Voter Attention and Distributive Politics *

Carl Heese †

Click here for latest version

November 5, 2019

Abstract

I study the role of heterogeneous attention among voters for distributive politics. Groups in the population pay systematically different attention to politics and acquire different levels of information. I study the effects of heterogenous attention when a reform may benefit one group at the expense of others. In the benchmark, when the information of voters is exogenous, a median voter theorem holds and a welfare-enhancing reform might not be adopted if it is not preferred by a majority. I show that the (endogenously) heterogeneous attention shifts election outcomes into a direction that is welfare-improving. Even when a welfare-enhancing reform is not preferred by a majority ex-post, under certain conditions, there are equilibria where the reform is adopted. The key driver of the results is that voters who are more severely affected by a proposed reform will pay more attention, consistent with empirical studies ("issue publics hypothesis", Converse (1964)). This information advantage translates into voting power precluding the majority from exerting its dominance.

*I am grateful for helpful discussions with Nageeb Ali, Daniel Krähmer, Stephan Lauermann, Philipp Strack and Thomas Tröger as well as comments from audiences at Bonn, Yale, the CRC TR 224 Junior Conference in April 2018, SAET 2018, CED 2019, SAET 2019, Stony Brook 2019. This work was supported by a grant from the European Research Council (ERC 638115) and the CRC TR224, Project B03. This draft is preliminary. Comments and suggestions are welcome.

†University of Bonn, Department of Economics, heese@uni-bonn.de.

1 Introduction

A rich literature in distributive politics seeks to understand when and if governments or elections select policies that maximize social welfare.[1] I introduce a novel consideration into this discussion —heterogenous attention to politics— based on the empirical observation that voters that care more about a political issue, will seek out more information about it ("issue publics hypothesis", Converse (1964)).[2]

I propose a model that allows to carve out the relation between the information level of different voter groups with conflicting interests and their voting power. The main result shows that, for a large class of settings, there are equilibria where the first-best outcome is elected, even if the outcome is not preferred by a majority of the voters ex-post (Theorem 1). In such situations, a minority of the voters is more severely affected by the reform, but will be better informed about it, and is able to translate the information advantage into voting power, thereby shifting the election outcome towards efficiency.

Examples of reforms with uncertain distributive consequence are numerous: a trade reform opens new markets for exporting firms but threatens the prospects in other sectors, [3] a public health policy reform makes certain treatments more accessible to citizens, while it implies price increases for a range of pharmaceuticals needed by other citizens, and a new tax policy reduces the tax burden for some citizens, but elevates it for others. In all these examples, most voters are ex ante uninformed about the consequences of the reform, e.g. which sectors gain from a trade reform, or what the complex implications of a tax reform are. However, they hold have private information about their exposure to the proposed reform, that is about the magnitude of their preference intensities: more affluent citizens have more money at stake, and medically dependent citizens are more strongly affected by changes in public health policy.

[1] See e.g. Fernandez & Rodrik (1991), Alesina & Rodrik (1994) and Persson & Tabellini (1994).

[2] See e.g. Carmines & Stimson (1980), Miller & Krosnick (2000), and Iyengar *et al.* (2008).

[3] There is a large literature on the political economy of trade policies, see e.g. Mayer (1984).

I consider a modified version of the canonical setting by Feddersen & Pesendorfer (1997). In this modified version, the voters information about the policies is endogenous.[4] There are two possible policies, a reform and a status quo. Voters' preferences over policies are heterogeneous and depend on an unknown, binary state in a general way (some voters may prefer the reform only in the first state, some prefer the reform only in the second state, and some "partisans" may prefer one of the policies independently of the state). The preferences are each voters' private information. In addition, all voters can receive information about the state in the form of a noisy signal and each voter freely chooses the precision of his private signal. More precise information is more costly. The election determines the outcome by a simple majority rule. Feddersen & Pesendorfer (1997) have shown that, when voters receive conditionally i.i.d. signals of some exogenous quality and preferences are "monotone",[5] in all equilibria of large elections the outcome preferred by the median voter is elected. In many situations, where voters have conflicting interests, this is not the first-best outcome: for example, when 51% of the citizens benefit from a reform marginally, but at the same time 49% of the citizens are impaired by it severely.

The main analysis describes the political power of the heterogenous voters. All equilibria can be represented by an index rule (Lemma 8). The *power index* of a group of voters sharing common interests is a relative measure of its electoral power and is, in particular, proportional to the size of the group and increasing in the welfare that is at stake for the group. When relevant political information can be acquired at cost that are "not too high", there is an equilibrium where competitive information acquisition shapes the election and where the election yields the outcome preferred by the group with the higher power index.[6] I will show that, when information cost are not too low,

[4] Similar to this paper, Ali *et al.* (2018) transports the informational approach to elections (Austen-Smith & Banks (1996), Feddersen & Pesendorfer (1997)) to the literature on distributive politics.

[5] The preference distribution of the voters is "monotone" if a higher belief in the first state entails that more voters prefer the reform. I study also non-monotone preferences.

[6] What will matter is how fast cost go to zero when a voter chooses an arbitrarily uninformative signal. The critical condition is that the first three derivatives of the cost

these equilibria lead to first-best outcomes; to put it differently, the information cost screen the voters intensities appropriately.

In the second part, I deepen the equilibrium analysis. By doing so, I point out determinants of efficiency and the political power of voter groups that are novel to the existing literature. First, I analyze the robustness of the equilibria with first-best outcomes when varying the voters cost of acquiring relevant information about the policy consequences. The more similar is the welfare at stake, that is the aggregate intensities, of the citizens who benefit relative to those who are harmed by a reform, the smaller the set of cost functions for which there are equilibria with first-best outcomes (Theorem 3). Intuitively, more similar intensities make it more difficult for the information cost to screen voter intensities appropriately. Second, I show that the electoral power of a voter group depends on the polarization of the preference intensities within the group. Whatever the distribution of the preferences in the population is, for any voter group sharing a common interest, there is a mean preserving-spread of the preference intensities within the group such that for almost no cost function, there is an equilibrium leading to the outcome preferred by the group in both states (Theorem 4).

In the last part of the paper, I study the situations where the voters differ not only in their exposure to the proposed reform, but also in their ability to access and interpret political information. Formally, the voters are subject to different cost functions. First, I provide an equivalence result: the extended setting with heterogenous cost is equivalent to a setting with homogenous cost: intuitively, the cost of information and preference intensities are strategically equivalent such that differences in cost translate into differences in intensities. My previous efficiency results therefore extend when the cost types are independent of the preference intensities (Theorem 7). Second, I vary the richness of the information choice of the voters in the extended setting and show that coarsening the voters information choice may have positive welfare effects. I compare the situation where the voters have only access to a binary set of signals, that is, voters choose between a given informative signal and an unin-

function are zero at the precision of the uninformative signal.

formative signal, to the previous setting where voters have access to a rich set of signals and can freely choose their signal precision without constraints. In the coarse setting, there are always equilibria where the welfare-maximizing outcome is elected in all states (Theorem 8), unlike in the setting with rich information choice where this depends on how similar the welfare at stake is for the citizens who benefit relative to those who are harmed by a reform. Hence This is particularly surprising, since a richer choice set for the voters should intuitively facilitate an appropriate screening of preference intensities in equilibrium, and not prevent it.

The paper contributes to the literature on information aggregation in large elections. Condorcet's Jury Theorem (1785) says that if voters have common interests but information is dispersed in the electorate, majority rule results in socially optimal outcomes. Information aggregates in the sense that electoral outcomes correspond to the choices of a fully informed welfare-maximizing social planner. Austen-Smith & Banks (1996), Feddersen & Pesendorfer (1998) have established a "modern" version of the Condorcet Jury Theorem in a setting where voters are strategic. However, as argued, elections might fail to elect socially optimal outcomes when voters have conflicting interests, for example when 51% of the citizens benefit from a reform marginally, but at the same time 49% of the citizens are impaired by it severely. This paper points at an empirical observation that has been overlooked, namely, that the dispersion of the voters' information is endogenous, and I show how this can imply socially optimal outcomes, independent of the distribution of the voters' preference intensities.

The paper contributes to the literature on elections with costly information acquisition by studying a general setup that, in particular, allows the voters to have conflicting interests like in distributive politics. The previous literature (e.g. as in Martinelli (2006) and Oliveros (2013)) has studied situations where all voters have common interests. For this case, I generalize the results of the literature by characterizing all equilibria of the voting game and by providing an analysis for general continuous preference distributions.[7]

[7]The settings in **?** and in **?** are a special case of my model.

The paper is related to work on elections with voting cost and vote-buying. Krishna & Morgan (2011) and Krishna & Morgan (2015) have shown that elections yield first-best outcomes when voting is voluntary and costly. These results are analogous to my findings for the setting with a coarse, binary information choice. The main model with a rich information choice is more closely related to the literature on vote buying. Lalley & Weyl (2018) have shown that equilibrium outcomes of a large electorate are first-best when each voter can buy any number of votes at a total price quadratic in the number bought. Similarly, in this model, the analysis will show that when the cost of information is arbitrarily close to 'cubic' as a function of the precision of the signal, there are equilibria with first-best outcomes for almost all preference distributions.

The paper contributes to the literature on efficient mechanisms with information acquisition: Bergemann & Välimäki (2002) have shown that in private-value mechanism design settings with transferable utility and information acquisition, Vickrey-Clarke-Groves mechanisms implement first-best outcomes. I show that in a setting where utility is non-transferable, majority rules may implement first-best outcomes and describe conditions when this is possible.

The rest of the paper is organized as follows. Section 1.1 provides a simple example that lays out the logic of the main result. Section 2 presents the model and the preliminary analysis. Section 3 discusses the power (index) of voter groups. Section 4 shows the existence of equilibria with first-best outcomes and discusses their robustness and the role of preference polarization. Section 5 characterizes all other equilibria and, in particular, shows existence of another equilibrium that converges to voting according to the prior belief" about the unknown state. In this equilibrium, the outcome that is preferred by the majority of the voters given the prior belief, is elected (Theorem 5). Section 6 studies heterogenous cost and a setting where the information choice of the voters is coarse. Section 7 discusses the relation to the literature further.

1.1 Two-type example

There are $2n+1$ voters. With probability $\lambda > \frac{1}{2}$, a voter is *aligned* and prefers the reform A over the status quo B only in α and B over A in β. Otherwise, a voter is *contrarian* and prefers A only in β. A contrarian voter gets utility of 1 when her preferred policy is elected and an aligned voter gets a small utility of $\epsilon > 0$ when her preferred policy is elected. Voters have a binary choice, to get a private, perfect signal about the state at a given cost $c > 0$ or an uninformative signal at no cost. The common prior about the state is uniform, i.e. $\Pr(\alpha) = \frac{1}{2}$. To maximize welfare, the election should implement what the contrarians want since they care much more.

Consider three scenarios: zero, intermediate and high cost. When cost are zero, i.e. $c = 0$, all voters get perfectly informed about the state and the outcome preferred by the median voter is elected in each state. When cost are very high, e.g. $c > 1$, nobody gets informed and the policy elected is independent of the state. When cost come from a certain intermediate range, they screen types such that only types with high intensities get informed. In fact, for such intermediate cost, there is an equilibrium where the aligned vote for each policy with the same probability and the contrarians vote for their preferred outcome in each state. Thus, when the electorate grows large, i.e. as $n \to \infty$, the outcome preferred by the contrarians, which represent a minority in expectation, is elected in both states and this outcome maximizes welfare.[8]

To see why such an equilibrium exists, note that without the private signal, a citizen is indifferent between voting for either of the policies: the event piv in which the citizen's vote affects the outcome, is equally likely in each state in the candidate equilibrium; in one state the reform wins with a margin of $(1-\lambda)$ and in the other state the reform loses with a margin $(1 - \lambda)$ in expectation, meaning that in each state, the election is equally close to being tied, implying $\Pr(\alpha|\text{piv}) = \frac{1}{2}$. When cost are intermediate, i.e. when $\epsilon \Pr(\text{piv}) < c < \Pr(\text{piv})$,

[8] This follows from a law of large number argument since the expected vote share in α is strictly smaller than $\frac{1}{2}$ and the expected vote share in β is strictly larger than $\frac{1}{2}$. I do not discuss the welfare effects of the cost for the example since it turns out that the aggregate cost are arbitrarily small in the equilibria of the main result when the electorate is large.

only the contrarians would like to buy the signal and the aligned voters hold a belief that makes them completely indifferent between policy A and B. Note that the closer the utility at stake for the aligned and the contrarians, the smaller the range of intermediate cost that screen the types.

2 Model

There are $2n+1$ voters (or citizens), two policies A and B, and two states of the world $\omega \in \{\alpha, \beta\} = \Omega$. The prior probability of α is $\Pr(\alpha) \in (0,1)$.

Voters have heterogeneous and state-dependent preferences. A voter's preference is described by a type $t = (t_\alpha, t_\beta)$, with $t_\omega \in [-1,1]$ the utility of A in ω. The utility of B is normalized to zero, so that t_ω is the difference of the utilities of A and B in ω. The types are identically distributed across voters according to a cumulative distribution function $H : [-1,1]^2 \to [0,1]$ that has a continuous density h. The own type is private information of the voter. Each voter observes a binary signal $s \in \{a, b\}$. The observed signal is the private information of the voter as well. The joint distribution of the type and the signal of a voter is independent of the distribution of the signals and the types of the other voters.

A strategy $\sigma = (x, \mu)$ of a voter consists of a function $x : [-1,1]^2 \to [0, \frac{1}{2}]$ and a function $\mu : [-1,1]^2 \times \{a,b\} \to [0,1]$. Here, $\frac{1}{2} + x(t)$ is the *precision* of the private signal of the voter of type t, i.e. $\frac{1}{2} + x(t) = \Pr(a|\alpha) = \Pr(b|\beta)$ and $\mu(t,s)$ is the probability that a voter of type t with signal s votes for A.

There is a strictly increasing, convex and twice continuously differentiably *cost function* $c : [0, \frac{1}{2}] \to \mathbb{R}_+$ and when choosing precision x, the voter bears a cost $c(x)$ where $c(0) = 0$. There is a $d > 1$ such that $\lim_{x\to\infty} \frac{c'(x)}{x^{d-1}} \in \mathbb{R}$. [9] I consider only non-degenerate strategies.[10]

[9]It will be a trivial insight from the preliminary results in the next section that without the condition $d > 1$, no voter acquires any information in equilibrium when n is sufficiently large; see in particular (19).

[10]A strategy σ is *degenerate* if $\mu(t,s) = 1$ for all (t,s) or if $\mu(t,s) = 0$ for all (s,t). When all voters follow the same degenerate strategy and there are at least three voters, if one

The voting game is as follows. First, nature draws the state and the profile of types $\mathbf{t}$ according to H. Second, after observing her type, each voter chooses a precision. Then, private signals realize. After observing her private signal, each voter simultaneously submits a vote for A or B. Finally, the submitted votes are counted and the majority outcome is chosen. This defines a Bayesian game. I analyze the Bayes-Nash equilibria of this game in symmetric strategies, henceforth called *equilibria*.

2.1 Preliminaries

2.1.1 Aggregate Preferences

A central object of the analysis is the *aggregate preference function*

$$\Phi(p) := \mathrm{Pr}_G(\{t : p \cdot t_\alpha + (1-p) \cdot t_\beta \geq 0\}), \tag{1}$$

which maps a belief $p \in [0,1]$ to the probability that a random type t prefers A under p. Figure 1 illustrates Φ. The dashed (blue) line corresponds to the plane of types $t = (t_\alpha, t_\beta)$ that are indifferent between policy A and policy B when holding the belief p. Voters having types to the north-east prefer A given p. The indifference plane has a slope of $\frac{-p}{1-p}$ and a change in p corresponds to a rotation of it. Given that H has a continuous density, Φ is continuously differentiable in p.

Voters having types t in the north-east quadrant prefer A for all beliefs and voters having types t in the south-west quadrant always prefer B (*partisans*). Voters having types t in the south-east quadrant prefer A in state α and B in β (*aligned voters*) and voters having types t in the north-west quadrant prefer B in state α and A in β (*contrarian voters*). I assume that

$$\Phi(0) < \frac{1}{2}, \text{ and } \Phi(1) > \frac{1}{2} \tag{2}$$

voter deviates to any other strategy, the outcome is the same. So, the degenerate strategies with $x = 0$ are trivial equilibria.

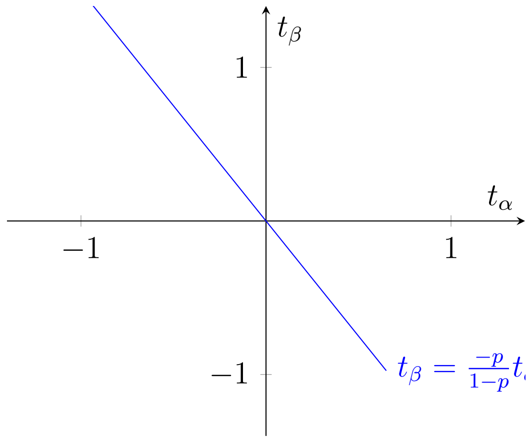

Figure 1: For any given belief $p = \Pr(\alpha) \in (0,1)$, the plane of types t with threshold of doubt $y(t) = p$ is $t_\beta = \frac{-p}{1-p}t_\alpha$.

such that the median-voter preferred outcome is A in α and B in β. In particular, this excludes the cases where there is a majority of partisans for one policy in expectation. Henceforth, I call distributions H that have a continuous density and satisfy (2) simply *preference distributions.* In the following, the set of aligned types is denoted $\ell = \{t : t_\alpha > 0, t_\beta < 0\}$ and the set of contrarian types is denoted $\{s = \{t : t_\alpha < 0, t_\beta > 0\}$.

2.1.2 Pivotal Voting and the Voter's Threshold of Doubt

How do the voters decide which alternative to vote for? Since the preferences of the voters are state-dependent, their beliefs about the state matter. This section shows how the aligned and contrarian voters form posterior beliefs about the state and that they follow a cutoff rule when deciding between A and B: the posterior has to exceed a type-dependent *threshold of doubt*.

From the viewpoint of a given voter and given any strategy $\sigma' = (x', \mu')$ used by the other voters, the pivotal event, piv, is the event in which the realized types and signals of the other $2n$ voters are such that exactly n of them vote for A and n for B. For this event, if she votes A, the outcome is A; if she votes B, the outcome is B. In any other event, the outcome is independent of her vote. Thus, a strategy is optimal if and only if it is optimal conditional on the pivotal event.

Take two strategies $\sigma' = (x', \mu')$ and $\sigma = (x, \mu)$. Let $\Pr(\alpha|s, \text{piv}; \sigma', n)$ denote the posterior probability of α conditional on *being pivotal* and conditional on s when the other voters use σ' and the given voter uses σ. I omit the dependence on σ since in most parts of the paper $\sigma = \sigma'$. Given some x, the function μ is part of a best response $\sigma = (x, \mu)$ to σ' if and only if for all $t = (t_\alpha, t_\beta)$ and for the signal precision $x(t)$,

$$\Pr(\alpha|s, \text{piv}; \sigma', n) \cdot t_\alpha + (1 - \Pr(\alpha|s, \text{piv}; \sigma', n)) \cdot t_\beta > 0 \Rightarrow \mu(s, t) = 1, \quad (3)$$

and

$$\Pr(\alpha|s, \text{piv}; \sigma', n) \cdot t_\alpha + (1 - \Pr(\alpha|s, \text{piv}; \sigma', n)) \cdot t_\beta < 0 \Rightarrow \mu(s, t) = 0, \quad (4)$$

that is, a voter supports A if the expected value of A conditional on being pivotal and s is strictly positive and supports B otherwise. Note that indifference holds only for a set of types that has zero measure since the type distribution H is continuous. For all other types, the best response is pure. It follows that there is no loss of generality to consider pure strategies with $\mu(t, s) \in \{0, 1\}$ for all (s, t).[11]

The relative size of the intensities t_α and t_β pins down how certain an aligned type must be of the state being α to choose A under the best response. Similarly, the relative size of the intensities t_α and t_β pins down how certain a contrarian type must be of the state being α to choose B under the best response. For any voter any aligned or contrarian type t, I call

$$y(t) = \frac{-t_\beta}{t_\alpha - t_\beta}. \quad (5)$$

the voter's *threshold of doubt*. Given (3) and (4), μ is part of a best response $\sigma = (x, \mu)$ to σ' if and only if for the signal precision $x(t)$,

[11] When writing "*for all*," I ignore zero measure sets here and in the following.

$$\forall t \in \ell : \Pr(\alpha|s, \text{piv}; \sigma', n) > y(t) \Rightarrow \mu(t,s) = 1, \tag{6}$$
$$\Pr(\alpha|s, \text{piv}; \sigma', n) < y(t) \Rightarrow \mu(t,s) = 0, \tag{7}$$

and

$$\forall t \in s : \Pr(\alpha|s, \text{piv}; \sigma', \sigma, n) > y(t) \Rightarrow \mu(t,s) = 0, \tag{8}$$
$$\Pr(\alpha|s, \text{piv}; \sigma', \sigma, n) < y(t) \Rightarrow \mu(t,s) = 1, \tag{9}$$

Aligned and contrarians simply compare the posterior against their threshold of doubt to decide which policy to vote for.

