T_N^0/\sigma_N^0) \to 0$ in (23) and (24), which implies that $\Pr(u(\bar{b}_N, \theta) = 1) \to \mu(1)$ in (22). Similarly, if $T_N^1 \to -\infty$ and $T_N^0 \to +\infty$ then $\Pr(u(\bar{b}_N, \theta) = 1) \to \mu(0)$.

A.9 Proposition 1

In an election with a poll, the probability of correct choice at the informative equilibrium is $\frac{e^{1/\lambda}}{1+e^{1/\lambda}}$ for any N by Theorem 1. In an election without a poll, the probability of correct choice at any (informative or uninformative) equilibrium converges to either $\mu(1)$, $\mu(0)$, or $\mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)$ as $N \to \infty$ by Lemma 6. It then suffices to show that there exists a small $\epsilon > 0$ such that for any $\lambda > 0$ and any μ such that $|\mu(1) - \frac{1}{2}| < \epsilon$,

$$\frac{e^{1/\lambda}}{1 + e^{1/\lambda}} > \max\Big\{\mu(1), \mu(0), \mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)\Big\}, \tag{31}$$

[15] In this analysis, we assume that $t_N^1 \neq 0$ and $t_N^0 \neq 0$. In this footnote, we discuss the other cases. If $t_N^1 = 0$ for all sufficiently large N then (23) is reduced to $\lim_{N\to\infty} \Pr(\bar{b}_N > \frac{1}{2} \mid \theta = 1) = \frac{1}{2}$. This case is included in the desired result (8) since $t^1 = 2\lim_{N\to\infty} \sqrt{N} t_N^1 = 0$. Similarly, if $t_N^0 = 0$ for all sufficiently large N then (24) is reduced to $\lim_{N\to\infty} \Pr(\bar{b}_N < \frac{1}{2} \mid \theta = 0) = \frac{1}{2}$, and this case is included in the desired result (8). Lastly, if $t_N^\theta = 0$ 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, $t_N^\theta \neq 0$.

where (t^1, t^0) is a solution to $\lambda\mu(0)(t^1 + t^0) = \phi(t^1)$ and $\lambda\mu(1)(t^1 + t^0) = \phi(t^0)$.

Since (t^1, t^0) is continuous in μ, it follows that $\mu(1)\Phi(t^1) + \mu(0)\Phi(t^0)$ is continuous in μ. To show the above result, it suffices to prove the result for the symmetric prior μ, $\mu(1) = \frac{1}{2}$.

Lemma B. *For each N, any election $\mathcal{Q}_N$ with the symmetric prior μ 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*

$$\lim_{N\to\infty} \Pr(u(\bar{b}_N, \theta) = 1) = \Phi(t_\lambda),$$

where $t_\lambda > 0$ is the unique solution to equation $\lambda t = \phi(t)$.

Proof. Under the symmetric prior μ, 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. By symmetry, $Q^*_N(1 \mid 1) = Q^*_N(0 \mid 0) > \frac{1}{2}$ and thus $t^1_N = t^0_N > 0$ and $T^1_N = T^0_N$. Since their limits coincide (i.e., $T^1 = T^0$, where $T^1 = \lim_{N\to\infty} T^1_N$ and $T^0 = \lim_{N\to\infty} T^0_N$), we have $t \equiv t^1 = t^2$, where $t^1 = 2T^1$ and $t^2 = 2T^0$. Substituting it into Lemma 6, we have the limit, $\lim_{N\to\infty} \Pr(u(\bar{b}_N, \theta) = 1) = \Phi(t)$, and the constraint, $\lambda t = \phi(t)$. ■

By Lemma B, if the prior μ is symmetric, (31) is equivalent to $\frac{e^{1/\lambda}}{1+e^{1/\lambda}} > \max\{\frac{1}{2}, \Phi(t_\lambda)\}$. Since $\Phi(t) > \frac{1}{2}$ for any $t > 0$, this is equivalent to $\frac{e^{1/\lambda}}{1+e^{1/\lambda}} > \Phi(t_\lambda)$. Hence, the proof is completed by the following lemma:

Lemma C. *For any $\lambda > 0$, $\frac{e^{1/\lambda}}{1+e^{1/\lambda}} > \Phi(t_\lambda)$.*

Proof. See Online Appendix B. ■

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.

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.

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.

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.

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.

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.