Discussion Paper Series – CRC TR 224

Discussion Paper No. 209
Project B 03

Voter Attention and Distributive Politics

Carl Heese [1]

September 2020

[1] University of Bonn, Department of Economics, heese@uni-bonn.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Voter Attention and Distributive Politics *

Carl Heese †

August 31, 2020

Abstract

This paper studies theoretically how endogenous attention to politics affects social welfare and its distribution. When information of citizens about uncertain policy consequences is exogenous, a median voter theorem holds. When information is endogenous, attention shifts election outcomes into a direction that is welfare-improving. For a large class of settings, election outcomes maximize a weighted welfare rule. The implicit decision weight of voters with higher utilities is higher, but less so, when information is more cheap. In general, decision weights are proportional to how informed voters are. The results imply that uninformed voters have effectively almost no voting power, that the ability to access and interpret information is a critical determinant of democratic participation, and that elections are susceptible to third-party manipulation of voter information.

1 Introduction

This paper studies how endogenous attention to politics affects social welfare and its distribution. It is guided by the empirical observation that voters that care more about a political issue will acquire more information about it.[1]

I propose a model of an election over a distributive reform with uncertain consequences.Examples of such reforms are numerous: a trade reform opens

*I am grateful for helpful discussions with Nageeb Ali, Johannes Hörner, Daniel Krähmer, Stephan Lauermann, Philipp Strack, and Thomas Tröger as well as comments from audiences at Bonn, Yale, the CRC TR 224 Conferences 2018 and 2019, CED 2019, SAET 2019, Stony Brook 2019, CMID 20, the World Congress of the Econometric Society 2020, and the Young Economist Symposium 2020. The author gratefully acknowledges financial support from the German Research Foundation (DFG) through CRC TR224 (Project B03) and financial support from the European Research Council (ERC 638115).

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

[1] This is known as the "issue publics hypothesis" (Converse, 1964). See e.g. Krosnick (1990) and Henderson (2014), and Carpini & Keeter (1996) for an overview about the American public's factual knowledge about politics.

new markets for exporting firms but threatens the prospects in other sectors;a public health policy reform makes certain treatments more accessible to some citizens, while implying price increases for a range of pharmaceuticals needed by others; and a new education reform benefits some children but affects others negatively. In all these examples, some voters are *ex-ante* uninformed about the consequences of the reform, e. g. which sectors gain from a trade reform, or if their child benefits from education reform. However, they hold private information about their exposure to the proposed reform, that is, about the magnitude of their preference intensities: older people care more about healthcare issues, while changes in education policy are more relevant to citizens with children (Iyengar *et al.* , 2008).

Are referenda and elections efficient mechanisms of collective choice in such situations? This paper considers a modified version of the canonical setting by Feddersen & Pesendorfer (1997). Relative to Feddersen & Pesendorfer (1997), the voters' information about the policies is endogenous *and* the setting allows that the voters have conflicting interests (distributive politics). 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, others may prefer the reform only in the second state, while some "partisans" may prefer one of the policies independently of the state). The preferences are each voter's private information. Besides, all voters can receive information about the state in the form of a noisy signal, and each voter freely chooses the precision of her private signal. More precise information is more costly. Upon receiving their signals, all citizens vote simultaneously. The election determines the outcome by simple majority rule. Feddersen & Pesendorfer (1997) have shown that when voters receive conditionally i.i.d. signals of some exogenous quality and their preferences are "monotone" in all equilibria of large elections the outcome preferred by the median voter is elected state-by-state.[2] In many situations where voters have conflicting interests this is not the first-best outcome: for example, when 51% of the citizens marginally benefit from a reform, while the other 49% are severely impacted by it.

In our setting, elections either lead to the full-information outcome, but otherwise lead to outcomes that are only preferred by a minority ex-post

[2]The preference distribution of the voters is "monotone" if a higher belief in the first state entails that more voters prefer the reform.

(Theorem 2). This is the case when a minority of the voters is more severely affected by the reform. As a consequence, the minority will be better informed. Importantly, the more information the minority voters acquire, the more they correlate their vote with the unknown state of the world, thereby pushing the outcome in their favorite direction in each state. When voters of the minority group acquire substantially more information than others, they coordinate well on voting for their preferred policy, and this policy will indeed be elected in each state.