2.1.3 Likelihood of the Pivotal Event

Intuitively, when the electorate grows large, the likelihood that a given voter is pivotal is increasingly small. In fact, for any strategy σ used by the voters, the likelihood of the pivotal event goes to zero as $n \to \infty$.
Take any strategy $\sigma = (x, \mu)$ of the voters. The probability that a voter of random type supports A in state $\omega \in \{\alpha, \beta\}$ is

$$\begin{aligned} q(\alpha;\sigma) &= \Pr_H(\{t : x(t) = 0\}) \int_t \mu(a,t) dHt(t|x(t) = 0) \qquad (10) \\ &+ \Pr(\{t : x(t) > 0\}) \Big[\int_t (\frac{1}{2} + x(t))\mu(t,a) + (\frac{1}{2} - x(t))\mu(t,b) dHt(t|x(t) > 0 \Big], \\ q(\beta;\sigma) &= \Pr_H(\{t : x(t) = 0\}) \int_t \mu(t,a) dHt(t|x(t) = 0) \qquad (11) \\ &+ \Pr(\{t : x(t) > 0\}) \Big[\int_t (\frac{1}{2} - x(t))\mu(t,a) + (\frac{1}{2} + x(t))\mu(t,b) dHt(t|x(t) > 0 \Big]. \end{aligned}$$

I refer to $q(\omega;\sigma)$ also as the *(expected) vote share* of A in ω. Since the type and the signal of a voter is independent of the types and signals of the other voters, the probability of a tie in the vote count is

$$\Pr(\text{piv}|\omega; \sigma, n) = \binom{2n}{n} (q(\omega;\sigma))^n (1 - q(\omega;\sigma))^n . \tag{12}$$

Given (12), the likelihood of the pivotal event is non-zero for any (non-degenerate) strategy. Hence, all partisans use the dominant strategy to choose an uninformative signal and vote their preferred policy. Given this simple behaviour of the partisans, in the following, oftentimes when I describe the behaviour of 'voters' I mean voters of aligned and contrarian types, slightly abusing the wording.

I use a Stirling approximation of the binomial coefficient and (12) to obtain[12] [13]

$$\Pr\left(\mathrm{piv}|\omega;\sigma,n\right) \approx 4^n (n\pi)^{-\frac{1}{2}}\Big[q(\omega;\sigma)(1-q(\omega;\sigma)\Big]^n. \tag{13}$$

Since the maximum of the function $q(1-q)$ on $[0,1]$ is $\frac{1}{4}$, for any strategy σ, $\lim_{n\to\infty}\Pr\left(\mathrm{piv}|\omega;\sigma,n\right)=0$ for all $\omega\in\{\alpha,\beta\}$.

In the next two sections, I characterize the optimal information acquisition behaviour of the voters. First, in Section 2.1.4, I characterize the choices of voter types t that acquire information, i.e. for which $x(t)>0$. In Section 2.1.6, I characterize which types acquire any information at all.

2.1.4 Choices of Informed Voters

Let $\sigma=(x,\mu)$ be a best response to σ'.

Since signal a is indicative of α and b of β, voters with a signal a believe state α to be more likely than voters with a signal b. In fact, given any $x>0$, the likelihood ratios are ordered as

$$\Pr\left(\alpha|a,\mathrm{piv};\sigma,n\right) > \Pr\left(\alpha|b,\mathrm{piv};\sigma,n\right). \tag{14}$$

To see why, note that the posterior likelihood ratio of the states conditional

[12]The notation $x_n \approx y_n$ describes that two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are *asymptotically equivalent* in the following sense: $\lim_{n\to\infty}\frac{x_n}{y_n}=1$.

[13]Stirling's formula yields $(2n)! \approx (2\pi)^{\frac{1}{2}}2^{2n+\frac{1}{2}}n^{2n+\frac{1}{2}}e^{-2n}$ and $(n!)^2 \approx (2\pi)n^{2n+1}e^{-2n}$. Consequently, $\binom{2n}{n} \approx (2\pi)^{-\frac{1}{2}}2^{2n+\frac{1}{2}}n^{-\frac{1}{2}} = 4^n(n\pi)^{-\frac{1}{2}}$.

on a signal $s \in \{a, b\}$ and the event that the voter is pivotal is

$$\frac{\Pr(\alpha|s, \text{piv}; \sigma, n)}{\Pr(\beta|s, \text{piv}; \sigma', n)} = \frac{\Pr(\alpha)}{\Pr(\beta)} \frac{\Pr(\text{piv}|\alpha; \sigma', n)}{\Pr(\text{piv}|\beta; \sigma', n)} \frac{\Pr(s|\alpha; \sigma)}{\Pr(s|\beta; \sigma)}, \tag{15}$$

if $\Pr(\text{piv}|\beta; \sigma', n) > 0$, where I used the conditional independence of the types and signals of the other voters from the signal of the given voter. The order of the likelihood ratios follows from $\Pr(a|\alpha; \sigma) = \Pr(b|\beta; \sigma) = \frac{1}{2} + x$.
Suppose that a voter acquires information, i.e. $x(t) > 0$, and suppose that she votes for the same policy after both signals, i.e. $\mu(t, a) = \mu(t, b)$. Then, she would be strictly better off by not paying for the information and simply voting for this policy. Therefore, (6)-(9), and (14) together imply the following result.

Lemma 1 *Aligned types that acquire information vote A only after a, and contrarian types that acquire information vote A only after b:*

$$\forall t \in \ell : x(t) > 0 \Rightarrow \mu(t, a) = 1 \ \textit{and} \ \mu(t, b) = 0, \tag{16}$$

$$\forall t \in \mathscr{s} : x(t) > 0 \Rightarrow \mu(t, a) = 0 \ \textit{and} \ \mu(t, b) = 1. \tag{17}$$

Let us turn to the question how much information voters acquire when they choose to acquire any at all. Recall that a voter only benefits from better information in the pivotal event when her vote affects the outcome of the election. A voter who chooses precision $x > 0$ is both pivotal and receives the 'correct' signal about the state with probability

$$\Pr(\text{piv}; \sigma', n)(\frac{1}{2} + x) \tag{18}$$

When receiving the correct signal in a state ω, the voter will choose the outcome that maximizes her utility in that state, given Lemma 1, where the utility difference of the utilities of A and B in ω is $|t_\omega|$. Thus, the expected marginal benefit from a higher precision x is $\Pr(\text{piv})\text{E}_\omega(|t_\omega||\text{piv}, s = \omega; \sigma', n)$, where '$s = \omega$' denotes the event of receiving the correct signal, slightly abusing notation. Since the correct signal is equally likely to be received in each state,

$\mathrm{E}_\omega(|t_\omega||\mathrm{piv}, s = \omega; \sigma', n) = \mathrm{E}_\omega(|t_\omega||\mathrm{piv}; \sigma, n)$.
Any type that decides to acquire some information, i.e. $x > 0$, compares marginal cost and marginal benefits for her choice of x. When the electorate is large, each voter is pivotal only with a small likelihood and marginal benefits become arbitrarily small, in particular smaller than the marginal cost $c'(\frac{1}{2})$ of the maximal precision. Then, no type acquires complete information and any optimal interior signal satisfies the first-order condition that equates marginal cost and marginal benefits,

$$c'(x) = \Pr(\mathrm{piv}; \sigma', n)\mathrm{E}_\omega(|t_\omega||\mathrm{piv}; \sigma', n). \tag{19}$$

Since c is strictly convex, there is a unique solution $x^*(t; \sigma', n)$. I use the implicit function theorem and prove the following result in the Appendix.

Lemma 2 *For any n large enough, there is a function $x^*(t; \sigma', n)$ that is continuously differentiable in t and*

$$x(t) > 0 \Rightarrow x(t) = x^*(t; \sigma', n).$$

2.1.5 Coordinate Change

Voter types differ in their *(total) intensity*

$$k(t) = |t_\alpha| + |t_\beta|. \tag{20}$$

In fact, the total intensity and the threshold of doubt $y(t) = \frac{-t_\beta}{t_\alpha - t_\beta}$ relate one-to-one to the type if we restrict to either the aligned or the contrarian types. Precisely,

$$\forall t \in \ell : t_\alpha = k(t)(1 - y(t)) \text{ and } t_\beta = -k(t)y(t), \tag{21}$$

$$\forall t \in s : t_\alpha = -k(t)(1 - y(t)) \text{ and } t_\beta = k(t)y(t). \tag{22}$$

The next section describes which types acquire information. In that section, I will sometimes treat types as pairs of a treshold of doubt y tand of a total intensity k. Note that $|t_\omega|$ and therefore the solution $x^*(t)$ to the first-order condition (19) only depends on $y(t)$ and $k(t)$ and not on the group, aligned or contrarians, that the type belongs to.

2.1.6 Information Acquisition Regions

When deciding if to acquire information, voters trade off the cost of information with the benefits. The *critical types* t with $y(t) = \Pr(\alpha|\text{piv}; \sigma', n)$ are indifferent between A and B without further information, given (6) - (9). On the other hand, voters with extreme treshold of doubts $y(t) \approx 0$ or $y(t) \approx 1$ have strong preferences for one alternative without further information. Intuitively, these extreme types value information less than the critical types. The following result verifies this intuition and shows that only types with intermediate threshold of doubts acquire information, including the critical types. Second, intuitively, types with higher (total) intensities have more incentives to acquire information. The next result also shows that the interval around the critical threshold $y(t) = \Pr(\alpha|\text{piv}; \sigma', n)$ of the types that get informed only depends on the (total) intensity k.

Lemma 3 *Let σ' be a strategy with $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma', n) \in (0,1)$. When n is large enough, there are functions $\phi_g : (0,2] \to [0,1]$ and $\psi_g : (0,2] \to [0,1]$ with $\phi_g < \Pr(\alpha|\text{piv}; \sigma, n) < \psi_g$ for $g \in \{\ell, s\}$ such that for $I_g(k) = [\phi_g(k), \psi_g(k)]$ and $I_g^0(k) = (\phi_g(k), \psi_g(k))$, any best response $\sigma = (x, \mu)$ to σ', any $g \in \{\ell, s\}$, and any type $(y, k) \in g$,*

$$x(y,k) > 0 \quad \Rightarrow \quad y \in I_g^0(k), \tag{23}$$

$$x(y,k) = 0 \quad \Rightarrow \quad y \notin I_g(k). \tag{24}$$

Figure 2 illustrates the functions ϕ_g and ψ_g which limit the regions of types that acquire information under the best response. In the following, I illustrate the result with an example. The general proof is provided in the Appendix.

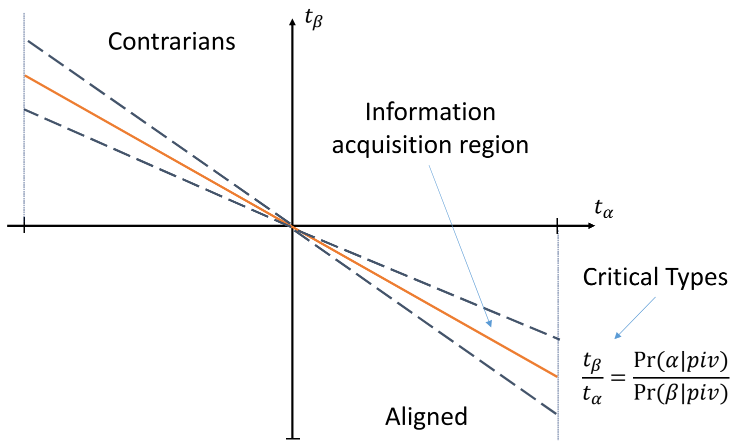

Figure 2: Information Acquisition regions of group g with boundaries the graphs of $\phi_g(k)$ and $\psi_g(k)$ (dashed lines).

Example with Uniform Types (version 1). Let the prior be uniform. All the types are aligned types. All types share the same total intensity of $k > 0$. The distribution F of the treshold of doubt is uniform.

I claim that there is a *symmetric equilibrium*, meaning that the expected margin of victory is the same in both states, i.e. $q(\alpha; \sigma_n^*) - \frac{1}{2} = \frac{1}{2} - q(\beta; \sigma_n^*)$. This implies that the election is equally close to being tied in both states, and voters do not learn anything from the pivotal event, i.e. $\Pr(\alpha|\text{piv}) = \frac{1}{2}$, given (12). The expected utility of a type who chooses precision $x > 0$ is given by the expected utility, K, from all the events when her vote does not affect the outcome, by the cost $c(x)$, and by the expected utility from the pivotal event. The expected utility from the pivotal event is

$$\Pr(\text{piv})\Big[t_\alpha \Pr(\alpha|\text{piv}; \sigma)(\frac{1}{2} + x) + t_\beta(\Pr(\alpha|\text{piv}; \sigma)(\frac{1}{2} - x)\Big]. \tag{25}$$

To see why, note that when the citizen is pivotal and votes B, she receives utility of zero. When she is pivotal and votes A, she receives utility of t_ω, depending on the state. Given Lemma 3, she only votes A after a; and she receives a in α with probability $\frac{1}{2} + x$ and in β with probability $\frac{1}{2} - x$.

Equilibrium conditions. Using that $\Pr(\alpha|\text{piv}) = \frac{1}{2}$ in any symmetric candidate equilibrium, and using $t_\alpha = k(1-y)$ and $t_\beta = -ky$ (see 21), the first-order condition, which equates marginal cost and marginal benefits is

$$\Pr(\text{piv})\frac{k}{2} = c'(x). \tag{26}$$

The first-order condition pins down the precision $x^* > 0$ chosen by the types that get partially informed; compare to Lemma 2 and (19). Which are these types? Using (21) and (25), the expected utility from an informed choice, with precision x^*, is lower than the expected utility from an uninformed vote for B if $\Pr(\text{piv})\frac{1}{2}\Big[(\frac{1}{2}+x)k(1-y) - (\frac{1}{2}-x)k_g y\Big] - c(x) \geq 0$. This is equivalent to $\Pr(\text{piv})\frac{k}{2}\Big[\frac{1}{2}(1-y) - \frac{1}{2}y\Big] + \Pr(\text{piv})\frac{k}{2}x\Big[(1-y)+y)\Big] - c(x) < 0$. Using the first-order condition, this is equivalent to $c'(x^*)(\frac{1}{2}-y) + c'(x^*)x^* - c(x^*) < 0$. Dividing by $c(x^*)$,

$$x^*(1 - \frac{c(x^*)}{x^*c'(x^*)}) \quad < \quad (y - \frac{1}{2}). \tag{27}$$

So, all types with $y > \frac{1}{2} + x^{**}$ prefer B where

$$x^{**} = x^*(1 - \frac{c(x^*)}{x^*c'(x^*)}). \tag{28}$$

Analogously, all types with $y < \frac{1}{2} - x^{**}$ prefer A. Using the symmetry of the signals and that the distribution of the threshold of doubt y is uniform, the best response σ is symmetric, i.e. the vote shares satisfy $q(\alpha;\sigma) - \frac{1}{2} = \frac{1}{2} - q(\beta;\sigma)$. An application of Kakutani's fixed point theorem implies that there is a symmetric equilibrium. In this equilibirum, only the types $y \in [\psi^g(k), \phi^g(k)]$ get partially informed with $\psi_g(k) = \frac{1}{2} - x^{**}$ and $\phi_g(k) = \frac{1}{2} + x^{**}$.

In the general case, which I prove in the Appendix, the boundaries $\psi_g(k)$

and $\phi_g(k)$ are implicitly determined by

$$\frac{1}{2} + x^{**}(\phi_g(k), k) = \frac{\Pr(\beta|\text{piv})\phi_g(k)}{\Pr(\alpha|\text{piv}(1-\phi_g(k)) + \Pr(\beta|\text{piv})\phi_g(k)}, \tag{29}$$

$$\frac{1}{2} - x^{**}(\psi_g(k), k) = \frac{\Pr(\beta|\text{piv})\psi_g(k)}{\Pr(\alpha|\text{piv}(1-\psi_g(k)) + \Pr(\beta|\text{piv})\psi_g(k)}, \tag{30}$$

where $x^{**}(y, k; \sigma, n) = x^*(y, k; \sigma, n)(1 - \frac{c(x^*(y,k;\sigma,n))}{x^*(y,k;\sigma,n)c'(x^*(y,k;\sigma,n))})$ is defined as in (28) and $x^*(y, k; \sigma, n)$ is the solution to the first-order condition (19). Note that (29) and (30) reduce to $\frac{1}{2}+x^{**}(y, k; \sigma, n) = \psi_g(k)$ and $\frac{1}{2}-x^{**}(y, k; \sigma, n) = \phi_g(k)$ when $\Pr(\alpha|\text{piv}) = \frac{1}{2}$, as in the uniform types example.

2.1.7 Inference in Large Elections

I record the intuitive fact that voters update toward the substate in which the vote share is closer to 1/2, that is, in which the election is closer to being tied in expectation.

Lemma 4 *Take any strategy σ for which $\Pr(\text{piv}|\beta;\sigma) \in (0,1)$. If*

$$\left|q(\alpha;\sigma) - \frac{1}{2}\right| < (\leq) \left|q(\beta;\sigma) - \frac{1}{2}\right|, \tag{31}$$

then

$$\frac{\Pr(\text{piv}|\alpha;\sigma,n)}{\Pr(\text{piv}|\beta;\sigma,n)} > (\geq) 1. \tag{32}$$

Proof. The function $q(1-q)$ has an inverse u-shape on $[0,1]$ and is symmetric around its peak at $q = \frac{1}{2}$, as is illustrated in Figure 3. So, $\left|q - \frac{1}{2}\right| < (\leq) \left|q' - \frac{1}{2}\right|$ implies that $q(1-q) > (\geq) q'(1-q')$. Thus, it follows from (12) that (31) implies (32). ■

I show that Lemma 4 extends in an extreme form as the electorate grows large ($n \to \infty$): the event that the election is tied is infinitely more likely in the state in which the election is closer to being tied in expectation. In fact, the likelihood ratio of the pivotal event diverges exponentially fast.

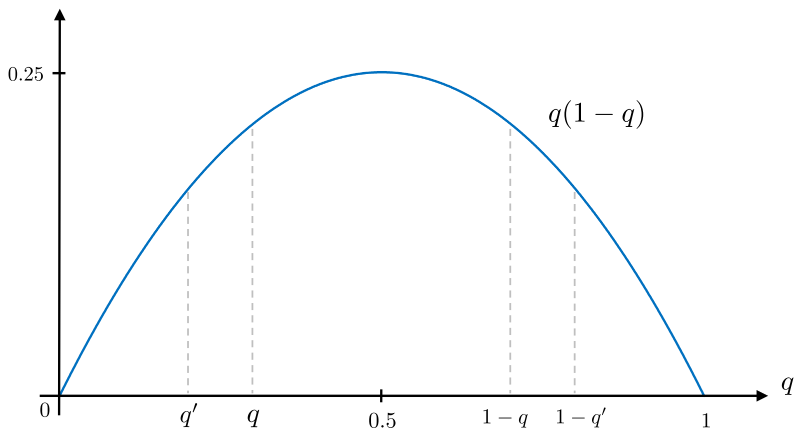

Figure 3: The function $q(1-q)$ for $q \in [0,1]$. If $|q - \frac{1}{2}| < |q' - \frac{1}{2}|$, then $q(1-q) > q'(1-q')$.

Lemma 5 *Consider any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$. If,*

$$\lim_{n\to\infty} \left| q(\alpha;\sigma_n) - \frac{1}{2} \right| < (>) \lim_{n\to\infty} \left| q(\beta;\sigma_n) - \frac{1}{2} \right|, \tag{33}$$

then, for any $\kappa \geq 0$,

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} n^{-\kappa} = \infty(0). \tag{34}$$

Proof. Let

$$k_n = \frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}.$$

From (12), the left-hand side of (34) is $\frac{(k_n)^n}{n^\kappa}$. The function $q(1-q)$ has an inverse u-shape on $[0,1]$ and is symmetric around its peak at $q = \frac{1}{2}$, as is illustrated in Figure 3. So, (33) implies that $\lim_{n\to\infty} k_n > 1$. So, $\lim_{n\to\infty} (k_n)^n = \infty$. Moreover, $(k_n)^n$ diverges exponentially fast and, hence, dominates the denominator n^κ, which is polynomial. ■

Last, I note that the posterior conditional on the pivotal event pins down the limit of the vote share in both states, as $n \to \infty$. To see why, recall that for any strategy sequence, the likelihood of the pivotal event converges to zero. Therefore, the first-order condition (19) implies that the precision of all types converges to zero uniformly. This implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv}, s) - \Pr(\alpha|\text{piv}; \sigma_n^*) = 0$, given the equilibrium precision $x(t)$ of any type. Hence, given (6)-(8) the limit vote share of A is given by the share of voters preferring A given the belief $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n)$, meaning that $\lim_{n\to\infty} q(\omega; \sigma_n^*) = \Phi(\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*))$. Using that Φ is continuous, I conclude:

Lemma 6 *Take any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$. For all n, let $\hat{\sigma}_n$ be any best response to σ_n. For any $\omega \in \{\alpha, \beta\}$,*

$$\lim_{n\to\infty} q(\omega; \hat{\sigma}_n) = \lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n)). \tag{35}$$

3 The Power of Voter Groups

In many elections we observe that many citizens are uninformed, while other citizens are much better informed. Generally, voters have information of heterogenous quality about the consequence of the election outcomes. This heterogeneity might have a systematic effect on election outcomes. In this section, I discuss the electoral power of the uninformed (Section 3.1) and the different types of informed citizens respectively.

For this, I consider a sequence of elections along which the electorate's size $2n + 1$ grows. For each n and a strategy σ_n, I calculate the probability that a policy $z \in \{A, B\}$ wins the support of the majority of the voters in state ω, denoted $\Pr(z|\omega; \sigma_n, n)$. I will be interested in the limits of $\Pr(z|\omega; \sigma_n^*, n)$ for equilibrium sequences $(\sigma_n^*)_{n\in\mathbb{N}}$. I am particularly interested in equilibrium sequences where citizens vote in an informed manner such that the election outcomes differ across the states,

$$\lim_{n\to\infty} \Pr(A|\alpha; \sigma_n, n) \neq \lim_{n\to\infty} \Pr(A|\beta; \sigma_n, n), \tag{36}$$

which I call *informative*.

When voters have conflicting interests, competitive information acquisition shapes the election. In Section 3.2, I analyze this competition between voters and show that there is a simple index rule that determines which voter group, aligned or contrarians, acquires more information and, second, in which state policy A is more likely to be elected. What matters is which voter group has the larger *power index*, which is defined as follows. For any $p \in [0, 1]$, any $g \in \{\ell, s\}$, let

$$W(g, p) = \Pr(g) f(p|t \in g) \mathrm{E}(k(t)^{\frac{2}{d-1}}|t \in g, y(t) = p), \tag{37}$$

where f is the density of the (conditional) distribution of the threshold of doubt. For any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and any $g \in \{s, \ell\}$,

$$W(g, \hat{p}) \tag{38}$$

is the *power index of group g*, where $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n, n)$. The power index is proportional to the likelihood of the critical types and the mean of the weighted intensities of the critical types as $n \to \infty$.

3.1 The Uninformed

This section provides a conceptual insight about information cost: not being informed in an election is similar to not voting. It turns out that in *any* informative equilibrium sequence, a random uninformed citizen votes almost $50 - 50$. The effect on the election outcome is therefore similar to abstention. In fact, I will show in a later part that the election outcome is random when only considering the votes of the uninformed.

What matters is that for some voters to be willing to acquire relatively much information, the election must be sufficiently close to being tied such that the likelihood of being the pivotal voter is relatively large from the perspective of a single voter. The following lemma verifies this intuition and shows that in any informative equilibrium sequence the *induced prior* $\Pr(\alpha|\text{piv}; \sigma_n^*)$ converges to a belief p^* given which a randomly drawn type would vote exactly $50-50$, i.e. $\Phi(p^*) = \frac{1}{2}$. Equivalently, the expected vote shares converge to $\frac{1}{2}$, given Lemma 6. Whenever, the expected vote share is not converging to $\frac{1}{2}$ as $n \to \infty$, the pivotal likelihood is exponentially small, see (13), and I show that the incentives to acquire information are then not sufficient for the equilibrium to be informative.[14]

Lemma 7 *For any informative equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$, there is p^* satisfying $\Phi(p^*) = \frac{1}{2}$ such that*

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) = p^*. \tag{39}$$

In the following, I illustrate the result with the running example. The proof is provided in the Appendix.

Uniform types Example. In the symmetric equilibrium of the running example, $\Pr(\alpha|\text{piv};\sigma_n^*) = \frac{1}{2}$ for all n, and $\Phi(\frac{1}{2}) = \frac{1}{2}$. This means, that without further information about the state, i.e. $c = \infty$, the election would be exactly tied in expectation and outcomes would be completely random. However, some voters acquire information, leading to informative election outcomes. The vote of the uninformed is, conversely, split exactly $50-50$ in expectation, those critical of policy A, with threshold of doubt $y > \frac{1}{2} + x^{**}$, vote B, and those critical of policy B, with threshold of doubt $y < \frac{1}{2} - x^{**}$, vote A. If we would just consider the uninformed votes, the outcome of the election would be random. Surprisingly, this observation generalizes to *all* settings considered. Let

$$\Pr(A|\omega;\sigma_n^*;\tilde{\pi}_n) \tag{40}$$

be the likelihood of outcome A in the game $\tilde{\pi}_n$ of $2n+1$ voters where a random citizen votes A with the same likelihood as an uninformed voter, given σ_n^*.

[14] The term induced prior was introduced by Bhattacharya (2013).

Observation 1 *Take any preference distribution H with $W(\ell,p) \neq W(s,p)$ for any $p \in [0,1]$ with $\Phi(p) = \frac{1}{2}$. For any informative equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$, in each state $\omega \in \{\alpha,\beta\}$, the outcome implied by the vote share of the uninformed voters is random, i.e.* $\Pr(A|\omega;\sigma_n^*;\tilde{\pi}_n) \in (0,1)$.

The proof is in Section 5, where I finish the characterization of all informative equilibrium sequences . The result is stated here already, to provide guidance and intuition for the upcoming analysis.

3.2 The Partially Informed: Competition in Information Acquisition

Suppose that there is an informative equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$. Given Observation 1, the election is indeterminate when considering only the votes of the uninformed. Intuitively, the election will therefore be decided by the voters that get partially informed. The partially informed citizens vote for their prefered policy with a likelihood of $\frac{1}{2}+x(t)$, and therefore, shift the vote share towards their prefered policy in each state. On the aggregate, the difference in the expected vote share for policy A across the states is

$$q(\alpha;\sigma_n^*) - q(\beta;\sigma_n^*) \;=\; 2\Big[\int_{t\in\ell} x(t)dH(t) - \int_{t\in s} x(t)dH(t)\Big]. \qquad (41)$$

where $\int_{t\in g} x(t)dH(t) = \Pr_H(g)\mathrm{E}_H(x(t)|t\in g)$ is the average precision acquired by a type of the group $g \in \{\ell, s\}$, weighted by the likelihood of a random type belonging to the group. We see that the order of the likelihood-weighted average precision determines the order of the vote shares in the states and therefore, in which state A is more likely to be elected.
The next result shows that there is a simple rule describing which group has the higher likelihood-weighted average precision: it is the group with the higher power index.

Lemma 8 *Take any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$. For all n, let $\hat{\sigma}_n$ be any best response to σ_n. If $\hat{p} = \lim_{n\to\infty}\Pr(\alpha|\mathrm{piv};\sigma_n,n) \in (0,1)$ the sequence of* *best responses satisfies*

$$\lim_{n\to\infty} \frac{\int_{t\in\ell} x(t)dH(t)}{\int_{t\in s} x(t)dH(t)} = \frac{W(\ell,\hat{p})}{W(s,\hat{p})}. \tag{42}$$

In the following, I illustrate the result with the running example and, then, provide a sketch of the proof in the main text. The general proof is in Section B of the Appendix.

Example with Uniform Types (version 2). I return to the example with uniform types of Section 2.1.6, but modify it slightly: each voter is aligned with probability $\lambda > \frac{1}{2}$ and contrarian with probability $(1-\lambda)$, the cost are polynomial $c(x) = x^d$ and the total intensity is the same for all aligned and contrarian voters respectively, but might depend on the voter group, that is $k(t) = k_\ell$ for all $t \in \ell$ and $k(t) = k_s$ for all $t \in s$. The prior is uniform and the conditional distribution of the threshold of doubt is the same for aligned and contrarian types respectively and uniform.
There is a symmetric equilibrium that is fully described by a pair $(x_\ell^*, x_s^*) \in (0,1)^2$ of signal precisions, similar to the symmetric equilibrium described in Section 2.1.6. The proof is verbatim with the required changes in notation and therefore omitted. In this equilibrium, $x_\ell^* > 0$ is the precision acquired by the aligned types that get partially informed, and $x_s^* > 0$ the precision acquired by the contrarian types that get partially informed. Both x_ℓ^* and x_s^* are pinned down by the first-order condition (see (19) and (26)). The types that get partially informed are those for which $y(t) \in [\frac{1}{2} - x_g^*(1 - \frac{c(x_g^*)}{x_g^* c'(x_g^*)}), \frac{1}{2} + x_g^*(1 - \frac{c(x_g^*)}{x_g^* c'(x_g^*)})]$. I conclude,

$$\begin{aligned}
\lim_{n\to\infty} \frac{\int_{t\in\ell} x(t)dH(t)}{\int_{t\in s} x(t)dH(t)} &= \lim_{n\to\infty} \frac{\Pr(\ell)}{\Pr(s)} \frac{(x_\ell^*)^2(1 - \frac{c(x_\ell^*)}{x_\ell^* c'(x_\ell^*)})}{(x_s)^2(1 - \frac{c(x^*)}{x_g^* c'(x^*)})} \\
&= \frac{\Pr(\ell)}{\Pr(s)} \frac{(x_\ell^*)^2}{(x_s)^2},
\end{aligned} \tag{43}$$

where, for the last line, I used $\frac{c(x)}{xc'(x)} = \frac{1}{d}$, given that $c(x) = x^d$. The first-order condition (26) pins down the ratio of the precision of aligned and contrarians, $\frac{x^*_\ell}{x^*_s} = (\frac{k_\ell}{k_s})^{\frac{1}{d-1}}$. Finally, the definition of the power indices $W(g)$, i.e. (38), together with (43) implies (42).

Sketch of Proof: General Case. The analysis for the general case follows along similar lines. However, one has to show that the order of several types of asymmetric behaviour is sufficiently small such that the results carry over asymptotically as $n \to \infty$. First, even when fixing the total intensity, different types choose different precisions in general. I show that this heterogeneity is sufficiently small and the information precision of all types that get partially informed is well approximated by the precision of the critical types, which satisfy $y = \Pr(\alpha|\text{piv}; \sigma_n, n)$. Using a Taylor approximation, for all k,

$$x(y,k) > 0 \Rightarrow \frac{x(y,k)}{x(\Pr(\alpha|\text{piv}; \sigma_n, n), k)} \approx 1. \tag{44}$$

The basic intuition is that only types arbitrarily close to the critical types acquire information when n is growing large. The second insight comes from a close inspection of the first-order condition (19). Changes in the marginal benefit on the right hand side due to changes in the threshold of doubt y are of an order of the marginal cost $c'(x(y,k))$. These changes relate to infinitely smaller changes in the equilibrium precision on the left hand side since the second derivative $c''(x(y,k))$ of the cost function is much larger than the first derivative $c'(x(y,k))$ when the precision is small.[15]
Second, the information acquisition region might be asymmetric around the critical types with $y = \Pr(\alpha|\text{piv}; \sigma_n, n)$. I show that this asymmetry is of a sufficiently small order. Similar to the uniform types example,

$$\frac{x(\hat{y},k)(1 - \frac{c(x(\hat{y},k))}{c'(x(\hat{y},k))x(\hat{y},k)})}{\Big[\phi_g(k) - \Pr(\alpha|\text{piv}; \sigma_n, n)\Big]} \approx \frac{x(\hat{y},k)(1 - \frac{c(x(\hat{y},k))}{c'(x(\hat{y},k))x(\hat{y},k)})}{\Big[\Pr(\alpha|\text{piv}; \sigma_n, n) - \psi_g(k)\Big]} \approx K, \tag{45}$$

[15]Consider the example of polynomial cost $c(x) = x^d$. Then $\lim_{x\to 0} \frac{c''(x)}{c'(x)} = \lim_{x\to\infty} \frac{d-1}{x} = \infty$.

where K is a constant that only depends on $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n,n) = \hat{p}$. To show this, I consider the linear (Taylor) approximations at $\hat{y}_n = \Pr(\alpha|\text{piv};\sigma_n^*)$ of the left hand side, $\chi(y)$, of the equations (29) and (30) that define the boundaries $\psi_g(k)$ and $\phi_g(k)$, . Since the left hand side takes the value $\frac{1}{2}$ at $\hat{y}_n$, this yields $\chi'(\hat{y}_n)\Big[\psi_g(k)-\hat{y}_n)\Big] \approx x^{**}(\psi_g(k))$ and $\chi'(\hat{y}_n)\Big[\hat{y}_n-\phi_g(k)\Big] \approx x^{**}(\psi_g(k))$.[16] The intuition here is again that only types arbitrarily close to the critical types acquire information when n is large such that the boundaries are arbitrarily close to $\hat{y}$ and the linear approximation is asymptotically precise. Finally, (45) follows from (44) and the Taylor approximations of (29) and (30).

Third, when the distribution of the thresholds is not uniform, the linear (Taylor) approximation of the distribution around the critical types is asymptotically precise to describe the mass of types that acquire information,

$$\frac{f(\hat{y},k|k'=k,g)\Big[\phi^g(k)-\psi^g(k)\Big]}{F(\psi^g(k)|k'=k,g)-F(\phi^g(k)|k'=k,g)} \approx 1. \tag{46}$$

where $F(-|k'=k,g)$ is the conditional distribution of the thresholds and $f(-|k'=k,g)$ its density.

Finally, given (44), (46), and (45),

$$\begin{aligned} &\int_{t\in g} x(t)dH(t) \\ \approx\ & \int_k f(\hat{y}_n,k|k'=k,g)2Kx(\hat{y}_n,k)^2(1-\frac{c(x(\hat{y}_n,k))}{c'(x(\hat{y}_n,k))x(\hat{y}_n,k)})dJ(k|g) \\ =\ & \Pr(g)f(\hat{y}_n|g)\mathrm{E}_k(x(y,k)^2(1-\frac{c(x(y,k))}{c'(x(y,k))x(y,k)})|y=\hat{y}_n,g). \end{aligned} \tag{47}$$

where $J(\cdot|g)$ is the conditional distribution of the total intensity of the types

[16] The approximation errors are of an order $\Big[\phi_g(k)-\hat{y}_n\Big]^2$ and $\Big[\psi_g(k)-\hat{y}_n\Big]^2$ respectively. Since only types arbitrarily close to the critical types acquire information as $n\to\infty$, it holds$\phi_g(k) \overset{n\to\infty}{\to} \psi_g(k)$ for all $k>0$. Hence, the error terms are negligible.

of group g. Thus,

$$\lim_{n\to\infty} \frac{\int_{t\in g} x(t)dH(t)}{\Pr(g)f(\hat{p}|g)\mathrm{E}_k\left[x(y,k)^2|y=\hat{p},g\right]^{\frac{2K(d-1)}{d}}} = 1 \tag{48}$$

where I used the notation $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}, \sigma_n^*) = p^*$ and, second, that $(1 - \frac{c(x(y,k))}{c'(x(y,k))x(y,k)}) \to \frac{d-1}{d}$ as $n \to \infty$, which follows from Lemma 14 in the Appendix and since $x(y,k) \to 0$ for all (y,k). The first-order condition (19) together with $\lim_{x\to 0} \frac{c'(x)}{x^{d-1}} \in \mathbb{R}$ implies $x(\hat{p},k) \approx k^{\frac{1}{d-1}} x(\hat{p},1)$ for all $k > 0$. Thus, it follows from (48) that the ratio of the power indices $W(g)$ defined by (38) is asymptotically equivalent to the ratio of the likelihood-weighted average precision $\int_{t\in g} x(t)dH(t)$, that is (42) holds.

Given Lemma 8, whenever the contrarians have a larger power index, policy A is more likely to be elected in β than in α. Clearly, this implies that the median-voter preferred outcome is less likely to be elected in one of the states, since the median voter prefers A only in α.

Corollary 1 *Let $W(\ell,p) < W(s,p)$ for any $p \in [0,1]$ with $\Phi(p) = \frac{1}{2}$. For any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$, there is a state $\omega \in \{\alpha, \beta\}$ where the median-voter preferred outcome is less likely to be elected as $n \to \infty$, i.e.*

$$\lim_{n\to\infty} \Pr(A|\alpha; \sigma_n^*, n) \leq \frac{1}{2} \quad \textit{or} \quad \lim_{n\to\infty} \Pr(B|\beta; \sigma_n^*, n) \leq \frac{1}{2}. \tag{49}$$

Proof. In the Appendix, Lemma 15 (a corollary of the proof of Lemma 7) shows that any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ satisfies $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n) \in \{\Pr(\alpha), \Phi^{-1}(\frac{1}{2})\}$. Since $\Phi(0) \neq \frac{1}{2}$, $\Phi(1) \neq \frac{1}{2}$, and $\Pr(\alpha) \in (0,1)$ by assumption, the corollary follows from Lemma 8. ■

4 Welfare-Maximizing Outcomes

In this section, I study the welfare properties of the election. I describe the welfare effect of the disparate information acquisition of voters. For this note that policy A is welfare-maximizing in state ω if $\mathrm{E}_H(t_\omega) \geq 0$.

The key insight is that people that care much about the election, will be better informed. I show that this shifts outcomes into a direction that improves social welfare, from the median towards the mean of the voters preferences. Somewhat surprisingly, a simple majority election can choose outcomes that are preferred by only a minority of the voters, but only if this is welfare-improving.

The main result of this section is that, for a large class of settings, elections have equilibria that yield outcomes maximizing social welfare in each state, unless information cost are too extreme, that is, too low or too high. I consider the following settings: recall from Section 2.1.5 that we can understand the conditional distribution of the aligned types or the contrarian types respectively simply as a distribution of pairs of total intensities k and threshold of doubts y. I consider preference distributions for which the conditional distribution of the threshold of doubt, $F(-|g)$ is independent of the voter group, i.e. for all $g \in \{\ell, s\}$,

$$F(-|g) = F. \tag{50}$$

and F is independent from the conditional distribution $J(-|g)$ of the total intensities of types $t \in g$, that is, for all $g \in \{s, \ell\}$,

$$J(-|g) \perp\!\!\!\perp F. \tag{51}$$

Further, the welfare at stake for the A-partisans is, in expectation, of the same magnitude as the welfare at stake for the B-partisans, i.e. for all $\omega \in \{\alpha, \beta\}$,

$$\mathrm{E}_H(|t_\omega||\omega, \{t : t_\alpha > 0, t_\beta > 0\}) = \mathrm{E}_H(|t_\omega||\omega, \{t : t_\alpha < 0, t_\beta < 0\}). \tag{52}$$

Clearly, information cost cannot screen the intensities of the partisans since all the partisans stay uninformed and simply vote their preferred policy. Condition (52) describes the settings where screening of the intensities of the partisans is obsolete.