We provide the result that election outcomes are as if the decision weight of a citizen is proportional to how informed she is, provided the cost of voters to acquire political information are not "too high". This has important implications: first, when information cost are extremely low, all voters are relatively well informed, the implicit decision weights of citizens are approximately equal, and election outcomes almost always lead to majority-preferred outcomes. Second, uninformed citizens have no voting power, similar to voters that abstain due to voting being costly. Third, elections may be susceptible to targeted informational interventions of third-parties, which we will discuss in Section 6.

The main result characterizes which group of voters sharing a common interest will win the election. For this, we aggregate the decision weights of the citizens to describe the *power* of voter groups with common interests. A group's power increases in its size and the group's welfare at stake. The main result shows that the group with the larger power wins the election in each state. Under an independence condition,this yields sharp welfare predictions: elections lead to outcomes that maximize a weighted welfare rule. The weight of a voter is higher when her utilities are higher, but less so, when information is more cheap. For example, for intermediate cost, each citizen's information and weight turn out to be proportional to her utility. Then, elections lead to utilitarian outcomes.[3]

The main result describes the properties of limit equilibria with state-dependent election outcomes in large electorates. Thereby, we show, in particular, existence of such *informative* limit equilibria. This is economically surprising since voters of a large electorate face a severe free-rider problem

[3] I also show that aggregate cost of the voters converge to 0 as the electorate grows large such that the equilibrium sequences with utilitarian outcomes imply first-best results, even when taking into account the cost of voters, see Lemma 10.

when acquiring private information is costly, much similar to the reasoning in Downs' paradox of voting (Downs, 1957). The existence of informative limit equilibria relies on the observation that information acquisition in elections can be complementary, which we discuss in Section 4.5.4. This complementarity also drives an equilibrium multiplicity. Citizens may coordinate on paying much or very few attention to politics (Theorem 4).

In Section 6, we provide several extensions: first, we discuss the role of polarization of utilities within voter groups, and show that a more polarized group has a smaller electoral power and sufficiently much polarization, ceteris paribus, will lead the group to lose the election (Theorem 6). Second, we provide an extension where the cost of information of voters is heterogeneous, capturing that citizens have different abilities to access and interpret political information. Third, we discuss the potential of manipulative information provision by third-parties and its effectiveness.

In Section 7, we discuss the paper's contribution to the literature on voting cost and vote buying, especially Krishna & Morgan (2011) and Lalley & Weyl (2018). We also discuss the contribution to the literature on distributive politics, especially Fernandez & Rodrik (1991), and to the literature on information aggregation in elections: both modifications relative to Feddersen & Pesendorfer (1997) that are made in this paper have been studied before, but not together: Martinelli (2006) has studied a variant with endogenous information, and shown that the median voter theorem also holds, but only if voters can acquire relevant political information at a cost that is "not too high," thereby establishing the first existence result for informative equilibrium sequences.[4] Bhattacharya (2013a) has shown that the median voter result generalizes to settings with conflicting interests.[5] Importantly, his model does not allow to study the role of the intensity of preferences since the result is invariant to scaling the intensities of specific groups of voters. The paper also relates to a literature studying the interaction of limited attention of voters and the policy choices of political platforms. Matějka & Tabellini (2017)

[4]Formally, what matters for the result is how fast cost goes to zero when a voter chooses an arbitrarily uninformative signal. Basically, the critical condition is that elasticity of the cost function at the precision of the uninformative signal is large enough. The same condition is necessary in this paper for the existence of limit equilibria with non-trivial state-dependent outcomes, as the electorate grows large.

[5]Bhattacharya (2013a) also shows that the result breaks down when preferences are non-monotone; in particular, even minimal non-monotonicities turn around welfare predictions.

study this question in a probabilistic voting model.[6] There, citizens who pay attention are more responsive to policy changes, and as a consequence, political candidates offer policies catered to more attentive citizens. What differs is that in their work, endogenous attention distorts equilibrium policies away from first-best policies; in other words, the welfare implication is in the opposite direction relative to this paper. Second, the mechanism how attention affects policy outcomes is distinct from this paper, where information implicitly allows voter groups to coordinate more strongly, thereby enhancing their electoral power.