Theorem 1 *For any preference distribution H that satisfies (50)- (52),* $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$*, and for which $W(\ell, p) \neq W(s, p)$ for any $p \in [0,1]$ with $\Phi(p) = \frac{1}{2}$, there is an open set of cost functions and an equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ given any cost function of this set for which the welfare-maximizing outcome is elected with probability converging to* 1 *as $n \to \infty$.*

Note that I consider cost functions 'close to each other' when they have a similar elasticity at $x = 0$.[17]

Welfare-Maximizing Outcomes. Given the independence assumptions (50) - (52), for each state ω, the outcome preferred by the voter group, aligned or contrarians, with the larger likelihood weighted mean of the total intensities is the welfare-maximizing outcome; that is, the welfare-maximizing outcome is A in α and B in β if

$$\Pr(\ell)\mathrm{E}(k(t)|t \in \ell) > \Pr(s)\mathrm{E}(k(t)|t \in s) \tag{53}$$

and otherwise B in α and A in β. To see why, consider, for example, state α. By definition, A is the welfare-maximizing outcome if $E_H(t_\alpha) > 0$. Given (52), this is equivalent to $\Pr(\ell)\mathrm{E}(k(t)(1-y(t))|t \in \ell) - \Pr(s)\mathrm{E}(k(t)(1-y(t))|t \in s) > 0$, given (21) and (22). Given (50) and (51), this is equivalent to (53).

The relevant statistic is the ratio of the power indices of the aligned and contrarians, that is $\frac{W(\ell,\hat{p})}{W(s,\hat{p})}$.To prove Theorem 1, I show a lemma: when information of low precision $x \approx 0$ is sufficiently cheap, which will be captured by a condition on the elasticity at zero, $\lim_{x\to 0} \frac{c'(x)x}{x}$, there is an equilibrium sequence where the outcome preferred by the group with the larger power index is elected with probability converging to 1 as $n \to \infty$.

Lemma 9 *Let* $\lim_{x\to 0} \frac{c'(x)x}{x} > 3$*. Take any preference distribution H for which $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$ and $W(\ell, p) \neq W(s, p)$ for any $p \in [0,1]$ with $\Phi(p) = \frac{1}{2}$. There*

[17] I equip the cost functions with the topology given by the Euclidean distance of their elasticities at zero, $\lim_{x\to\infty} \frac{c'(x)x}{c(x)} = d$.

is an informative equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ *satisfying*

$$\lim_{n\to\infty} \Pr(A|\alpha;\sigma_n^*,n) = \lim_{n\to\infty} \Pr(B|\beta;\sigma_n^*,n) = \begin{cases} 0 & \text{if } W(\ell,\hat{p}) < W(s,\hat{p}), \\ 1 & \text{if } W(\ell,\hat{p}) > W(s,\hat{p}). \end{cases} \tag{54}$$

where $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*,n)$.

First, Lemma 14 in the Appendix shows that $\lim_{x\to 0} \frac{c'(x)x}{x} = d$. Now, lemma 9 implies Theorem 1 as follows. Given the independence assumptions (50) and (51),

$$\begin{aligned} W(\ell,\cdot) &> W(s,\cdot) \\ \Leftrightarrow \Pr(\ell)\mathrm{E}(k(t)^{\frac{2}{d-1}}|t\in\ell) &> \Pr(s)\mathrm{E}(k(t)^{\frac{2}{d-1}}|t\in s). \end{aligned} \tag{55}$$

Comparing (53) and (55), clearly, when $d > 3$ is sufficiently close to $d' = 3$, given the equilibrium sequence of Lemma 9 the welfare-maximizing outcome is elected in each state. Thus, the lemma implies that for each preference distribution H with $W(\ell) \neq W(s)$, there is an open set of elasticities $d > 3$ and equilibrium sequences, given d, such that the welfare-maximizing outcome is elected in each state.

4.1 Proof: Welfare-Maximizing Outcomes

Section 4.1.1 prepares the proof of Lemma 9 by analyzing which cost functions allow for sufficient information acquisition by the voters, as needed for election outcomes to be *informative*, $\lim_{n\to\infty} \Pr(A|\alpha;\sigma_n,n) \neq \lim_{n\to\infty} \Pr(B|\alpha;\sigma_n,n)$ and determinate in both states, $\lim_{n\to\infty} \Pr(B|\omega;\sigma_n,n) \in \{0,1\}$

4.1.1 Preliminaries: Value and Cost of Information

Value of Information. The value of information is proportional to the likelihood of the pivotal event, given (18). How large is this likelihood? Since the vote share in a state is the empirical mean of $2n+1$ i.i.d. Bernoulli variables which take the value 1 with probability $q(\omega;\sigma_n)$, an application of the local central limit[18] tells how likely events close to the expected vote share are. In particular, the next result shows that when the expected vote share in a state is sufficiently close to the majority threshold $\frac{1}{2}$, the pivotal likelihood is proportional to the standard deviation of the vote share, which is of order $n^{-\frac{1}{2}}$.[19] For any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and any n, let $s(\omega;\sigma_n)$ be the standard deviation of the vote share. Let

$$\delta_\omega = \lim_{n\to\infty} \frac{1}{s(\omega;\sigma_n)}\Big[q(\omega;\sigma_n) - \frac{1}{2}\Big] \tag{56}$$

be the normalized distance of the expected vote share to the majority threshold as $n \to \infty$.

Lemma 10 *For any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and for all $\omega \in \{\alpha,\beta\}$,*

$$\lim_{n\to\infty} \Pr(\text{piv}|\omega;\sigma_n)s(\omega;\sigma_n)^{-1} \;=\; \phi(\delta_\omega), \tag{57}$$

where ϕ the probability density function of the standard normal distribution.

The proof is omitted since the result directly follows from the local central limit theorem.

Figure 4 provides intuition for Lemma 12.

Cost of Information. It depends on the cost of information, how the value of information translates into information acquisition and election outcomes. In what follows, I describe a condition on the cost function that allows for informative and determinate outcomes under the best response to any strategies with vote shares sufficiently close to $\frac{1}{2}$. Conversely, I give a condition when this is not the case; then, intutively, determinate and informative

[18] See Gnedenko (1948), and McDonald (1980) for the local limit theorem for triangular arrays of integer-valued variables.

[19] The number of A-votes is distributed according to a binomial distribution with parameters n and $q_{\omega,n}$. Hence, its variance is $nq(\omega;\sigma_n)(1-q(\omega;\sigma_n))$, and the standard deviation $(nq(\omega;\sigma_n)(1-q(\omega;\sigma_n))^{\frac{1}{2}}$. Consequently, the standard deviation of the vote share is distributed according to $\frac{1}{n}\mathcal{B}(n,q(\omega;\sigma_n))$, so its standard deviation is $(\frac{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}{n})^{\frac{1}{2}}$.

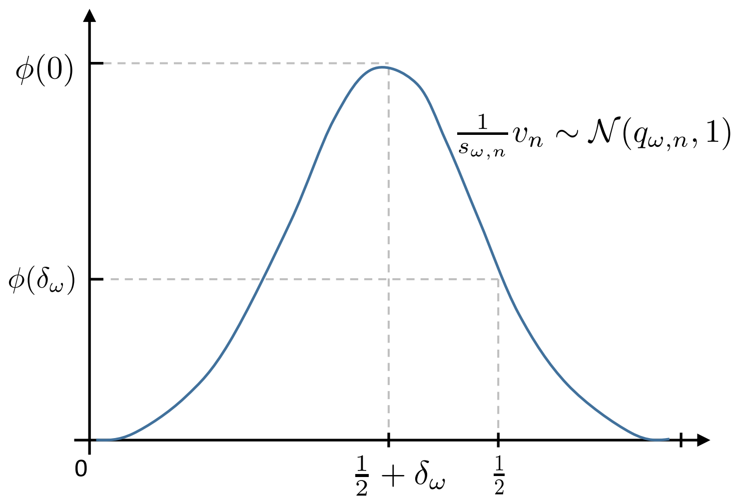

Figure 4: Conditional on ω, the realized vote share v_n follows a binomial distribution with parameters $q(\omega;\sigma_n^*)$ and $2n+1$. It follows from the Central Limit Theorem for triangular arrays that the distribution of the normalized vote share converges a normal distribution.

outcomes are not possible under the best response to *any* strategy since the value of information will just be lower if the election is less close to being tied. It turns out that the key statistic is the elasticity of the cost function for low levels of precision $x \approx 0$. To see why, in a first step, I show the following result.

Lemma 11 *Take a sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ with $\delta_\alpha \in \mathbb{R}$ and suppose that $W(\ell,\hat{p}) \neq W(s,\hat{p})$ for $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n,n)$. Let $\hat{\sigma_n}$ be a best reponse to σ_n.*

1. *If $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$, then*

$$\lim_{n\to\infty}\frac{q(\alpha;\hat{\sigma_n}) - q(\beta;\hat{\sigma_n})}{s(\alpha;\sigma_n)} = \infty. \tag{58}$$

2. *If $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$, then*

$$\lim_{n\to\infty}\frac{q(\alpha;\hat{\sigma_n}) - q(\beta;\hat{\sigma_n})}{s(\alpha;\sigma_n)} = 0. \tag{59}$$

The proof is provided in Section C of the Appendix.

How do the information acquisition and expected vote shares translate into election outcomes? Take a strategy $(\sigma_n)_{n\in\mathbb{N}}$ with $\delta_\alpha \in \mathbb{R}$ and let $d > 3$. Given Lemma 11, it might be that the expected vote shares are arbitrarily many standard deviations away from the majority threshold under the best response, i.e. $\delta_\omega((\hat{\sigma}_n)_{n\in\mathbb{N}}) \in \{\infty, -\infty\}$ for all $\omega \in \{\alpha, \beta\}$. The following result characterizes the distribution of the election outcomes in state ω as $n \to \infty$ for any given sequence of strategies as a function of $\delta_\omega \in \mathbb{R} \cup \{\infty, -\infty\}$. The result implies that outcomes are determinate in ω as $n \to \infty$ when $\delta_\omega \in \{\infty, -\infty\}$.

Lemma 12 *Consider any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and any state $\omega \in \{\alpha, \beta\}$. The probability that A gets elected in ω converges to*

$$\lim_{n\to\infty} \Pr(A|\omega_i; \sigma_n) = \Phi(\delta_\omega),$$

where $\Phi(\cdot)$ is the cumulative distribution of the standard normal distribution.

The proof relies on an application of the central limit theorem and is provided in the Appendix. Figure 4 illustrates Lemma 12.

In the next section, I prove Lemma 9, using the previous lemmas 11 and 12. Thereby, I show the existence of informative equilibrium sequences with both determinate and different outcomes in the states, given $\lim_{n\to\infty} \frac{c'(x)x}{c(x)} > 3$.

4.1.2 Fixed Point Argument

First, I provide a useful compact representation of equilibrium.

Equilibrium Vote Shares. It follows from the analysis of the best response in Section 2 that, for n large enough, an equilibrium is a (non-degenerate) strategy $\sigma = (x, \mu)$ that satisfies (6)- (9), with $\sigma' = \sigma$ and (19) for all types t with $x(t) > 0$.

I claim that equilibrium can be alternatively characterized in terms of the vector of the expected vote shares of outcome A in state α and β, i.e.

$$\mathbf{q}(\sigma) = (q(\alpha;\sigma), q(\beta;\sigma)). \tag{60}$$

Note that for any σ and any $\omega \in \{\alpha, \beta\}$, the vote share $q(\omega;\sigma)$ pins down the likelihood of the pivotal event conditional on ω, given (12). Given (6)-(9) and (19), the vector of the pivotal likelihoods is a sufficient statistic for the best response, and therefore $\mathbf{q}(\sigma)$ as well. Given some vector of expected vote shares $\mathbf{q} = (q(\alpha), q(\beta)) \in (0,1)$, let $\sigma^{\mathbf{q}}$ be the best response to $\mathbf{q}$. So, σ^* is an equilibrium if and only if $\sigma^* = \sigma^{\mathbf{q}(\sigma^*)}$. Conversely, an equilibrium can be described by a vector of vote shares $\mathbf{q}^*$ that is a fixed point of $\mathbf{q}(\sigma^{\cdot})$, that is[20]

$$\mathbf{q}^* = \mathbf{q}\left(\sigma^{\mathbf{q}^*}\right). \tag{61}$$

In the following, I use the notation $\Pr(\alpha|\text{piv};\mathbf{q})$ to denote the posterior consistent with (12) and the vote shares $\mathbf{q}$, and also similar notation like $s(\omega;\mathbf{q})$, analogous to the previous notation.

Proof of Lemma 9. First, I simplify the problem of finding an informative equilibrium sequence further to a problem of finding fixed points of one-dimensional functions.

For this, I define a constrained best response $\hat{q}(\omega, \sigma^{\mathbf{q}})$ as follows. Let $\hat{p} \in [0,1]$ be the minimal belief for which $\Phi(\hat{p}) = \frac{1}{2}$ and $\Phi'(p^*) > 0$.[21] Consider the case when $\Pr(\alpha) < \hat{p}$ and $W(\ell, \hat{p}) < W(s, \hat{p})$. Take $\delta > 0$ small enough such that Φ is strictly increasing on $[\hat{p}-\delta, \hat{p}+\delta]$ and $\Pr(\alpha) < \hat{p} - \delta$. Now, if $\Pr(\alpha|\text{piv};\mathbf{q}) \in [\hat{p}-\delta, \hat{p}+\delta]$, then let $\hat{q}(\omega, \sigma^{\mathbf{q}}) = q(\omega, \sigma^{\mathbf{q}})$. If $\Pr(\alpha|\text{piv};\mathbf{q}) < \hat{p}-\delta$, then let $\hat{q}(\omega, \sigma^{\mathbf{q}})$ be the vote share of the best response to any strategy σ' with

[20] The ability to write an equilibrium as a finite-dimensional fixed point via (61) is a significant advantage. A similar reduction to finite dimensional equilibrium beliefs has been useful in other settings as well; see Bhattacharya (2013), Ahn & Oliveros (2012) and Heese & Lauermann (2017).

[21] Recall that $\Phi(0) < \frac{1}{2}$ and $\Phi(1) > \frac{1}{2}$ by assumption and that Φ is continuously differentiable and not constant on any open interval. This implies the existence of $\hat{p}$.

induced prior $\Pr(\alpha|\text{piv};\sigma',n) = \hat{p} - \delta$ and for which the likelihood of the pivotal event is $\Pr(\text{piv}|\mathbf{q},n)$. Conversely, if $\Pr(\alpha|\text{piv};\mathbf{q}) > \hat{p}+\delta$, then let $\hat{q}(\omega,\sigma^{\mathbf{q}})$ be the vote share of the best response to any strategy σ' with induced prior $\Pr(\alpha|\text{piv};\sigma',n) = \Phi(\hat{p}+\delta)$ and for which the likelihood of the pivotal event is $\Pr(\text{piv}|\mathbf{q},n)$. The constrained best response is a 'truncation' of $q(\omega;\sigma^{\mathbf{q}})$ and therefore continuous in $\mathbf{q}$.

Step 1 *For any $\epsilon > 0$ small enough, any $\frac{1}{2} - \frac{\epsilon}{2} \leq q(\alpha) \leq \frac{1}{2}$, and any n large enough, there is $q_n^*(\beta) \geq \frac{1}{2}$ such that*

$$\begin{aligned} q(\alpha) &= \hat{q}(\alpha;\sigma^{(q(\alpha),q_n^*(\beta))}) && (62)\\ &= q(\alpha;\sigma^{(q(\alpha),q_n^*(\beta))}). && (63) \end{aligned}$$

and $q^*(\beta)$ is continuous in $q(\alpha)$.

Take any $q(\alpha) \in [\frac{1}{2}, \frac{1}{2} - \frac{\epsilon}{2}]$. Let $\mathbf{q} = (q(\alpha), q(\beta))$ in the following.

Step 1.1 *If $q(\beta) = \frac{1}{2} + \epsilon$, then, for ϵ small enough and n large enough,*

$$\hat{q}(\alpha;\sigma^{\mathbf{q}}) > q(\alpha). \tag{64}$$

The election is more close to being tied in α, and, by Lemma 5, voters become convinced that the state is α, that is $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\mathbf{q},n) = 1$. It follows from the definition of $\hat{q}(\omega;\sigma^{\mathbf{q}})$ and Lemma 6 that $\lim_{n\to\infty} \hat{q}(\alpha;\sigma^{\mathbf{q}}) = \Phi(\hat{p}+\delta)$. Finally, (64) follows when ϵ is small enough, since $\Phi(\hat{p}+\delta) > \frac{1}{2}$.

Step 1.2 *If $q(\beta) = \frac{1}{2}$, then for ϵ small enough and any n,*

$$\hat{q}(\alpha;\sigma^{\mathbf{q}}) < q(\alpha). \tag{65}$$

The election is more close to being tied in β, and, by Lemma 4, voters update towards β, that is $\Pr(\alpha|\text{piv};\mathbf{q},n) \leq \Pr(\alpha)$. Since $\Pr(\alpha) < \hat{p} - \delta$, it follows from the definition of $\hat{q}(\alpha;\sigma^{\mathbf{q}})$ and Lemma 6 that $\lim_{n\to\infty} \hat{q}(\alpha;\sigma^{\mathbf{q}}) = \Phi(\hat{p}-\delta)$. Finally, (65) follows for $\epsilon > 0$ small enough since $\Phi(\hat{p}-\delta) < \frac{1}{2}$.

Since $\hat{q}(\alpha;\sigma^{\mathbf{q}})$ is continuous in $q(\beta)$, it follows from Step 1.1, Step 1.2, and the intermediate value theorem that, for n large enough, there is $q^*(\beta)$ such that (62) holds. It follows from the implicit function theorem that $q^*(\beta)$ is continuous. Now, suppose that $\Pr(\alpha|\text{piv};\sigma^{(q(\alpha),q_n^*(\beta))},n) \notin [\hat{p}-\delta,\hat{p}+\delta]$. Then, Lemma 6 together with the definition of $\hat{q}(\alpha;\sigma^{\mathbf{q}})$ implies that $\lim_{n\to\infty}\hat{q}(\alpha;\sigma^{\mathbf{q}}) \in \{\Phi(\hat{p}-\delta),\Phi(\hat{p}+\delta)\}$. This contradicts with (62) since $\Phi(\hat{p}-\delta) < q(\alpha) < \Phi(\hat{p}+\delta)$ for $\epsilon > 0$ small enough. Hence, $\lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma^{(q(\alpha),q^*(\beta))},n) \in [\hat{p}-\delta,\hat{p}+\delta]$ and therefore, given the definition of the 'truncation' $\hat{q}$, (62) implies (63).

In what follows, I show that, for any n large enough, there is a vote share $q_n^*(\alpha)$ such that $\mathbf{q}_n^* = (q_n^*(\alpha), q^*(\beta))$ is a fixed point of $\mathbf{q}(\sigma^{\cdot})$, thereby constructing equilibria $\sigma^{\mathbf{q}_n^*}$ of the voting game. Given (62), it is sufficient to show the following.

Step 2 *For any n large enough, there is $q_n^*(\alpha)$ such that*

$$q(\beta) = q(\beta;\sigma^{(q(\alpha),q^*(\beta))}). \tag{66}$$

I consider sequences of vote shares $(\mathbf{q}_n)_{n\in\mathbb{N}}$ with $\mathbf{q}_n = (q(\alpha)_n, q^*(\beta)_n)$.

Step 2.1 *If $\lim_{n\to\infty}(\frac{1}{2}-q(\alpha)_n)s(\alpha;\mathbf{q}_n) \in \mathbb{R}^{\geq 0}$, then, for n large enough,*

$$q^*(\beta)_n < q(\beta;\sigma^{\mathbf{q}_n}). \tag{67}$$

Recall that $\lim_{x\to 0}\frac{c'(x)x}{c(x)} = d$ (see Lemma 16). Now, (63) together with the assumption $\lim_{n\to\infty}(\frac{1}{2}-q(\alpha)_n)s(\alpha;\mathbf{q}_n) \in \mathbb{R}^{\geq 0}$ implies

$$\delta(\alpha)((\sigma^{\mathbf{q}_n})_{n\in\mathbb{N}}) = \lim_{n\to\infty}(\frac{1}{2}-q(\alpha;\sigma^{\mathbf{q}_n}))s(\alpha;\mathbf{q}_n) \in \mathbb{R}^{\geq 0}. \tag{68}$$

Then, Lemma 6 implies $\lim_{n\to\infty}\Pr(\alpha|\text{piv};\mathbf{q}_n) \in \Phi^{-1}(\frac{1}{2})$. Given the definition of the 'truncated' best response $\hat{\sigma}$,

$$\lim_{n\to\infty}\Pr(\alpha|\text{piv};\mathbf{q}_n) = \hat{p}. \tag{69}$$

Since I consider the case where $W(\ell, \hat{p}) < W(s, \hat{p})$, it follows from Lemma 8 and (41) that $q(\alpha; \sigma^{\mathbf{q}_n}) < q(\beta; \sigma^{\mathbf{q}_n})$ for n large enough. Given (68), Lemma 11 implies that

$$\lim_{n\to\infty} \left[q(\beta; \sigma^{\mathbf{q}_n}) - \frac{1}{2}\right] |s(\alpha; \mathbf{q}_n) = \infty. \tag{70}$$

Given (69) and since $\hat{p} \in (0, 1)$, the inference from the pivotal event is bounded. Therefore, Lemma 10 implies that

$$(q^*(\beta) - \frac{1}{2})s(\beta; \mathbf{q}_n) \in \mathbb{R} \tag{71}$$

since otherwise $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\mathbf{q})}{\Pr(\alpha|\text{piv};\mathbf{q})} = \frac{\Phi(\delta_\alpha((\sigma^{\mathbf{q}_n})_{n\in\mathbb{N}}))}{\Phi(\delta_\beta((\sigma^{\mathbf{q}_n})_{n\in\mathbb{N}}))} = 0$. Note that $\lim_{n\to\infty} \frac{s(\alpha;\mathbf{q}_n)}{s(\beta;\mathbf{q}_n)} = 1$.[22] Therefore, (70) and (71) imply (67) for n large enough.

Step 2.2 *If $q(\alpha) = \frac{1}{2} - \frac{\epsilon}{2}$, then, for n large enough,*

$$q_n^*(\beta) > q(\beta; \sigma^{\mathbf{q}_n}). \tag{72}$$

Lemma 6 together with (62) implies that $\lim_{n\to\infty} q(\beta; \sigma^{\mathbf{q}}) = \frac{1}{2} - \epsilon$. Since $q^*(\beta) \geq \frac{1}{2}$, clearly, (72) holds for n large enough.
Finally, since $q(\beta; \sigma^{\mathbf{q}})$ is continuous in $q(\alpha)$, an application of the intermediate value theorem shows that, for all n large enough, there is $q^*(\alpha) < \frac{1}{2}$ for which (66) holds. This finishes the proof of Step (2.2). The corresponding strategies $\sigma_n^* = \sigma^{(q_n^*(\alpha), q^*(\beta))}$ form an equilibrium sequence .

It remains to show that the election chooses the outcome preferred by the contrarians in each state with probability converging to 1, given the equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$. Since $q(\alpha; \sigma_n^*) \leq \frac{1}{2} \leq q(\beta; \sigma_n^*)$, Lemma 10 implies that it remains to show that $|\delta_\omega| = \infty$ for all states ω.
Suppose that $\delta_\omega \in \mathbb{R}$ for some $\omega \in \{\alpha, \beta\}$. I claim that this implies that $\delta_\omega \in \mathbb{R}$ for all states. Otherwise, Lemma 12 implies that the inference from the pivotal

[22] The standard deviation of the vote share in ω is $(\frac{q(\omega))(1-q(\omega)_n}{n})^{\frac{1}{2}}$. Note that $\lim_{n\to\infty} q(\alpha)_n - q(\beta)_n = 0$, which implies that the ratio of the standard deviations converges to 1 as $n \to \infty$.

event is not bounded, that is $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\sigma_n^*)}{\Pr(\alpha|\text{piv};\sigma_n^*)} = \frac{\Phi(\delta_\alpha)}{\Phi(\delta_\beta)} \in \{0,1\}$. But then $\lim_{n\to\infty} q(\omega;\sigma_n^*) \in \{\Phi(0),\Phi(1)\}$ and $\delta_\omega = \infty$ for all states since $\Phi(0) < \frac{1}{2}$ and $\Phi(1) > \frac{1}{2}$. Second, when $\delta_\alpha \in \mathbb{R}$, Lemma 11 implies that there is a state ω for which $\delta_\omega = \infty$. However, this contradicts with the observation that $\delta_\omega \in \mathbb{R}$ for all states. Consequently, it must be that $|\delta_\omega| = \infty$ for all states ω. This finishes the proof of Lemma 9 for the case when $\Pr(\alpha) < \hat{p}$ and $W(\ell,\hat{p}) < W(s,\hat{p})$. The proof of the other cases is analogous.

Note that in the same way that I constructed an equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ with $q(\alpha;\sigma_n^*) \leq \frac{1}{2} \leq q(\beta;\sigma_n^*)$ in the case when $\hat{p} > \Pr(\alpha)$ and $W(\ell,\hat{p}) < W(s,\hat{p})$, I can construct an equilibrium sequence with $q(\alpha;\sigma_n^*) \leq q(\beta;\sigma_n^*) \leq \frac{1}{2}$ for n large enough, in this case; essentially since the likelihood of the pivotal event as a function of the vote shares $\mathbf{q}$ is symmetric around $\frac{1}{2}$. The proof is verbatim and only the relevant inequality signs need to be flipped. Conversely, if $\hat{p} > \Pr(\alpha)$ and $W(\ell,\hat{p}) > W(s,\hat{p})$, I can construct an equilibrium sequence with $\frac{1}{2} \leq q(\alpha;\sigma_n^*) \leq q(\beta;\sigma_n^*)$ for n large enough. By symmetry of the other cases, I conclude:

Theorem 2 *Let* $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$. *Take any preference distribution* H *for which* $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$ *and* $W(\ell,p) \neq W(s,p)$ *for any* $p \in [0,1]$ *with* $\Phi(p) = \frac{1}{2}$. *There is an equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *satisfying*

$$\lim_{n\to\infty} \Pr(A|\alpha;\sigma_n^*,n) = \lim_{n\to\infty} \Pr(B|\beta;\sigma_n^*,n)$$
$$= \begin{cases} 0 & \textit{if } \operatorname{sgn}(W(\ell,\hat{p}) - W(s,\hat{p})) \neq \operatorname{sgn}(\Pr(\alpha)) - \hat{p}) \\ 1 & \textit{if } \operatorname{sgn}(W(\ell,\hat{p}) - W(s,\hat{p})) = \operatorname{sgn}(\Pr(\alpha)) - \hat{p}). \end{cases} \tag{73}$$

where $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*,n)$.

4.2 The Degree of Political Conflict

In this section, consider the settings witout partisans,

$$\Pr(\ell) = 1 - \Pr(s) > 0, \tag{74}$$

mainly for the simplicity of exposition. Theorem 1 shows that information cost can screen intensities perfectly and this implies that the contrarians, which represent a minority in expectation, might win the election. Conversely, without information cost, all voters get perfectly informed and elections choose the median-voter preferred outcome in each state.

When the welfare at stake for both groups is very similar, screening becomes more difficult. Intuitively, this benefits the group of the aligned voters, which is larger in expectation. For any preference distribution H satisfying (50) - (76) and $\Pr(\mathcal{s}) > 0$, the *degree of conflict* is

$$C(H) = \Big[\Big| \frac{\Pr(\ell)}{\Pr(\mathcal{s})} \frac{\mathrm{E}(k(t)|t \in \ell)}{\mathrm{E}(k(t)|t \in \mathcal{s})} - 1 \Big| \Big]^{-1}. \tag{75}$$

The degree of conflict is larger, when the welfare at stake for the voter groups, aligned and contrarians, is more similar in expectation, thereby capturing how contested the election is.

The next result verifies the intuition and shows that the election being more contested is typically in the interest of the group that is larger in expectation. For this result, I focus on the informative equilibrium sequence of Lemma 9[23] and preference distributions such that all voters of the same group g share the same (total) intensity $k_g > 0$, that is, for all $g \in \{\mathcal{s}, \ell\}$ there is a $k_g > 0$ such that

$$t \in g \Rightarrow k(t) = k_g. \tag{76}$$

I show that, when the degree of conflict is arbitrarily large, the outcome preferred by the contrarians, which are a minority in expectation, is not elected, except for an arbitrarily small set of cost functions.

Theorem 3 *There is a function $\epsilon : \mathbb{R}^{>0} \to \mathbb{R}^{>0}$ with $\lim_{\kappa\to\infty} \epsilon(\kappa) = 0$ such that the following holds. Take $\kappa > 0$ and any preference distribution H with $C(H) > \kappa$, and satisfying (50), (74) (76), and $W(\ell, p) \neq W(\mathcal{s}, p)$ for any*

[23] The later characterization of all equilibrium sequences in Section 5.2 shows that, generically, there are no other informative equilibrium sequences when $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$.

$p \in [0,1]$ *with* $\Phi(p) = \frac{1}{2}$. *Then, for a measure* $1 - \epsilon(\kappa)$ *of cost functions with* $\lim_{x\to\infty} \frac{c'(x)x}{c(x)}$, *there is an equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *such that*

$$\lim_{n\to\infty} \Pr(A|\alpha;\sigma_n^*) = \lim_{n\to\infty} \Pr(B|\beta;\sigma_n^*) = 1, \tag{77}$$

Proof. Fix the elastictiy $d = \lim_{x\to\infty} \frac{c'(x)x}{c(x)} > 3$ and the preference distribution. Take any equilibrium sequence $(\sigma_n^*)_{n\in\ltimes}$ and let $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*, n)$. Suppose that $\frac{W(\ell,\hat{p})}{W(s,\hat{p})} < 1$. Given (50) and (76),

$$\frac{W(\ell,\hat{p})}{W(s,\hat{p})} < 1 \Leftrightarrow \left[\frac{k_\ell}{k_s}\right]^{\frac{2}{d-1}} < \frac{\Pr(s)}{\Pr(\ell)} \tag{78}$$

Given the definition (75), $\kappa \to \infty$ implies $\left[\frac{k_\ell}{k_s}\right] \to \frac{\Pr(s)}{\Pr(\ell)}$. Since $0 < \frac{\Pr(s)}{\Pr(\ell)} < 1$, for any κ large enough, the ratio $\left[\frac{k_\ell}{k_s}\right]^{\frac{2}{d-1}}$ is strictly increasing in d and the derivative has a positive lower bound.[24] Thus, for κ arbitrarily large, $\frac{W(\ell,\hat{p})}{W(s,\hat{p})} < 1$ can only hold for elasticities d arbitrarily close to $d = 3$. Consequently, Lemma 9 implies the theorem. ■

4.3 Polarization of Preferences

This section argues that groups of voters that share common interests are less likely to win an election when the preference intensities vary more strongly across the voters of the group.

For any preference distribution H and $g \in \{\ell, s\}$, let F_H^g be the conditional distribution of the threshold of doubt $y(t)$ of the types $t \in g$, and let J_H^g be the conditional distribution of the (total) intensities $k(t)$ of the types $t \in g$. A

[24]Calculation shows $\frac{\partial a^{\frac{2}{d-1}}}{\partial d} = -\frac{2a^{\frac{2}{d-1}} log(a)}{(d-1)^2}$ which is strictly bounded above zero when $0 < a < 1$.

distribution H is a *g-intensity spread of H* if

$$\begin{aligned} H(-|t \in g') &= H'(-|t \in g') \text{ for } g' \neq g \in \{s, \ell\}, && (79) \\ F_H^g &= F_{H'}^g, && (80) \\ J_H^g &<_{\text{mps}} J_{H'}^g, && (81) \end{aligned}$$

where (81) means that $J_{H'}^g$ is a mean-preserving spread of J_H^g.

The next result shows that a voter group acquires less information relative to other groups when the preference intensitities are more dispersed within the group. In other words, when the intensities are more polarized, this aggravates the free-rider problem of the group relative to the free-rider problem of the other groups. Given Lemma 8, the ratio $\frac{W_H(\ell,\hat{p})}{W_H(s,\hat{p})}$ of the power indices of the voter groups pins down the ratio of the equilibrium amount of information acquired by the voter groups, where for clarity I added the preference distribution as a subscript.

Lemma 13 *Let* $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$. *Let* $g \in \{s, \ell\}$. *Take preference distributions* H, H' *satisfying (50) - (52). Take any equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *given* H *and any equilibrium sequence* $(\hat{\sigma}_n)_{n\in\mathbb{N}}$ *given* H' *such that* $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n) = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \hat{\sigma_n}, n)$. *If* H' *is a g-intensity spread of* H, *for* $g' \neq g \in \{s, \ell\}$,

$$\frac{W_{H'}(g,\hat{p})}{W_{H'}(g',\hat{p})} < \frac{W_H(g,\hat{p})}{W_H(g',\hat{p})}. \tag{82}$$

The proof is in Section C of the Appendix.

I use the previous results of Lemma 8 and Lemma 9 that characterize equilibrium outcomes depending on the order of the power indices and show that when the intensities within a voter group are sufficiently dispersed, there is no equilibrium sequence where the outcome preferred by the group is more likely to be elected in both states. Second, there is an equilibrium sequence where the outcome that is preferred by the group is *never* elected as $n \to \infty$.

Theorem 4 *Let* $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$. *Let* $g = \ell$ $(g = s)$. *Take any preference distribution* H *satisfying (50) - (52). There is a* g*-intensity spread* H' *of* H *such that:*

1. *For all equilibrium sequences* $(\sigma_n^*)_{n\in\mathbb{N}}$ *given* H',

$$\lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*,n) \le (\ge)\frac{1}{2} \quad \text{or} \quad \lim_{n\to\infty}\Pr(B|\beta;\sigma_n^*,n) \le (\ge)\frac{1}{2}.$$

2. *There is an equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *given* H' *such that*

$$\lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*,n) = \lim_{n\to\infty}\Pr(B|\beta;\sigma_n^*,n) = 0(1).$$

The proof is in Section C of the Appendix.

5 Characterization of All Equilibria

5.1 Tyranny of the Uninformed Majority

The two-type example in the introduction illustrated how information acquisition can be complementary in election settings and how this complementarity entails a multiplicity of equilibria.

In this section, I show that, generically, there is an equilibrium sequence where the outcomes are as if all voters would have no access to further information about the state when the electorate is large. The theorem implies that when signals of low precision are sufficiently cheap, i.e. when $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$, generically, there are multiple types of equilibrium sequences: an informative equilibrium sequence (see Lemma 9) and an equilibrium sequence that is not informative.

Theorem 5 *Let* $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$. *There exists an equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *for which*

$$\lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*,n) = \lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*,n) = \begin{cases} 1 & \text{if } \Phi(\Pr(\alpha)) > \frac{1}{2}, \\ 0 & \text{if } \Phi(\Pr(\alpha)) < \frac{1}{2}, \end{cases} \tag{83}$$

Proof. Recall that equilibrium can be alternatively characterized in terms of the vector of the expected vote shares of outcome A in state α and β, i.e. (60). Let

$$\mathrm{B}_\delta = \{\mathbf{q} = (q(\alpha), q(\beta)) : |\mathbf{q} - (\Phi(\Pr(\alpha)), \Phi(\Pr(\alpha)))| < \delta. \tag{84}$$

Given $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$, there is $\delta > 0$ small enough such that the vote shares in B_δ are all larger than $\frac{1}{2}$ or all smaller than $\frac{1}{2}$. Hence, it follows from (13) that the likelihood of the pivotal event is exponentially small, given any such $\mathbf{q}$. The verbatim argument of the proof of Lemma 7 implies (109), i.e.

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}, n) = \Pr(\alpha). \tag{85}$$

The intuition is that the precision of each voter is exponentially small since the value of information is proportional to the likelihood of the pivotal event. Consequently, the difference of the vote shares in the two states is exponentially small, which implies that that the pivotal event contains no information about the relative probability of α and β as the electorate grows large. Consequently, voting according to the prior belief is optimal as the electorate grows large.
Now, take any $\mathbf{q} \in \mathrm{B}_\delta$. Lemma 6 together with (85) implies that the vote shares of $\sigma^{\mathbf{q}}$ are again in $\mathbf{B}_\delta$ when n is large enough, i.e. $\mathbf{q}(\sigma^{\mathbf{q}}) \in \mathrm{B}_\delta$. An application of Kakutani's fixed point theorem shows that there is a sequence of equilibrium vote shares $(\mathbf{q}^*_n)_{n\in\mathbb{N}}$, i.e. vote shares satisfying (60), and for all $\omega \in \{\alpha, \beta\}$,

$$\lim_{n\to\infty} q^*_n(\omega) = \Phi(\Pr(\alpha)). \tag{86}$$

The theorem follows from the weak law of large numbers, given $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$.
■

I state a corollary of the proof of Theorem 5 that will be useful for the characterization of all equilibrium sequences in Section 5.2.

Corollary 2 *Take any equilibrium sequence* $(\sigma^*_n)_{n\in\mathbb{N}}$. *Either* $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma^*_n, n) =$

$\Pr(\alpha)$ *or* $\Phi(\hat{p}) = \frac{1}{2}$ *for* $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n)$.

5.2 Characterization Result

In this section, I finish the characterization of equilibrium sequences.

Theorem 6 *Take any preference distribution H satisfying $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$ and $W(\ell, p) \neq W(s, p)$ for all $p \in \Phi^{-1}(\frac{1}{2})$.*

1. *If $\lim_{x\to 0} \frac{c'(x)x}{c(x)} < 3$, all equilibrium sequences satisfy (83).*
2. *If $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$, there are three types of equilibrium sequences. There is an equilibrium sequence satisfying (54). There is an equilibrium satisfying (73) , and there is an equilibrium sequence satisfying (83). Any equilibrium sequence satisfies either (54), (73) or (83).*

Let $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$. Then, Lemma 9 establishes the existence of an equilibrium sequence satisfying (54). Theorem 5 establishes the existence of an equilibrium sequence satisfying (83). Theorem 2 establishes the existence of an equilibrium sequence satisfying (73). The remaining parts of the proof of Theorem 6 are in Section D of the Appendix.

Using the characterization of equilibrium sequences, I prove Observation 1 in Section D of the Appendix.