A Two-Type Example

The following extreme setting shows how a minority can overcome the dominance of a majority by correlating their vote more strongly with the state than the majority. Thereby, we illustrate how utilitarian outcomes can be elected, even when a majority of the voters do not prefer the outcome *ex-post*.[7] There are $2n+1$ voters. With probability $1 > \lambda > \frac{1}{2}$, 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 in β and B in α.[8] An aligned voter gets a small utility of $\epsilon > 0$ when her preferred policy is adopted, while a contrarian voter gets a utility of 1 when her preferred policy is adopted. Each voter can either 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., $\Pr(\alpha) = \frac{1}{2}$. Let ϵ is sufficiently small such that $\epsilon\lambda < (1-\lambda)$; hence, in order to maximize utilitarian welfare, the election should choose the contrarians' prefered policy.

Consider three scenarios: zero, intermediate and high cost. When cost is zero, i.e. $c = 0$, all voters become perfectly informed about the state and the outcome preferred by the median voter is elected in each state. When the cost is very high, e. g. $c > 1$, nobody gets informed, and the policy elected is independent of the state. Now, suppose that only the contrarians receive the

[6] See also Hu & Li (2018).

[7] We do not discuss the welfare effects of the cost for the example since it turns out that the aggregate cost is arbitrarily small in the equilibria of the main model when the electorate is large, $n \to \infty$ (see Section 6.4).

[8] The terminology used to label the voter types carries no economic meaning whatsoever but only relates to the notation. Aligned voters prefer the outcome that is "aligned" with the state.

perfect signal, and vote for their prefered outcome in each state; the aligned have no information about the state and vote for each policy with the same probability, i.e. $50-50$. Then, in each state, the outcome preferred by the contrarians is elected as the electorate grows large. We claim that this behaviour is an equilibrium for an intermediate range of cost c. The relevant observation is that the value of information is higher for the contrarians since $\epsilon < 1$. As a consequence, there are intermediate levels of cost that exceed the value of information for the aligned, but not for the contrarians.[9]

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 $t = (t_\alpha, t_\beta)$, where $t_\omega \in [-1,1]$ is the utility of A in ω. The utility of B is normalized to zero, so that t_ω is the difference between 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. A voter's type is her private information. Each voter privately observes a binary signal $s \in \{a, b\}$ about the state. 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 conditional on the state.

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 $x(t) \in [0, \frac{1}{2}]$ of her signal, that is $\frac{1}{2} + x(t) = \Pr(a|\alpha) = \Pr(b|\beta)$. 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.

There is a strictly increasing, strictly convex, and twice continuously differen-

[9] For completeness, note that without the private signal, a citizen is indifferent between voting for either of the policies. First, recall that the prior is $\Pr(\alpha) = \frac{1}{2}$. Second, the citizens do not infer anything about the state from conditioning on being pivotal for the election outcome. This is because the event in which the citizen's vote affects the outcome is equally likely in each state in the candidate equilibrium since in β the reform wins with the same margin of $\left[\lambda\frac{1}{2} + (1-\lambda)\right] - \frac{1}{2} = \frac{1}{2}(1-\lambda)$ and in α the reform loses with a margin $\frac{1}{2}(1-\lambda)$ in expectation.

tiable 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 $d > 1$ such that[10]

$$\lim_{x \to \infty} \frac{c'(x)}{x^{d-1}} \in \mathbb{R}. \tag{1}$$

A strategy $\sigma = (x, \mu)$ of a voter consists of a function $x : [-1, 1]^2 \to [0, \frac{1}{2}]$ mapping types to signal precisions and of a function $\mu : [-1, 1]^2 \times \{a, b\} \to [0, 1]$ mapping types and signals to probabilities to vote A, i.e., $\mu(t, s)$ is the probability that a voter of type t with signal s votes for A. I consider only non-degenerate strategies.[11] I analyze the Bayes-Nash equilibria of the Bayesian game of voters in symmetric strategies, henceforth called *equilibria*.

2.1 Preferences

Figure 1 shows the area of possible preference types. 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*). To simplify the exposition, in the rest of the paper, we only consider strategies σ where the partisans use the (weakly) dominant strategy to vote for their preferred policy.[12]

Aggregate Preferences. A central object of the analysis is the *aggregate preference function*

$$\Phi(p) = \mathrm{Pr}_H(\{t : p \cdot t_\alpha + (1 - p) \cdot t_\beta \geq 0\}), \tag{2}$$

which maps a belief $p \in [0, 1]$ about the state being α to the probability that a

[10]It will be a direct 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 (18).

[11]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 voter deviates to any other strategy, then the outcome is the same. Therefore, the degenerate strategies with $x = 0$ are trivial equilibria.

[12]In fact, for any non-degenerate strategy, I show that the likelihood that a given voter is pivotal for the election outcome is non-zero (see Section 3.1) such that voting for the preferred policy while not acquiring any information is the unique strict best response for all partisans.

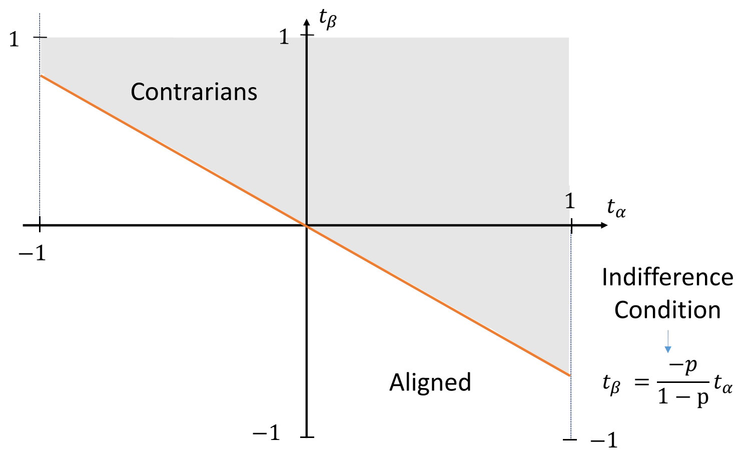

Figure 1: For any given belief $p = \Pr(\alpha) \in (0,1)$, the set of types t with a threshold of doubt $y(t) = p$ is given by $t_\beta = \frac{-p}{1-p} t_\alpha$. Voter types north-east of the indifference line (shaded area) prefer A given p.

random type t prefers A given p. Figure 1 illustrates Φ: the (colored) line corresponds to the set 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 (shaded area); these types have mass $\Phi(p)$. The indifference set has a slope of $\frac{-p}{1-p}$ and an increase in p corresponds to a clockwise rotation of it. Given that H has a continuous density, Φ is continuously differentiable in p.

I assume that

$$\Phi(0) < \frac{1}{2}, \text{ and } \Phi(1) > \frac{1}{2} \tag{3}$$

such that the median-voter preferred outcome is A in α and B in β. In particular, this excludes the cases when there is a majority of partisans for one policy in expectation. I also make the genericity assumption that Φ is not constant on any open interval.[13] Henceforth, I will call distributions H that have a continuous density and satisfy (3) simply *preference distributions*. The set of the aligned types is denoted $L = \{t : t_\alpha > 0, t_\beta < 0\}$ and the set of the contrarian types is denoted $C = \{t : t_\alpha < 0, t_\beta > 0\}$ and $g \in \{L, C\}$ is the generic symbol for a voter group, aligned or contrarians.

[13]This assumption is known from the literature, see Bhattacharya (2013b).

Threshold of Doubt and Preference Intensity It is useful to view types as information about, first, the relative preference intensities across states,

$$y(t) = \frac{-t_\beta}{t_\alpha - t_\beta}, \tag{4}$$

and, second, the *total intensity*,

$$k(t) = t_\alpha - t_\beta. \tag{5}$$

For any aligned type t, $y(t)$ and $k(t)$ together uniquely pin down t.[14] Similarly, for any contrarian type t, $y(t)$ and $k(t)$ together uniquely pin down t. Recall that a strategy describes a voting choice and an information choice for each type. Section 3 shows that the *threshold of doubt* $y(t)$ determines the voting choice of (non-partisan) types, and the total intensity determines the information choice.

3 Citizens' Votes and Information

3.1 Threshold of Doubt Pins Down Vote

Take any strategy $\sigma = (x, \mu)$ of the voters. The probability that a voter of random type votes for A in state $\omega \in \{\alpha, \beta\}$ is denoted $q(\omega; \sigma)$. A simple calculation shows that

$$q\left(\alpha;\sigma\right) \;=\; \int_{t\in[-1,1]^2} (\frac{1}{2} + x(t))\mu(t,a) + (\frac{1}{2} - x(t))\mu(t,b)dHt, \tag{6}$$

and

$$q\left(\beta;\sigma\right) \;=\; \int_{t\in[-1,1]^2} (\frac{1}{2} - x(t))\mu(t,a) + (\frac{1}{2} + x(t))\mu(t,b)dHt. \tag{7}$$

I also refer to $q\left(\omega;\sigma\right)$ as the *(expected) vote share* of A in ω.