6 Extensions

6.1 Heterogenous Cost

In this section, I show that cost of information and preference intensities are strategically equivalent such that the previous results extend to a setting where the information cost of the voters are heterogenous. I show, in particular, that there are equilibria with first-best outcomes when the heterogeneity of the cost

is independent of the ordinal preference type of the voters.
Let the information cost of the voters depend on a private type γ. For a given cost function c, a voter of *effort type* γ pays $c(\gamma, x) = \gamma c(x)$ for a signal of precision x. The effort type γ is distributed independently and identically across voters according to some distribution G, with density g, the *effort type distribution*, and independently of the the signals of the voters and the preference types of the other voters. The support of G is bounded below by a strictly positive constant.
The model with heterogenous information cost is equivalent to a model with homogenous cost and a suitable distribution of preferences since the best response of an aligned or contrarian voter with effort type γ, total intensity k and threshold of doubt y is the same as that of the voter with effort type $\gamma' = 1$, total intensity $\frac{k}{\gamma}$ and threshold of doubt y, given the characterization of the best response, (6)-(9), (19), (29) and (30).
Consider the settings with heterogenous cost where the effort type is independent of the voter's own preference type, that is $G \perp\!\!\!\perp H$, and where the distribution H of the preference types satisfies (50) - (52). Then, the corresponding preference distribution with homogenous cost types satisfies (50) - (52) also. As a consequence, Theorem 1 extends to settings with heterogenous information cost. With the appropriate extension of the definitions to the heterogenous cost model:

Theorem 7 *Take any preference distribution H that satisfies (50)- (52) and $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$, and for which $W(\ell, p) \neq W(s, p)$ for any $p \in [0, 1]$ with $\Phi(p) = \frac{1}{2}$. Take any effort type distribution G that is independent of H. There is an open set of cost functions $c(\cdot)$ and an equilibrium sequence $(\sigma_n^*)_{n \in \mathbb{N}}$ given c for which the welfare-maximizing outcome is elected with probability converging to 1 as $n \to \infty$.*

I conjecture that theorem 7 extends to settings where the distribution of effort types has a support with lower end $\gamma = 0$, and more generally, that the equilibrium representation through power indices (as in Lemma 8) extends, given effort types that are independent of the ordinal preference types: for any fixed γ, certainly, the ratio of the power indices captures the ratio of the information acquired. But then, this remains true, when integrating over γ.

6.2 Rich vs Coarse Choice of Information Quality

In this section, I illustrate that the 'richness' of the informational choice set of the voters matters. For this, I compare the previous results of the model where voters choose the quality of their information from a continuous set to a setting where the informational choice of the citizens is coarse. Interestingly, in the setting with a coarse information choice, there are *always* equilibrium sequences where, in each state, the welfare-maximizing outcome is elected, whereas in the continuous setting this depends on the degree of conflict κ, that is how similar the welfare at stake is for voters that gain from policy A and those that are harmed by it; see Section 4.2. This finding suggests that the richness of the informational choice present in modern societies might sometimes be harmful to social welfare. This is particularly surprising, since a richer choice set for the voters should intuitively facilitate the screening of intensities.

Binary Precision Setting. A random voter is of an aligned type with probability $\lambda = \Pr(\ell) > \frac{1}{2}$, and of a contrarian type with probability $1 - \lambda = \Pr(s)$. All types share a common threshold of doubt $y(t) = \frac{1}{2}$. Each voter can choose an uninformative signal at no cost or a binary, symmetric signal of precision $\frac{1}{2} + x$ with $x > 0$ at a cost $c \geq 0$, $\Pr(a|\alpha) = \Pr(b|\beta) = \frac{1}{2} + x$. The cost c are drawn independently and identically across voters from the uniform distribution and are the private information of the voter. The cost of a voter is independent of the (preference) types and the signals of the voters. The common prior of the voters is uniform.

Now, take any symmetric strategy σ_n, meaning that $|q(\alpha;\sigma_n) - \frac{1}{2}| = |q(\beta;\sigma_n) - \frac{1}{2}|$. Given (12), this implies that $\Pr(\alpha|\text{piv};\sigma_n, n) = \frac{1}{2}$. The expected utility from the pivotal event of an aligned voter who chooses the informative signal is

$$\Pr(\text{piv})\Big[k_\ell \Pr(\alpha|\text{piv};\sigma)(\frac{1}{2}+x) - k_\ell(\Pr(\alpha|\text{piv};\sigma)(\frac{1}{2}-x)\Big] - c \tag{87}$$
$$= \Pr(\text{piv})\frac{k_\ell}{2}x - c. \tag{88}$$

To see why, note that when the citizen is pivotal and votes B, she receives utility of zero. When she is pivotal and votes A, she receives utility of k_ℓ or $-k_\ell$, depending on the state. Note that Lemma 1 extends to this setting. Given Lemma 3, she only votes A after a; and she receives a in α with probability $\frac{1}{2}+x$ and in β with probability $\frac{1}{2}-x$.
Thus, it is optimal for an aligned voter to choose the informative signal if and only if $c \leq \Pr(\text{piv})\frac{k}{2}x$. Given that the cost are uniformly distributed, the likelihood-weighted mean precision of the aligned types is

$$\begin{aligned}\int_{t\in\ell} x(t)dH(t) &= \int_{t\in\ell} \frac{k(t)}{2}\Pr(\text{piv})x dH(t)\\ &= \Pr(\text{piv})x\lambda \mathrm{E}_H(k(t)|t\in\ell).\end{aligned} \tag{89}$$

Similarly, the likelihood-weighted mean precision of the contrarian types is

$$\int_{t\in s} x(t)dH(t) = \Pr(\text{piv})x(1-\lambda)\mathrm{E}_H(k(t)|t\in s). \tag{90}$$

When the types that choose the uninformative signal vote $50-50$ for policy A and B, the vote shares of the best response are symmetric to $\frac{1}{2}$. An application of Kakutani's fixed point theorem implies that, for all n, there is a symmetric equilibrium σ_n^*. In this equilibrium, the vote shares are ordered as follows,

$$q(\alpha;\sigma_n^*) > q(\beta;\sigma_n^*) \tag{91}$$
$$\Leftrightarrow \lambda\mathrm{E}_H(k(t)|t\in\ell) > (1-\lambda)\mathrm{E}_H(k(t)|t\in s). \tag{92}$$

Note that the preference distribution satisfies (51) and (50) such that the welfare-maximizing outcome is given by (53). Thus, (91) implies that the welfare-maximizing policy is more likely to be elected in each state. The next result shows that the welfare-maximizing outcome is elected with probability converging to 1 in each state as the electorate grows large.

Theorem 8 *Consider the setting with a binary choice of the signal precision. For any preference distribution H with $(1-\lambda)\mathrm{E}_H(k(t)|t \in s) \neq \lambda \mathrm{E}_H(k(t)|t \in \ell)$, there is an equilibrium sequence $(\sigma_n)_{n\in\mathbb{N}}$, for which the welfare-maximizing outcome is elected with probability converging to* 1.

Proof. Consider a sequence $(\sigma_n)_{n\in\mathbb{N}}$ of symmetric equilibria, as I just constructed. Clearly, Lemma 12 is valid in this setting such that it remains to show that $\delta_\omega \in \{\infty, -\infty\}$ for all $\omega \in \{\alpha, \beta\}$. By symmetry of the equilibria, $\delta_\alpha = -\delta_\beta$. Suppose that $\delta_\alpha \in \mathbb{R}$. Given (89) and (90),

$$\begin{aligned} & (q(\alpha) - \frac{1}{2})s(\alpha;\sigma_n^*)^{-1} \\ = \quad & \frac{1}{2}\Pr(\mathrm{piv})x(\Big[\mathrm{E}_H(k(t)|t \in \ell) - (1-\lambda)\mathrm{E}_H(k(t)|t \in s)\Big] \end{aligned}$$

Multiplication of both sides with the inverse of the standard deviation $s(\alpha;\sigma_n^*)^{-1} = \frac{q(\omega;\sigma_n^*)(1-q(\omega;\sigma_n^*)}{2n+1}$ and taking limits $n \to \infty$ and using Lemma 10 yields

$$\delta_\alpha = \lim_{n\to\infty} s(\alpha;\sigma_n^*)^{-1}\phi(\delta_\alpha)x(\Big[\mathrm{E}_H(k(t)|t \in \ell) - (1-\lambda)\mathrm{E}_H(k(t)|t \in s)\Big] \quad (93)$$

The right hand side diverges, and, by assumption, the left hand side does not diverge, $\delta_\alpha \in \mathbb{R}$. This yields a contradiction. Hence, $\delta_\alpha = -\delta_\beta \in \{\infty, -\infty\}$, and the theorem follows from Lemma 12 and (91). ■

6.3 Aggregate Cost

I show that the sum of the voters' cost converges to zero in all equilibrium sequences when $\lim_{x\to 0}\frac{c'(x)x}{c(x)} \neq 3$. This shows that the equilibrium sequences with first-best outcomes imply first-best results, even when taking into account the cost of voters.

Theorem 9 *Let* $\lim_{x\to 0}\frac{c'(x)x}{c(x)} \neq 3$. *Take any equilibrium sequence* $(\sigma_n)_{n\in\mathbb{N}}$ *and* *let x_i be the realisation of the precision of voter $i \in \{1, \dots, 2n+1\}$. Then,*

$$\lim_{n\to\infty} \Big[\sum_{i=1,\dots,2n+1} c(x_i) \Big] = 0. \tag{94}$$

The proof is provided in Section E of the Appendix.

7 Related Literature

Voting Cost and Vote-Buying Literature. The paper is related to work on elections with voting cost and vote-buying. Krishna & Morgan (2011) and Krishna & Morgan (2015) have shown that elections yield first-best outcomes when voting is voluntary and costly. These results are most similar to the binary precision setting of Section 6.2 where voters have a binary choice between an informative signal and an uninformative signal and there is always an equilibrium where the first-best outcome is elected. The similarity in results is somewhat surprising, since in the binary precision setting all citizens vote (and have strict incentives to do so) such that, intuitively, the free-rider problem is more severe relative to the voting cost models.

The main model of this paper is more closely related to the literature on vote buying. Lalley & Weyl (2018) have shown that equilibrium outcomes of a large electorate are first-best when each voter can buy any number of votes at a total price quadratic in the number bought. Similarly, in this model, there are equilibria with first-best outcomes for almost all preference distributions when the cost of information is arbitrarily close to 'cubic' as a function of the precision of the signal, i.e. $d = 3 + \epsilon$ for some $\epsilon > 0$.

Appendices

A Information Acquisition

A.1 Proof of Lemma 2

The benefits of information acquisition depend on how often the voters' ballot decides the election, that is the likelihood of the pivotal event. Consider an aligned type. The expected utility of an aligned type who chooses precision $x > 0$ is given by the expected utility, K, from all the events when her vote does not affect the outcome, by the cost, and by the expected utility from the pivotal event. The expected utility from the pivotal event is

$$\Pr(\text{piv})\Big[t_\alpha \Pr(\alpha|\text{piv};\sigma)(\frac{1}{2}+x) + t_\beta(\Pr(\alpha|\text{piv};\sigma)(\frac{1}{2}-x)\Big]. \tag{95}$$

To see why, note that when the citizen is pivotal and votes B, she receives utility of zero. When she is pivotal and votes A, she receives utility of t_ω, depending on the state. Given Lemma 3, it is optimal to vote A only after a; and she receives a in α with probability $\frac{1}{2}+x$ and in β with probability $\frac{1}{2}-x$. Equating marginal cost and marginal benefits gives

$$c'(x) = \Pr(\text{piv}|\sigma')\Big[\Pr(\alpha|\text{piv};\sigma')t_\alpha - \Pr(\beta|\text{piv};\sigma')t_\beta\Big]. \tag{96}$$

I claim that there is a unique solution $x^*(t;\sigma',n)$ to (96) when n is large enough. First, the marginal cost are strictly increasing since c is strictly convex. So, any solution to (96) is unique. Since $\lim_{x\to 0}\frac{c'(x)}{x^{d-1}} \in \mathbb{R}$ for some $d > 1$ and since c is continuously differentiable, marginal cost at zero are zero, i.e. $c'(0) = 0$. Given $c'(0) = 0$, $c'(1) > 0$ and since $0 \le \Pr(\text{piv}|\sigma', n) < c'(1)$ for any n large enough, given (13), it follows from the intermediate value theorem that there is a solution to (96). It follows from the implicit function theorem that $x^*(t;\sigma',n)$ is continuously differentiable. The argument for the contrarian types is analogous. This finishes the proof of the lemma.

A.2 Proof of Lemma 3.

Lemma 14 $\lim_{x\to 0}\frac{xc'(x)}{c(x)} = d$.

Proof. Recall that $\lim_{n\to\infty}\frac{c'(x)}{x^{d-1}} \in \mathbb{R}$. Since $\lim_{x\to 0} c(x) = c(0) = 0$, it follows from l' Hospital's rule that $\lim_{x\to 0}\frac{c(x)}{x^d} = \lim_{x\to 0}\frac{1}{d}\frac{c'(x)}{x^{d-1}}$. Therefore, $\lim_{n\to\infty}\frac{c(x)}{c'(x)x} = \lim_{x\to 0}\frac{c(x)}{x^d}\frac{x^d}{c'(x)x} = \frac{1}{d}$. ■

Step 1 *There is $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$: take any strategy σ' and any type t. It is a best response for a type t to acquire information $x(t) > 0$ if and only if*

$$\frac{1}{2} + x^{**}(t) \geq \frac{\Pr(\beta|\text{piv};\sigma',n)\phi_g(k)}{\Pr(\alpha|\text{piv};\sigma',n)(1-\phi_g(k)) + \Pr(\beta|\text{piv};\sigma',n)\phi_g(k)}, \quad (97)$$

$$\frac{1}{2} - x^{**}(t) \leq \frac{\Pr(\beta|\text{piv};\sigma',n)\psi_g(k)}{\Pr(\alpha|\text{piv};\sigma',n)(1-\psi_g(k)) + \Pr(\beta|\text{piv};\sigma',n)\psi_g(k)}, \quad (98)$$

where $x^{**}(t;\sigma',n) = x^*(t;\sigma',n)(1 - \frac{c(x^*(t;\sigma',n))}{x^*(t;\sigma',n)c'(x^*(t;\sigma',n))})$ and $x^*(t;\sigma',n)$ is the unique solution to the first-order condition (19).

Consider an aligned type. Recall Lemma 2 which says that if an aligned type acquires information, then $x(t) = x^*(t;\sigma',n)$. Recall Lemma 1 which says that if an aligned type acquires information, the type votes A after a and B after b. So, the comparison of the expected utility from voting A or B without further information and the expected utility from choosing $x = x^*(t;\sigma',n)$ and voting A after a and B after b minus the cost $c(x^*(t;\sigma',n))$, pins down the aligned types that acquire information.

Using (95), an aligned type prefers choosing precision $x = x^*(t;\sigma',n)$ over voting A without further information if

$$\begin{aligned}&\Pr(\text{piv}|\sigma',n)\Big[\Pr(\alpha|\text{piv};\sigma',n)(\frac{1}{2}+x)t_\alpha + \Pr(\beta|\text{piv};\sigma',n)(\frac{1}{2}-x)t_\beta\Big] - c(x)\\ \geq\ &\Pr(\text{piv}|\sigma',n)\Big[\Pr(\alpha|\text{piv};x,\sigma',n)t_\alpha + \Pr(\beta|\text{piv};\sigma',n)t_\beta\Big]. \quad (99)\end{aligned}$$

I rewrite (99) as

$$
\begin{aligned}
& \Pr(\text{piv}|\sigma', n)\Big[(\frac{1}{2}+x)\Big[\Pr(\alpha|\text{piv};\sigma',n)t_\alpha - \Pr(\beta|\text{piv};\sigma',n)t_\beta\Big] + \Pr(\beta|\text{piv};\sigma',n)t_\beta\Big] - c(x) \\
\geq \ & \Pr(\text{piv}|\sigma', n)\Big[\Pr(\alpha|\text{piv};\sigma',n)t_\alpha - \Pr(\beta|\text{piv};\sigma',n)t_\beta + 2\Pr(\beta|\text{piv};\sigma',n)t_\beta\Big] \qquad (100)
\end{aligned}
$$

Let $x = x^*(t;\sigma',n)$ in the following. Plugging (96) into (100) gives

$$
\begin{aligned}
& (\frac{1}{2}+x)c'(x) - c(x) + \Pr(\text{piv}|\sigma',n)\Pr(\beta|\text{piv};\sigma')t_\beta \\
\geq \ & c'(x) + 2\Pr(\text{piv}|\sigma',n)\Pr(\beta|\text{piv};\sigma',n)t_\beta. \qquad (101)
\end{aligned}
$$

I divide by $c'(x)$ rearrange, and use (96) again,

$$
(\frac{1}{2}+x) - \frac{c(x)}{c'(x)} \geq 1 + \frac{\Pr(\beta|\text{piv};\sigma',n)(t_\beta)}{\Pr(\alpha|\text{piv};\sigma',n)t_\alpha + \Pr(\beta|\text{piv};\sigma',n)(-t_\beta)}. \qquad (102)
$$

I use (21) and (22),

$$
(\frac{1}{2}+x) - \frac{c(x)}{c'(x)} \geq 1 + \frac{-\Pr(\beta|\text{piv};\sigma',n)y(t)}{\Pr(\alpha|\text{piv};\sigma',n)(1-y(t)) + \Pr(\beta|\text{piv};\sigma',n)y(t)}. \qquad (103)
$$

Rearranging gives (98). In the same way one shows that an aligned type prefers choosing precision $x = x^*(t;\sigma',n)$ over voting B without further information only if (97) holds. The argument for the contrarian types is analogous. This finishes the proof of the first step.

Fix a group of voter types, aligned or contrarians. Recall that any type of the group is uniquely determined by the pair $(y(t), k(t))$, see (21) and (22).

Step 2 *For any n large enough: for any $g \in \{s, \ell\}$ and for any $k > 0$,*

$$
\begin{aligned}
& \frac{1}{2} + x^{**}(y,k) - \chi(y), \quad \textit{and} \qquad (104) \\
& \frac{1}{2} - x^{**}(y,k) - \chi(y). \qquad (105)
\end{aligned}
$$

are strictly increasing in $y \in (0,1)$, where $\chi(y) = \frac{\Pr(\beta|\text{piv};\sigma',n)y}{\Pr(\beta|\text{piv};\sigma',n)y + \Pr(\alpha|\text{piv};\sigma',n)(1-y)}$. For any $\epsilon > 0$, there is $\delta > 0$ such that the derivatives of (104) and (105) are *bounded below by δ.*

Consider the derivatives of the summands individually: recall that $x^*(y,k;\sigma',n)$ is implicitly defined by the first-order condition 19. First, it follows from implicit differentiation[25] that the derivative of $x^*(y,k;\sigma',n)$ with respect to y converges to zero uniformly as $n \to \infty$. Intuitively, changes in the marginal benefit on the right hand side of (19) due to changes in the threshold of doubt y are of an order of the marginal cost $c'(x(y,k))$. These changes relate to infinitely smaller changes in the equilibrium precision $x^*(y,k;\sigma',n)$ on the left hand side since the derivative of the marginal cost, $c''(x(y,k))$, is much larger than the marginal cost $c'(x(y,k))$ when the precision is small. Lemma 14 together with (28) implies that the derivative of $x^{**}(y,k;\sigma',n)$ with respect to y converges to zero uniformly as $n \to \infty$.
Second, note that χ is continuously differentiable in y; moreover, for any $\epsilon > 0$, there is $\delta > 0$ such that $\chi'(y) > \delta$ for any $y \in (\epsilon, 1-\epsilon)$ and any n.[26] I conclude that for n large enough, $\frac{1}{2} + x^{**}(y,k) - \chi(y)$ and $\frac{1}{2} - x^{**}(y,k) - \chi(y)$ are strictly increasing.

Step 3 *For any n large enough, there are $\phi_g(k), \psi_g(k)$ with $\phi_g(k) < \hat{y} < \psi_g(k)$ such that (29) and (30) hold, and*

$$\chi(y) \leq \frac{1}{2} + x^{**}(y,k) \quad \Leftrightarrow \quad y \leq \psi_g(k), \tag{106}$$

$$\chi(y) \geq \frac{1}{2} - x^{**}(y,k) \quad \Leftrightarrow \quad y \geq \phi_g(k). \tag{107}$$

Note that $\chi(\hat{y}) = \frac{1}{2}$ and $x^{**}(\hat{y},k) > 0$ for $\hat{y} = \Pr(\alpha|\text{piv};\sigma',n)$. Thus, $\chi(\hat{y}) \leq \frac{1}{2} + x^{**}(\hat{y},k)$. The claim of the step follows from Step 2 and since $\lim_{n\to\infty} x^{**}(\hat{y},k) =$

[25] For any continuously differentiable $F : [0,1] \times [0,\frac{1}{2}] \to \mathbb{R}$ and an (implicit) function $g : [0,1] \to [0,\frac{1}{2}]$ with $F(a,g(a)) = 0$, $g'(a) = \frac{\frac{\partial F(x,y)}{\partial x}|_{(x,y)=(a,g(a))}}{\frac{\partial F(x,y)}{\partial y}|_{(x,y)=(a,g(a))}}$, which is an implication of the chain rule of differentiation.