Pivotal Voting. Take a single citizen, and fix a strategy σ' of the other voters. The given citizen's vote determines the outcome only in the event when the votes of the other citizens tie, denoted piv. Thus, a strategy is

[14] For $t \in L$, $y(t)k(t) = t_\beta$, and $(1 - y(t))k(t) = t_\alpha$.

optimal if and only if it is optimal conditional on the pivotal event piv. The probability that the votes of the other citizens tie in ω is

$$\Pr\left(\text{piv}|\omega;\sigma',n\right)=\binom{2n}{n}\left(q\left(\omega;\sigma'\right)\right)^{n}\left(1-q\left(\omega';\sigma\right)\right)^{n}. \tag{8}$$

since conditional on the state, the type and the signal of a voter is independent of the types and the signals of the other voters. For any type t of the given citizen, and given the precision choice $x(t)$, let $\Pr(\alpha|s,\text{piv};\sigma',n)$ be the posterior probability of α conditional on having received the private signal s and conditional on *being pivotal* when the other voters use σ'. We conclude that, μ is part of a best response $\sigma=(x,\mu)$ if and only if for all $t=(t_\alpha,t_\beta)$ and for the signal precision $x(t)$,

$$\Pr(\alpha|s,\text{piv};\sigma',n)\cdot t_\alpha+(1-\Pr(\alpha|s,\text{piv};\sigma',n))\cdot t_\beta>0\Rightarrow\mu\left(s,t\right)=1, \tag{9}$$
$$\Pr(\alpha|s,\text{piv};\sigma',n)\cdot t_\alpha+(1-\Pr(\alpha|s,\text{piv};\sigma',n))\cdot t_\beta<0\Rightarrow\mu\left(s,t\right)=0, \tag{10}$$

that is, a voter supports A if the expected value of A conditional on being pivotal and s is strictly positive and otherwise supports B. Note that for each aligned type $t\in L$, (9) and (10) are equivalent to

$$\Pr(\alpha|s,\text{piv};\sigma',n)>y(t)\Rightarrow\mu(t,s)=1, \tag{11}$$
$$\Pr(\alpha|s,\text{piv};\sigma',n)<y(t)\Rightarrow\mu(t,s)=0; \tag{12}$$

and for all contrarian types $t\in C$, (9) and (10) are equivalent to

$$\Pr(\alpha|s,\text{piv};\sigma',\sigma,n)>y(t)\Rightarrow\mu(t,s)=0, \tag{13}$$
$$\Pr(\alpha|s,\text{piv};\sigma',\sigma,n)<y(t)\Rightarrow\mu(t,s)=1, \tag{14}$$

We see that $y(t)$ is the unique belief that a makes a voter of type t indifferent, thereby qualifying the name *threshold of doubt*.

3.2 Preference Intensity Pins Down Information Level

What is the marginal value of information to a citizen? Take an aligned voter, and fix the likelihood $x(t)>0$ of her receiving a correct signal about the state. At the end of this section, we establish that she votes A after a and B after b

(Lemma 1), that is, she votes for her preferred policy in each state whenever receiving a "correct signal" . When she is not pivotal, the policy elected is independent of her private precision $x(t)$. In the pivotal event, using Lemma 1, her expected utility from the elected policy is

$$\Pr(\text{piv}|\sigma', n)\Pr(\alpha|\text{piv};\sigma)(\frac{1}{2}+x(t))t_\alpha \tag{15}$$

in state α, and

$$\Pr(\text{piv}|\sigma', n)\Pr(\beta|\text{piv};\sigma)(\frac{1}{2}-x(t))t_\beta \tag{16}$$

in state β, where we used that the utility from B is normalized to zero.[15] Therefore, the marginal benefit of a higher precision $x(t)$ is

$$\begin{aligned} & MB(x(t);\sigma', n) \\ = \; & \Pr(\text{piv}|\sigma', n)(\Pr(\alpha|\text{piv};\sigma)t_\alpha - \Pr(\beta|\text{piv};\sigma)t_\beta) \\ = \; & \Pr(\text{piv}|\sigma', n)k(t)c_1(y(t)) \end{aligned} \tag{17}$$

for $c_1(y(t)) = \Pr(\alpha|\text{piv};\sigma)(1-y(t)) + \Pr(\beta|\text{piv};\sigma)y(t)$, where we used that $t_\alpha = k(t)(1-y(t))$ and $t_\beta = k(t)y(t)$ for the last equation. We see that the total intensity $k(t)$ is decisive. Finally, for any type t for which it is optimal to acquire some information, $x(t) > 0$, the precision is pinned down by equating marginal benefits and marginal cost,

$$c'(x(t)) = MB(x(t);\sigma', n) \tag{18}$$

It follows from the strict convexity of c, that for any t, there is a unique solution to (18), denoted $x^*(t;\sigma', n)$. Moreover, $x^*(t;\sigma', n)$ is continuously differentiable by an application of the implicit function theorem, and, given

[15] Similarly, in the pivotal event, a contrarian's expected utility when choosing $x(t)$ is

$$\Pr(\text{piv};\sigma', n)\Pr(\alpha|\text{piv};\sigma)(\frac{1}{2}-x(t))t_\alpha$$

in state α, and

$$\Pr(\text{piv};\sigma', n)\Pr(\beta|\text{piv};\sigma)(\frac{1}{2}+x(t))t_\beta$$

in state β.

(1),

$$x^*(t;\sigma,n) \approx MB(x(t);\sigma',n)^{\frac{1}{d-1}}. \tag{19}$$

Lemma 1 *Take any strategy σ'. The function μ is part of a best response $\sigma = (x,\mu)$ if and only if*

$$\forall t \in L : x(t) > 0 \Rightarrow \mu(t,a) = 1 \;\; \textit{and} \;\; \mu(t,b) = 0, \tag{20}$$
$$\forall t \in C : x(t) > 0 \Rightarrow \mu(t,a) = 0 \;\; \textit{and} \;\; \mu(t,b) = 1. \tag{21}$$

The proof is in the Appendix.

3.3 Information Acquisition Region

The *critical types* t with $y(t) = \Pr(\alpha|\text{piv};\sigma',n)$ are indifferent between A and B without further information, given (11) - (14). Lemma 2 shows that, for each total intensity $k = k(t) \in [0,2]$, only types in a certain interval around the critical types acquire information.

Lemma 2 *Let σ' be a strategy with $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma',n) \in (0,1)$. When n is large enough, for any $k \in (0,2)$ and any $g \in \{L,C\}$ there are $\phi_g^-(k) < \Pr(\alpha|\text{piv};\sigma',n) < \phi_g^+(k)$ for such that for any best response $\sigma = (x,\mu)$ to σ' and any type $t \in g$ with $k(t) = k$,*

$$x(t) > 0 \quad \Leftrightarrow \quad y(t) \in [\phi_g^-(k), \phi_g^+(k)], \tag{22}$$

The proof is in the Appendix. Figure 2 illustrates the functions ϕ_g^- and ϕ_g^+. For intuition: one can show that $y(t) \geq \phi_g^-(k)$ if and only if

$$\frac{\Pr(\alpha|\text{piv})}{\Pr(\beta|\text{piv})} \frac{\frac{1}{2} - x^{**}(t)}{\frac{1}{2} + x^{**}(t)} \leq \frac{y(t)}{1 - y(t)}, \tag{23}$$

and $y(t) \leq \phi_g^+(k)$ if and only if

$$\frac{y(t)}{1 - y(t)} \leq \frac{\Pr(\alpha|\text{piv})}{\Pr(\beta|\text{piv})} \frac{\frac{1}{2} + x^{**}(t)}{\frac{1}{2} - x^{**}(t)}, \tag{24}$$

for $x^{**}(t;\sigma,n) = x^*(t;\sigma,n)(1 - \frac{c(x^*(y,k;\sigma,n))}{x^*(t;\sigma,n)c'(x^*(t;\sigma,n))})$, where $x^*(t;\sigma,n)$ is the solution to the first-order condition (18). Thus, to decide if to acquire any