[26] For any $p \in (0,1)$, $\frac{\partial}{\partial y}\left(\frac{py}{py+(1-p)(1-y)}\right) = \frac{(1-p)p}{(p(2y-1)-y+1)^2}$. Thus, for any $\epsilon > 0$, there is $\delta > 0$ such that for all $p \in (\epsilon, 1-\epsilon)$, $\frac{\partial}{\partial y}\left(\frac{py}{py+(1-p)(1-y)}\right) > \delta$. The assumption $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma',n) \in (0,1)$ implies that, moreover, there is $\delta > 0$ such that $\chi'(y)$ is uniformly bounded below by a positive constant for any n.

0.

Finally, Lemma 3 follows from Step 1 and Step 3. Note that it follows from the implicit function theorem that the functions $\phi_g(k)$ and $\psi_g(k)$ are continuously differentiable in k.

B The Power of Voter Groups

B.1 Proof of Lemma 7

Suppose that $\lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv};\sigma_n^*)) \neq \frac{1}{2}$. Then, the pivotal likelihood is exponentially small when n is large, given (13). Thus, the first-order condition (19) implies that the precision $x(t)$ of each voter is exponentially small. It follows from (41) that the difference in the vote shares is exponentially small,

$$q(\alpha;\sigma_n^*) - q(\beta;\sigma_n^*) = \lim_{n\to\infty} 2\Big[\int_{t\in\ell} x(t)dH(t) - \int_{t\in s} x(t)dH(t)\Big] < z^n c_3 \quad (108)$$

for some constant $c_3 \neq 0$ and n large enough. Consider the voter's inference about the relative likelihood of α and β. Intuitively, given (108), the pivotal event contains no information about the relative probability of α and β as the electorate grows large. I claim that (108) implies

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*, n) = \Pr(\alpha). \quad (109)$$

To see why, consider the likelihood ratio

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n^*, n)}{\Pr(\text{piv}|\beta;\sigma_n^*, n)} = \Big[\frac{q(\alpha;\sigma_n^*)(1-q(\alpha;\sigma_n^*)}{q(\beta;\sigma_n^*)(1-q(\beta;\sigma_n^*)}\Big]^n. \quad (110)$$

The inequality (108) states that the difference in vote shares is exponentially small such that $(1-\frac{1}{n^2})^n \leq \frac{\Pr(\alpha|\text{piv};\sigma_n^*,n)}{\Pr(\beta|\text{piv};\sigma_n^*,n)} \leq (1+\frac{1}{n^2})^n$ for all n large enough. Then, the description $\lim_{n\to\infty}(1-\frac{x}{n})^n = e^x$ of the e-function implies that the likelihood ratio converges to 1 and in turn (109).

If $\Phi(\Pr(\alpha)) = \frac{1}{2}$, Lemma 7 follows from (109) since Φ is continuous. Suppose that $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$. Note that (109) implies that $\lim_{n\to\infty} q(\omega;\sigma_n^*) = \Phi(\Pr(\alpha))$ for $\omega \in \{\alpha, \beta\}$, given Lemma 6 and since Φ is continuous. Then, the weak law of large numbers implies $\lim_{n\to\infty} \Pr(A|\alpha;\sigma_n^*, n) = \lim_{n\to\infty} \Pr(B|\beta;\sigma_n^*, n) \in \{0,1\}$. This contradicts with the equilibrium sequence being informative; see the definition (36). I conclude, that any informative equilibrium sequence must satisfy (39). This finishes the proof of Lemma 7.

I record the following corollary of the proof of Lemma 7.

Lemma 15 *For any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$, $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) \in (0,1)$.*

Proof. Suppose, for example that $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) = 0$. Then, Lemma 6 implies that $\lim_{n\to\infty} q(\omega;\sigma_n^*) = \Phi(0)$ for any $\omega \in \{\alpha,\beta\}$. Since $\Phi(0) \neq \frac{1}{2}$, the proof of Lemma 7 implies (109) which is a contradiction to $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) = 0$ since $\Pr(\alpha) > 0$. The assumption $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) = 1$ leads to a contradiction in the same way. This finishes the proof of Lemma 15. ■

B.2 Proof of Lemma 8

I show a more general lemma, which implies Lemma 8.

Lemma 16 *Take any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$. For all n, let $\hat{\sigma}_n$ be any best response to σ_n. If $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n, n) \in (0,1)$, for any $g \in \{\ell, s\}$, the sequence of best responses satisfies*

$$\int_{t\in g} x(t)dH(t) \;\; \approx \;\; c_4\Big[\Pr(\text{piv}|\sigma_n^*, n)\Big]^{\frac{2}{d-1}} W(g,\hat{p})$$

for some constant $c_4 > 0$ that does not depend on g.

Proof. For the simplicity of the exposition, in the following, we consider cost functions $c(x) = \gamma x^d$ for some $d > 1$. More generally, recall that any cost function considered in this paper satisfies $\lim_{x\to 0} \frac{c'(x)}{x^{d-1}}$ for some $d > 1$ such that l'Hospital's rule implies that $\lim \frac{c(x)}{x^d} \in \mathbb{R}$. The proof for a general cost function is verbatim except that one has to replace the equality $c(x) = \gamma x^d$ with the approximation $c(x) \approx \gamma x^d$.

Notation. Recall that, when fixing the voter group $g \in \{\ell, s\}$, we can view the conditional distribution $H(-|t \in g)$ of the preference types $t \in g$ as a distribution of the threshold of doubt $y(t)$ and the total intensity $k(t)$ of the types $t \in g$, see Section 2.1.5. The marginal distribution $F(-|t \in g)$ of the threshold of doubt of the types $t \in g$ has the density $f^g(y) = \int_{k\in[0,2]} h^g(y,k)dk$ where h^g is the density of $H(-|t \in g)$. The marginal distribution J of the total intensity of the voter types $t \in g$ has the density $j^g(k) = \int_{y\in[0,1]} h^g(y,k)dy$. Similarly, let the densities of the conditional distributions $F(y'|k = k', t \in g) = \int_{y\le y'} \frac{h^g(y,k)}{j^g(k')}$ and $J(k'|y = y', t \in g) = \int_{k\le k'} \frac{h^g(y,k)}{f^g(y')}$ be $f^g(y|k = k')$ and $j^g(k|y = y')$, for any $k' \in [0, 2]$ and any $y' \in [0, 1]$. In the following, I consider $x(t)$, $x^*(t; \sigma_n, n)$ and $x^{**}(t; \sigma_n, n)$ as functions of (y, k). Recall from Section 2.1.5 that $x^*(t; \sigma_n, n)$ and $x^{**}(t; \sigma_n, n)$ only depend on the treshold of doubt and the total intensity, and are independent of the group that the type t belongs to.

Take any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n, n) \in (0, 1)$. The first step shows that the region $[\phi_g(k), \psi_g(k)]$ of types that acquire information, i.e. for which $x(y, k) > 0$, is sufficiently symmetric around the critical types with $\hat{y}_n = \Pr(\alpha|\text{piv}; \sigma_n)$. This first step implies (45), as in the sketch of proof in the main text.

Step 1 *For all* $k \in (0, 2)$,

$$\lim_{n\to\infty} \frac{\psi_g(k) - \hat{y}_n}{x^{**}(\hat{y}, k; \sigma_n, n)} = -\lim_{n\to\infty} \frac{\phi_g(k) - \hat{y}_n}{x^{**}(\hat{y}_n, k; \sigma_n, n)} = \lim_{n\to\infty} \frac{1}{\chi'(\hat{y}_n)}. \tag{111}$$

where $\hat{y}_n = \Pr(\alpha|\text{piv}; \sigma_n)$.

Let $\chi(y) = \frac{\Pr(\beta|\text{piv};\sigma_n,n)y}{\Pr(\beta|\text{piv};\sigma',n)y+\Pr(\alpha|\text{piv};\sigma_n,n)(1-y)}$, which is the left hand side of (29) and (30). Recall that $\chi(\hat{y}_n) = \frac{1}{2}$. A Taylor expansion of χ at $\hat{y}_n = \Pr(\alpha|\text{piv}; \sigma_n)$

implies

$$\phi_g(k) - \hat{y}_n = -\frac{1}{\chi'(\epsilon_{1,n})} x^{**}(\phi_g(k), k) \tag{112}$$

for some $\epsilon_{1,k,n} \in [\phi_g(k), \hat{y}_n]$. I plug in (28),

$$\phi_g(k) - \hat{y}_n = -\frac{1}{\chi'(\epsilon_{1,k,n})} x^*(\phi_g(k), k; \sigma_n, n) \Big[1 - \frac{x^*(\phi_g(k), k; \sigma_n, n) c'(x^*(\phi_g(k), k; \sigma_n, n))}{c(x^*(\phi_g(k), k; \sigma_n, n))}\Big] \tag{113}$$

A Taylor expansion of $x^*(y, k; \sigma_n, n)$ at $\hat{y}_n = \Pr(\alpha|\text{piv}; \sigma_n, n)$ yields

$$x^*(\phi_g(k), k; \sigma_n, n) = x^*(\hat{y}_n, k; \sigma_n, n) + (x^*(\epsilon_{2,k,n}, k; \sigma_n, n))' \Big[\phi_g(k) - \hat{y}_n\Big]. \tag{114}$$

for some $\epsilon_{2,n} \in [\phi_g(k), \hat{y}_n)]$. As in the proof of Lemma 16, implicit differentiation shows that the derivative of the function $x^*(y, k; \sigma_n, n)$ with respect to y converges to zero uniformly. Therefore, (113), (114) and together imply

$$\lim_{n\to\infty} \frac{\phi_g(k) - \hat{y}_n}{x^{**}(\hat{y}, k; \sigma_n, n)} = -\lim_{n\to\infty} \frac{1}{\chi'(\epsilon_{1,n})} = -\lim_{n\to\infty} \frac{1}{\chi'(\hat{y}_n)}. \tag{115}$$

where the last equality follows from the continuity of χ' and since $\epsilon_{1,k,n} \to \hat{y}_n$ as $n \to \infty$. The analogous argument for $\psi_g(k)$ shows that

$$\lim_{n\to\infty} \frac{\psi_g(k) - \hat{y}_n}{x^{**}(\hat{y}_n, k; \sigma_n, n)} = \lim_{n\to\infty} \frac{1}{\chi'(\hat{y}_n)}, \tag{116}$$

which finishes the proof of (111).

The next step shows (44), as in the sketch of the proof in the main text.

Step 2 *For all* $k \in (0, 2)$,

$$x(y, k) > 0 \Rightarrow \frac{x(y, k)}{x(y^*, k)} \approx 1. \tag{117}$$

where the convergence is uniform across (y, k).

Take any (y, k) such that $x(y, k) > 0$. Lemma 3 implies that $y \in [\phi_g(k), \psi_g(k)]$.

A Taylor expansion of $x^*(y,k;\sigma_n,n)$ at $\hat{y}_n = \Pr(\alpha|\text{piv};\sigma_n,n)$ implies

$$x(y,k) = x^*(\hat{y}_n,k;\sigma_n,n) + (x^*(\epsilon_{y,k,n},k;\sigma_n,n))'(y-\hat{y}_n) \tag{118}$$

for some $\epsilon_{y,k,n} \in [y,\hat{y}_n]$. Since $\lim_{n\to\infty} x^*(\hat{y},k;\sigma_n,n) = 0$, Lemma 14 implies that

$$\lim_{n\to\infty} \frac{x^*(\hat{y}_n,k;\sigma_n,n)}{x^{**}(\hat{y}_n,k;\sigma_n,n)} = \frac{d}{d-1}. \tag{119}$$

Hence, Step 1 implies $\lim_{n\to\infty} \frac{x^*(\hat{y}_n,k;\sigma_n,n)}{\psi_g(k)-\hat{y}_n} \in \mathbb{R}$, and, thus $\lim_{n\to\infty} \frac{x^*(\hat{y}_n,k;\sigma_n,n)}{y-\hat{y}_n} \in \mathbb{R}$ Recall that the derivative of $x^*(y,k;\sigma_n,n)$ with respect to y converges to zero uniformly. Therefore, I conclude that (118) yields (117).

The next step shows (46), as in the sketch of the proof in the main text.

Step 3 *For all $k' \in (0,2)$ and any $g \in \{\ell, s\}$,*

$$\frac{f^g(\hat{y}_n,k;\sigma_n,n|k=k')\Big[\phi_g(k')-\psi_g(k')\Big]}{F(\psi_g(k)|k(t)=k,t\in g) - F(\phi_g(k)|k=k',t\in g)} \approx 1. \tag{120}$$

A Taylor expansion of $F(-|k=k',t\in g)$ at $\hat{y} = \Pr(\alpha|\text{piv};\sigma_n,n)$ implies

$$F(\phi(k)|k=k',t\in g) = F(\hat{y}|k=k',t\in g) + f^g(\epsilon_{3,k',n}|k=k')\Big[\phi_g(k')-\hat{y}_n\Big], \tag{121}$$
$$F(\psi(k)|k=k',t\in g) = F(\hat{y}|k=k',t\in g) + f^g(\epsilon_{4,k',n}|k=k')\Big[\psi_g(k')-\hat{y}_n\Big] \tag{122}$$

for some $\epsilon_{3,k,n} \in [\phi_g(k'),\hat{y}_n]$ and some $\epsilon_{4,k,n} \in [\hat{y}_n,\psi_g(k')]$. Since $\lim_{n\to\infty} \phi_g(k') - \phi_g(k) = 0$, the continuity of $f^g(-|k=k')$ implies

$$\lim_{n\to\infty} \frac{f^g(\epsilon_{3,k',n}|k=k')}{f^g(\hat{y}_n|k=k')} = 1, \tag{123}$$
$$\lim_{n\to\infty} \frac{f^g(\epsilon_{4,k',n}|k=k')}{f^g(\hat{y}_n|k=k')} = 1. \tag{124}$$

Finally, (123) and (124) together with (121) and (122) imply (120).

The last two steps finish the proof of Lemma 8.

Step 4 *For any $g \in \{s, \ell\}$ and any $k' \in [0,2]$,*

$$\lim_{n\to\infty} \frac{\int_y x(y,k')dF(y|k=k',t\in g)}{x^*(\hat{y}_n,k';\sigma_n,n)^2} = \lim_{n\to\infty} \frac{2f^g(\hat{y}_n|k=k')}{\chi'(\hat{y}_n)}\frac{d-1}{d}. \tag{125}$$

I use the first three steps,

$$\begin{aligned}
& \lim_{n\to\infty}\int_y x(y,k')dF(y|k=k',t\in g) \\
\approx\ & \Big[F(\psi^g(k')|k=k') - F(\phi^g(k')|k=k')\Big]x^*(\hat{y}_n,k';\sigma_n,n) \\
\approx\ & f^g(\hat{y}_n,k'|k=k')\Big[\phi^g(k') - \psi^g(k')\Big]x^*(\hat{y}_n,k';\sigma_n,n) \\
\approx\ & 2\frac{f^g(\hat{y}_n,k'|k=k')}{\chi'(\hat{y}_n)}x^{**}(\hat{y}_n,k';\sigma_n,n)x^*(\hat{y}_n,k';\sigma_n,n) \\
\approx\ & 2\frac{f^g(\hat{y}_n,k'|k=k')}{\chi'(\hat{y}_n)}x^*(\hat{y}_n,k';\sigma_n,n)^2\frac{d-1}{d},
\end{aligned}$$

where I used Step 2 for the second line, Step 3 for the third line, Step 1 for the fourth line and (119) for the last line.

Step 5 *For all $g \in \{\ell, s\}$,*

$$\int_{t\in g} x(t)dH(t) \approx \frac{2(d-1)}{d}\frac{x^*(\hat{p},1;\sigma_n,n)^2}{\chi'(\hat{p})}\Pr(g)f^g(\hat{p})\mathrm{E}(k(t)^{\frac{2}{d-1}}|t\in g, y(t)=\hat{p}).$$

for $\hat{p} = \lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma_n,n)$.

First, I change coordinates,

$$\int_{t\in g} x(t)dG(t) = \Pr(g)\int_{k'\in[0,2]}\int_{y\in[0,1]} x(y,k')f^g(y,k')dydk' \tag{126}$$

and, then, apply (125) to obtain

$$\begin{aligned}
&\int_{t\in g} x(t)dG(t) \\
\approx\ & \Pr(g)\int_{k'\in[0,2]} j^g(k')x^*(\hat{y}_n,k')^2\frac{f^g(\hat{y}_n|k=k')}{\chi'(\hat{y}_n)}\frac{2(d-1)}{d}dk'. \quad (127)
\end{aligned}$$

It follows from the first-order condition (19) and since $c(x)=\gamma x^d$ that for all $k'\in[0,2]$,

$$x^*(\hat{y}_n,k';\sigma_n,n)=(k')^{\frac{1}{d-1}}x^*(\hat{y}_n,1;\sigma_n,n). \quad (128)$$

Combining (127) and (128),

$$\int_{t\in g} x(t)dG(t) \approx \Pr(g)\frac{2(d-1)}{d\chi'(\hat{y}_n)}x^*(\hat{y}_n,1;\sigma_n,n)^2\int_{k'\in[0,2]}(k')^{\frac{2}{d-1}}j^g(k')f^g(\hat{y}_n|k=k')dk'.$$

I rewrite,

$$\begin{aligned}
\int_{t\in g} x(t)dG(t) &\approx \Pr(g)\frac{2(d-1)}{d\chi'(\hat{y}_n)}x^*(\hat{y}_n,1;\sigma_n,n)^2f^g(\hat{y}_n)\int_{k'\in[0,2]}(k')^{\frac{2}{d-1}}\frac{h^g(\hat{y}_n,k')}{f^g(\hat{y}_n)}dk' \\
&= \Pr(g)\frac{2(d-1)}{d\chi'(\hat{y}_n)}x^*(\hat{y}_n,1;\sigma_n,n)^2f^g(\hat{y}_n)\int_{k'\in[0,2]}(k')^{\frac{2}{d-1}}j^g(k'|y=\hat{y}_n)dk' \\
&= \Pr(g)\frac{2(d-1)}{d\chi'(\hat{y}_n)}x^*(\hat{y}_n,1;\sigma_n,n)^2f^g(\hat{y}_n)\mathrm{E}_H((k')^{\frac{2}{d-1}}|y=\hat{y}_n) \\
&\approx c_{4,n}\Pr(g)f^g(\hat{p})\mathrm{E}_H((k')^{\frac{2}{d-1}}|y=\hat{p}) \quad (129)
\end{aligned}$$

where $c_{4,n}=\frac{2(d-1)}{d\chi'(\hat{y}_n)}x^*(\hat{y}_n,1;\sigma_n,n)^2$. I used Lemma 7 and the continuity of f^g, and $f^g(k'|y=y')$ in y' for the statement on the last line. The continuity of χ' and $x^*(\cdot,1;\sigma_n,n)$ in y implies $\lim_{n\to\infty}\frac{c_4}{c_{4,n}}=1$ for $c_4=\frac{2(d-1)}{d}\frac{x^*(\hat{p},1;\sigma_n,n)^2}{\chi'(\hat{p})}>0$. This finishes the proof of Step 4. ∎

C Welfare-Maximizing Outcomes

C.1 Proof of Lemma 11

Recall that $d = \lim_{x\to 0} \frac{c'(x)x}{c(x)}$ (see Lemma 16). What will matter is if the mean square of the precision of the critical types, i.e. $\mathrm{E}_k\Big[x(y,k)^2|y = \hat{p}\Big]$, is of an order larger or smaller than the pivotal likelihood. Consider the first-order condition, (19),

$$\begin{aligned} c'(x^*(t;\sigma_n,n)) &= \Pr(\text{piv}|\sigma_n,n)\mathrm{E}_\omega(|t_\omega||\text{piv};\sigma_n,n). \\ \Rightarrow x^*(t;\sigma_n,n)^2 &\approx \Big[\Pr(\text{piv}|\sigma_n,n)\frac{\mathrm{E}_\omega(|t_\omega||\text{piv};\sigma_n,n)}{d}\Big]^{\frac{2}{d-1}}. \end{aligned} \tag{130}$$

We see that the critical elasticity is $d = 3$. For $d > 3$, the squared precision is of an order larger than the likelihood of the pivotal event, $\lim_{n\to\infty} \frac{x^*(t;\sigma_n,n)^2}{\Pr(\text{piv}|\sigma_n,n)} = \infty$. Let $d > 3$ in the following. Then, given (48) and the observation just made, the likelihood-weighted average precision of a voter group is of an order larger than the likelihood of the pivotal event, $\lim_{n\to\infty} \frac{\int_{t\in g} x(t)dH(t)}{\Pr(\text{piv}|\sigma_n,n)} = \infty$. Take any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ with $\delta_\alpha \in \mathbb{R}$. Then, the likelihood of the pivotal event in α is of the order of the standard deviation of the vote share, $s(\alpha;\sigma_n)$, given Lemma 10. I conclude that the likelihood-weighted average precision of each voter group is as large as arbitrarily many standard deviations of the vote share in α, as $n \to \infty$, i.e. $\lim_{n\to\infty} \frac{\int_{t\in g} x(t)dH(t)}{s(\alpha;\sigma_n)} = \infty$. Hence, the first item of the lemma follows from (41), Lemma 8 and the assumption that $W(\ell,\hat{p}) \neq W(s,\hat{p})$. The proof of the second item is analogous.