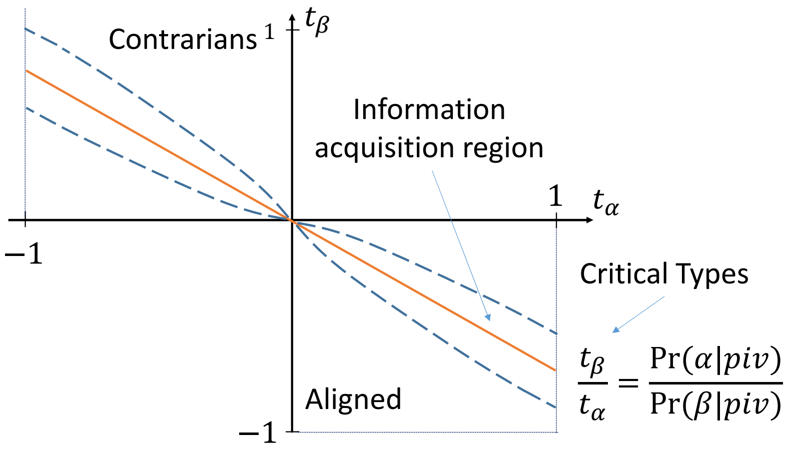

Figure 2: Information Acquisition regions of group g with boundaries the graphs of $\phi_g^-(k)$ and $\phi_g^+(k)$ (dashed lines).

information, a voter discounts the precision $x^*(t;\sigma,n)$ of her optimal informative signal by a certain cost factor, and then considers if, given the discounted precision, one of the signals, a or b, sways her opinion on which policy to vote for or if none of the signals sways her.

For the later equilibrium analysis, it is key to observe that the region of types acquiring information vanishes as $n \to \infty$. This observation will allow us to understand the aggregate information acquisition of large electorates by local approximations.

Lemma 3 *Take any σ'. Take the best response $\sigma = (x, \mu)$. Then, for any $k \in [0, 2]$,*

$$\lim_{n\to\infty} \phi_g^+(k) = \lim_{n\to\infty} \Pr(\alpha|\text{piv};\sigma',n) = \lim_{n\to\infty} \phi_g^-(k). \tag{25}$$

We claim that $\lim_{n\to\infty} \Pr(\text{piv}|\sigma,n) = 0$, that is, the pivotal likelihood goes to zero as $n \to \infty$. In fact, a Stirling approximation of the binomial coefficient and (8) yields[16] [17]

$$\Pr(\text{piv}|\omega;\sigma,n) \approx 4^n (n\pi)^{-\frac{1}{2}} \Big[q(\omega;\sigma)(1-q(\omega;\sigma)\Big]^n, \tag{26}$$

[16] 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: $\lim_{n\to\infty} \frac{x_n}{y_n} = 1$.

[17] 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, $\binom{2n}{n} \approx (2\pi)^{-\frac{1}{2}} 2^{2n+\frac{1}{2}} n^{-\frac{1}{2}} = 4^n (n\pi)^{-\frac{1}{2}}$.

and $\lim_{n\to\infty} \Pr(\text{piv}|\sigma, n) = 0$ follows from (26) since $q(1-q)$ is bounded above by $\frac{1}{4}$ on $[0,1]$. Importantly, this implies

$$x^*(t;\sigma,n) \to 0, \tag{27}$$

given (19) and (17). Hence, $x^{**}(t;\sigma,n) \to 0$. This, together with (23) and (24) implies (25).

4 Informative Equilibrium Sequences

In the following, for the ease of the exposition, we take Φ to be strictly monotone. The results in the general case do not differ qualitatively, and are provided in Section 6. We consider a sequence of elections along which the electorate's size $2n+1$ grows. For each n and a strategy σ_n, we calculate the probability that a policy $z \in \{A, B\}$ wins the support of the majority of the voters in state ω, denoted $\Pr(z|\omega;\sigma_n,n)$. We are interested in the limits of $\Pr(z|\omega;\sigma_n^*,n)$ for equilibrium sequences $(\sigma_n^*)_{n\in\mathbb{N}}$. We are particularly interested in equilibrium sequences where citizens vote in an informed manner such that the election outcomes differ across the states,

$$\lim_{n\to\infty} \Pr(A|\alpha;\sigma_n,n) \neq \lim_{n\to\infty} \Pr(A|\beta;\sigma_n,n), \tag{28}$$

which we call *informative.*[18]