C.2 Proof of Lemma 12

Proof. Let $q_n = q(\omega,\sigma_n)$. By using the normal approximation[27]

[27] For this normal approximation, we cannot rely on the standard central limit theorem, because q_n varies with n. Recall that for any undominated strategy, types t with $t_\alpha > 0, t_\beta > 0$ vote A and types t with $t_\alpha < 0, t_\beta < 0$ vote B. Hence, since the type distribution has a strictly positive density, there exists $\epsilon > 0$ such that $\epsilon < q_n < 1-\epsilon$ for all $n \in \mathbb{N}$. As a consequence, we can apply the Lindeberg-Feller central limit theorem (see Billingsley (2008), Theorem 27.2). To see why, one checks that a sufficient condition for the the Lindeberg condition is that $(2n+1)q_n(1-q_n) \to \infty$ as $n \to \infty$ since this implies that for n

$$\mathcal{B}(2n+1, q_n) \simeq \mathcal{N}((2n+1)q_n, (2n+1)q_n(1-q_n)),$$

we see that the probability that A wins the election in ω converges to

$$\Phi(\frac{\frac{1}{2}(2n+1) - (2n+1) \cdot q_n}{((2n+1)q_n(1-q_n))^{\frac{1}{2}}}).$$

Taking limits $n \to \infty$, gives

$$\begin{aligned}
& \lim_{n\to\infty} \Phi(\frac{\frac{1}{2}(2n+1) - (2n+1) \cdot q_n}{(2n+1)q_n(1-q_n))^{\frac{1}{2}}}) \\
= & \lim_{n\to\infty} \Phi(\frac{(2n+1)\frac{1}{2} - (2n+1)(\frac{1}{2} + (q_n - \frac{1}{2}))}{((2n+1)^{\frac{1}{2}}(q_n(1-q_n))^{\frac{1}{2}}}) \\
= & \lim_{n\to\infty} \Phi((q_n - \frac{1}{2})\Big[\frac{(2n+1)}{q_n(1-q_n)}\Big]^{\frac{1}{2}}) \\
= & \Phi(\delta_\omega),
\end{aligned}$$

where the equalities on the last two lines hold both when $\delta_\omega \in \{\infty, -\infty\}$ and when $\delta_\omega \in \mathbb{R}$. For the equality on the last line, I used that the standard deviation of the vote share is given by $s(\omega; \sigma_n) = (\frac{q_n(1-q_n)}{2n+1}^{\frac{1}{2}})$. ■

C.3 Proof of Lemma 13

Proof. Recall that $\lim_{x\to 0} \frac{c'(x)x}{c(x)} = d$ (Lemma 14). The basic intuition is that that power index $W(g, \hat{p})$ of the group g is proportional to $\mathrm{E}_H(k^{\frac{2}{d-1}})$. Since $k^{\frac{2}{d-1}|t\in g}$ is strictly concave, given $d > 3$, an application of Jensen's inequality shows that for any g-intensity spread H' of H and $g \neq g' \in \{\ell, s\}$,

$$\mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g) < \mathrm{E}_H(k^{\frac{2}{d-1}}|t \in g), \tag{131}$$

$$\mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g') = \mathrm{E}_H(k^{\frac{2}{d-1}}|t \in g'). \tag{132}$$

sufficiently large the indicator function in the condition takes the value zero.

First, note that by the definition of a g-intensity spread, $\Pr_{H'}(\ell) = \Pr_H(\ell)$ and $\Pr_{H'}(s) = \Pr_H(s)$; see (79)-(80). Second, $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n) = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \hat{\sigma_n}, n)$ and (80) together imply $f_H^g(\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n)) = f_{H'}^g(\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \hat{\sigma_n}, n))$. Third, since H and H' satisfy (51), $\mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g) = \mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g, y(t) = \hat{p})$ and $\mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g') = \mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g', y(t) = \hat{p})$. These observations together with (131) and (132) imply (82), given the definition (38) of the power indices. ■

C.4 Proof of Lemma 4

Consider the group of the aligned voters. Given Lemma 8 and Lemma 9, it remains to show that for any H there is a ℓ-intensity spread such that $\frac{W_{H'}(\ell,\hat{p})}{W_{H'}(s,\hat{p})} < 1$ across all equilibrium sequences. Note that $W_H(s\hat{p}) = W_{H'}(s, \hat{p})$ for any ℓ-intensity spread H' of H. I show that for any H and for any $\epsilon > 0$, there is an ℓ-intensity spread H' of H such that for all equilibrium sequences $(\sigma_n^*)_{n\in\mathbb{N}}$ given H',

$$W_{H'}(\ell, \hat{p}) < \epsilon. \tag{133}$$

Note that the density $f^g(y)$ of the threshold of doubt is continuous and therefore bounded. Given (50) and (51), it is therefore sufficient to show that for any $\epsilon > 0$, there is a a g-intensity spread with

$$\mathrm{E}_{H'}(k^{\frac{2}{d-1}}|t \in g) < \epsilon. \tag{134}$$

For this, consider g-intensity spreads $H'(\kappa)$ of H such that the conditional distributions $J^g_{H'(\kappa)}$ of the intensities k of the types of group g are concentrated in neighbourhoods around 0 and $K > 0$, that is $\Pr_{J^g_{H'(\kappa)}}(\kappa \leq k \leq \kappa + \delta) + \Pr_{J^g_{H'(\kappa)}}(0 \leq k \leq \epsilon) \geq 1 - \delta$. Since the mean of the intensities is the same under the g-intensity spread, the iterated law of expectation implies $\lim_{\delta\to 0} \Pr_{J^g_{H'(\kappa)}}(\kappa \leq k \leq \kappa + \delta)\kappa = E_{J^g_H}(k)$. Hence,

$$\lim_{\delta\to 0} \mathrm{E}_{J^g_{H'(\kappa)}}(k^{\frac{2}{d-1}}) = \frac{E_{J^g_H}(k)}{\kappa}\kappa^{\frac{2}{d-1}} \overset{\kappa\to\infty}{\to} 0, \tag{135}$$

where I used that $d = \lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$ and hence $\frac{2}{d-1} < 1$. Finally, (135) shows that $W_{H'(\kappa)}(\ell, \hat{p}) < \epsilon$ when κ is large enough and δ small enough. This finishes the proof of the theorem.

D Characterization of All Equilibria

D.1 Proof of Theorem 6

Case $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$**.**

I show that any equilibrium sequence satisfies either (54), (73) or (83).

Take any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$. It follows from Corollary 2 that $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*) \in \Phi^{-1}(\frac{1}{2}) \cup \{\Pr(\alpha)\}$. Given Lemma 8, the order of the vote shares is given by the order of $W(\ell, \hat{p})$ and $W(s, \hat{p})$. Consider the case when $W(\ell, \hat{p}) < W(s, \hat{p})$ such that there are the following two cases:

Case 1 *For any n large enough, $\frac{1}{2} \leq q(\alpha; \sigma_n^*) \leq q(\beta; \sigma_n^*)$.*

Suppose that $\delta(\alpha) \in \mathbb{R}$. Then, Lemma 11 implies that $\delta(\alpha) = \infty$, a contradiction. Hence $\delta(\alpha) = \delta(\beta) = \infty$. Lemma 12 implies that the equilibrium sequence satisfies (2).

Case 2 *For any n large enough, $q(\alpha; \sigma_n^*) \leq \frac{1}{2} \leq q(\beta; \sigma_n^*)$.*

Suppose that $\delta(\alpha) \in \mathbb{R}$. This implies $\delta(\beta) \in \mathbb{R}$. To see why, note that otherwise Lemma 10 implies that $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\sigma_n^*,n)}{\Pr(\alpha|\text{piv};\sigma_n^*,n)} = \frac{\Phi(\delta_\alpha)}{\Phi(\delta_\beta)} = 1$. Then, Lemma 6 implies that $\lim_{n\to\infty} q(\alpha; \sigma_n^*) = \Phi(1)$. This yields a contradiction to $q(\alpha; \sigma_n^*) \leq \frac{1}{2}$ since $\Phi(1) > \frac{1}{2}$ by assumption. On the other hand, Lemma 11 implies that $\delta(\alpha) \in \{\infty, -\infty\}$ or $\delta(\beta) \in \{\infty, -\infty\}$. Clearly, this contradicts the earlier observation that $\delta(\omega) \in \mathbb{R}$ for any $\omega \in \{\alpha, \beta\}$. In the same way, the assumption $\delta(\beta) \in \mathbb{R}$ leads to a contradiction. Consequently, $-\delta(\alpha) = \delta(\beta) = \infty$ and Lemma 12 implies that the equilibrium sequence satisfies (54).

This finishes the proof for the case when $W(\ell, \hat{p}) < W(s, \hat{p})$. The proof for the case when $W(\ell, \hat{p}) > W(s, \hat{p})$ is analogous.

Case $\lim_{x\to 0}\frac{c'(x)x}{c(x)} < 3$**.**

Take any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$. First, recall Corollary 2, which implies that $\hat{p} = \lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma_n^*) \in \Phi^{-1}(\frac{1}{2}) \cup \{\Pr(\alpha)\}$. If $\hat{p} = \Pr(\alpha)$, Lemma 6 together with $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$ and the law of large numbers imply that $(\sigma_n^*)_{n\in\mathbb{N}}$ satisfies (83). Let $\hat{p} \in \Phi^{-1}(\frac{1}{2})$ in the following.

Step 1 *Let* $\delta_{\omega,n} = (q(\omega;\sigma_n^*) - \frac{1}{2})s(\omega;\sigma_n^*)^{-1}$ *for* $\omega \in \{\alpha,\beta\}$*. Then,*

$$\lim_{n\to\infty}\delta_{\alpha,n} - \delta_{\beta,n} = 0. \tag{136}$$

Note that $s(\omega;\sigma_n^*) = \frac{q(\omega;\sigma_n^*)(1-q(\omega;\sigma_n^*)}{2n+1}$ such that $\hat{p} \in \Phi^{-1}(\frac{1}{2})$ together with Lemma 6 implies $\lim_{n\to\infty}\frac{s(\alpha;\sigma_n^*)}{s(\beta;\sigma_n^*)} = 1$. Note that Lemma 11 provides an upper bound for the limit of the difference of the vote shares measured in standard deviations, since it considers the case when the vote share in one or both of the states is arbitrarily close to $\frac{1}{2}$. Consequently, Lemma 11 implies (136).

Step 2 $\lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma_n^*) = \Pr(\alpha)$.

Consider the ratio of the likelihoods of the pivotal event in the two states,

$$\begin{aligned}
&\frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\beta;\sigma_n^*,n)} \\
= \;&\Big[\frac{q(\alpha;\sigma_n^*)(1-q(\alpha;\sigma_n^*)}{q(\beta;\sigma_n^*)(1-q(\beta;\sigma_n^*)}\Big]^n. \\
= \;&\Big[1 - \frac{(q(\alpha;\sigma_n^*)-\frac{1}{2})^2 - (q(\beta;\sigma_n^*)-\frac{1}{2})^2}{(q(\beta;\sigma_n^*)-\frac{1}{2})^2 - \frac{1}{4}}\Big]^n \\
= \;&\Big[1 - \frac{1}{2n+1}(\frac{q(\alpha;\sigma_n^*)(1-q(\alpha;\sigma_n^*))}{(q(\beta;\sigma_n^*)-\frac{1}{2})^2-\frac{1}{4}}\delta_{\alpha,n}^2 - \frac{q(\beta;\sigma_n^*)(1-q(\beta;\sigma_n^*))}{(q(\beta;\sigma_n^*)-\frac{1}{2})^2-\frac{1}{4}}\delta_{\beta,n}^2)\Big]^n.
\end{aligned}$$

Let

$$x_n = (\frac{q(\alpha;\sigma_n^*)(1-q(\alpha;\sigma_n^*))}{(q(\beta;\sigma_n^*)-\frac{1}{2})^2-\frac{1}{4}}\delta_{\alpha,n}^2 - \frac{q(\beta;\sigma_n^*)(1-q(\beta;\sigma_n^*))}{(q(\beta;\sigma_n^*)-\frac{1}{2})^2-\frac{1}{4}}\delta_{\beta,n}^2. \tag{137}$$

The likelihood ratio simplifies to

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\beta;\sigma_n^*,n)} = \Big[1 - \frac{1}{2n+1}x_n\Big]^n. \tag{138}$$

I rewrite,

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\beta;\sigma_n^*,n)} = (\Big[1 - \frac{1}{2n+1}x_n\Big]^n - e^{-\frac{1}{2}x_n}) + e^{-\frac{1}{2}x_n} \tag{139}$$

and analyse the two summands separately in the following. First, note that Lemma 6 together with $\hat{p} \in \Phi^{-1}(\frac{1}{2})$ implies $\lim_{n\to\infty} q(\omega;\sigma_n^*) = \frac{1}{2}$ for $\omega \in \{\alpha,\beta\}$. Thus, Step 1 and (137) together imply

$$\lim_{n\to\infty} x_n = 0. \tag{140}$$

This yields

$$\lim_{n\to\infty} e^{-\frac{1}{2}x_n} = 1, \tag{141}$$

Second, using the Lemmas 4.3 and 4.3 in Durrett (1991) [p.94], for all $y \in \mathbb{R}$,

$$|(1 - \frac{y}{(2n+1)})^n - e^{-y}| \leq \frac{y^2}{(2n+1)^3} \tag{142}$$

Finally, it follows from (139) - (142) that

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\beta;\sigma_n^*,n)} = 1, \tag{143}$$

which was to be shown.

Now, the result of Step 2 is a contradiction to $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) \in \Phi^{-1}(\frac{1}{2})$ since $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$ by assumption. Consequently, for all equilibrium sequences $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) = \Pr(\alpha)$. It follows from the law of large numbers that all equilibrium sequences satisfy (83).

D.2 Proof of Observation 1

Take any informative equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$. Let $\hat{q}(\omega;\sigma_n^*)$ be the likelihood that a randomly drawn uninformed type votes A in ω; $\hat{q}(\omega;\sigma_n^*)$ is the *vote share of the uninformed*. Let $\tilde{q}(\omega;\sigma_n^*)$ be the likelihood that a randomly drawn uninformed type votes A in ω; $\tilde{q}(\omega;\sigma_n^*)$ is the *vote share of the informed*. Now, first, Theorem 6 implies that $(\sigma_n^*)_{n\in\mathbb{N}}$ satisfies (54). Therefore, $q(\alpha;\sigma_n^*) < \frac{1}{2} < q(\beta;\sigma_n^*)$ for all n large enough or $q(\beta;\sigma_n^* < \frac{1}{2} < q(\alpha;\sigma_n^*)$ for all n large enough. Second, since the precision of the signals of the voters is symmetric across the states, Claim 1 implies that the vote shares in the two states α and β of the informed voters are symmetric to $\frac{1}{2}$. Third, suppose that $\lim_{n\to\infty} |\hat{q}(\omega;\sigma_n^*) - \frac{1}{2})|n^{-\frac{1}{2}} = \infty$. Then, Lemma 12 implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma_n^*) \in \{0,1\}$, given the symmetry of the vote share of the informed. This contradicts with Lemma 7. Finally, Observation 1 follows from Lemma 12.

E Extensions

E.1 Proof of Theorem 9

To be inserted.

References

Ahn, David S, & Oliveros, Santiago. 2012. Combinatorial voting. *Econometrica*, **80**(1), 89–141.

Alesina, Alberto, & Rodrik, Dani. 1994. Distributive politics and economic growth. *The quarterly journal of economics*, **109**(2), 465–490.

Ali, S Nageeb, Mihm, Maximilian, & Siga, Lucas. 2018. *Adverse Selection in Distributive Politics*. Tech. rept. Working paper, Penn State University.

Austen-Smith, David, & Banks, Jeffrey S. 1996. Information aggregation, rationality, and the Condorcet jury theorem. *American political science review*, **90**(1), 34–45.

Bergemann, Dirk, & Välimäki, Juuso. 2002. Information acquisition and efficient mechanism design. *Econometrica*, **70**(3), 1007–1033.

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

Billingsley, Patrick. 2008. *Probability and measure*. John Wiley & Sons.

Carmines, Edward G, & Stimson, James A. 1980. The two faces of issue voting. *American Political Science Review*, **74**(1), 78–91.

Converse, Philip E. 1964. The nature of belief systems in mass publics (1964). *Critical review*, **18**(1-3), 1–74.

Durrett, R. 1991. Probability: Theory and Examples, Wadsworth & Brooks/Cole, Pacific Grove. *MR1068527*.

Feddersen, Timothy, & Pesendorfer, Wolfgang. 1997. Voting behavior and information aggregation in elections with private information. *Econometrica: Journal of the Econometric Society*, 1029–1058.

Feddersen, Timothy, & Pesendorfer, Wolfgang. 1998. Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political science review*, **92**(1), 23–35.

Fernandez, Raquel, & Rodrik, Dani. 1991. Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, 1146–1155.

Gnedenko, Boris Vladimirovich. 1948. On a local limit theorem of the theory of probability. *Uspekhi Matematicheskikh Nauk*, **3**(3), 187–194.

Heese, Carl, & Lauermann, Stephan. 2017. *Persuasion and Information Aggregation in Elections*. Tech. rept. Working Paper.

Iyengar, Shanto, Hahn, Kyu S, Krosnick, Jon A, & Walker, John. 2008. Selective exposure to campaign communication: The role of anticipated agreement and issue public membership. *The Journal of Politics*, **70**(1), 186–200.

Krishna, Vijay, & Morgan, John. 2011. Overcoming ideological bias in elections. *Journal of Political Economy*, **119**(2), 183–211.

Krishna, Vijay, & Morgan, John. 2015. Majority rule and utilitarian welfare. *American Economic Journal: Microeconomics*, **7**(4), 339–375.

Lalley, Steven, & Weyl, E Glen. 2018. Nash equilibria for quadratic voting. *Available at SSRN 2488763*.

Martinelli, César. 2006. Would rational voters acquire costly information? *Journal of Economic Theory*, **129**(1), 225–251.

Mayer, Wolfgang. 1984. Endogenous tariff formation. *The American Economic Review*, **74**(5), 970–985.

McDonald, David R. 1980. On local limit theorem for integer-valued random variables. *Theory of Probability & Its Applications*, **24**(3), 613–619.

Miller, Joanne M, & Krosnick, Jon A. 2000. News media impact on the ingredients of presidential evaluations: Politically knowledgeable citizens are guided by a trusted source. *American Journal of Political Science*, 301–315.

Oliveros, Santiago. 2013. Aggregation of endogenous information in large elections.

Persson, Torsten, & Tabellini, Guido. 1994. Is inequality harmful for growth? *The American economic review*, 600–621.