4.1 Information Weighted Majority

What will matter in informative equilibria is if the aligned voters or the contrarian voters acquire more information, that is if

$$\int_{t\in L} x(t)dH(t) > \int_{t\in C} x(t)dH(t). \tag{29}$$

The precision $x(t)$ of a voter will play the role of an implicit *decision weight* of each voter. We will show that, in large electorates, in all states, the policy preferred by the aligned is elected when the sum of their decision weights is

[18] For large classes of settings, any typical efficiency measure, for example, full-information equivalence or utilitarian efficiency, requires equilibrium sequences to be informative.

larger than that of the contrarians, and vice versa. A heuristic explanation is this: when all citizens acquire some information, $x(t) > 0$,

$$q(\alpha; \sigma_n^*) = \Big[\int_{t \in L} \frac{1}{2} + x(t) dH(t) + \int_{t \in C} \frac{1}{2} - x(t) dH(t) \Big], \tag{30}$$

$$q(\beta; \sigma_n^*) = \Big[\int_{t \in L} \frac{1}{2} - x(t) dH(t) + \int_{t \in C} \frac{1}{2} + x(t) dH(t) \Big], \tag{31}$$

given Lemma 1. Hence,

$$\begin{aligned} & q(\alpha; \sigma_n^*) > \frac{1}{2} > q(\beta; \sigma_n^*) \\ \Leftrightarrow \quad & \int_{t \in L} x(t) dH(t) > \int_{t \in C} x(t) dH(t). \end{aligned} \tag{32}$$

Thus, whenever (29) holds, a majority of citizens votes for A in α and B in β, that is for the outcomes prefered by the aligned. What this heuristic does not capture though is that the uninformed types, $x(t) = 0$, may play a role in the election unless they randomize their vote $50 - 50$.

In Section 4.2, we describe $\int_{t \in g} x(t) dH(t)$ in terms of the primitives of the model, thereby uncovering how the properties of a voter group g determine the endogenous information and the electoral power of the group. In Section 4.3, we state and prove the main result, characterizing all informative equilibrium sequences, thereby showing the somewhat surprising implication that the uninformed types (mis)coordinate on voting $50 - 50$ in the aggregate.

4.2 Information and Power of Voter Groups

The following result shows that, when n is large, the information $\int_{t \in g} x(t) dH(t)$ acquired by a voter group is proportional to the mass of the critical types in the voter group and proportional to a weighted mean of the intensities of these critical types. The weight of the intensities depends on the limit elasticity of the cost function, $d = \lim_{x \to 0} \frac{c'(x)x}{c(x)}$, which can be interpreted as a measure of how "cheap" information of low precision is.[19] The proof uses Lemma 3 and is provided in Section 4.2.1 and Section 4.2.2.

[19]For illustration, take e.g. $c_d(x) = x^d$. Then $\lim_{x \to 0} \frac{c_d(x)}{c_{d'}(x)} = \infty$ if $d' > d$.

Lemma 4 *Let $g \in \{L, C\}$. Take any strategy σ'. Let $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma', n) \in (0,1)$, and*

$$W(g, \hat{p}) = \underbrace{\Pr(\{t : t \in g\}) \Pr(\{t : y(t) = \hat{p}\}|t \in g)}_{\text{likelihood of critical types}} \underbrace{\mathrm{E}(k(t)^{\frac{2}{d-1}}|y(t) = \hat{p}, t \in g)}_{\text{weighted mean intensity of critical types}}.$$

For the best response $\sigma = (x, \mu)$ to σ',

$$\int_{t\in g} x(t) dH(t) \approx W(g, \hat{p}) \Pr(\text{piv}|\sigma', n)^{\frac{2}{d-1}} c_2, \tag{33}$$

where $c_2 > 0$ is a constant independent of g.

Note that, in the following, we sometimes denote types by (y, k) instead of t.

4.2.1 How Many Voters Acquire Information and How Much

Fix $k = k(t)$. Given Lemma 3, when n is large, only types close to critical type with $y(t) = \Pr(\alpha|\text{piv}; \sigma, n)$ acquire information, $x(t) > 0$. We show that, as a consequence, all such types choose asymptotically equivalent precisions as $n \to \infty$. In the following, we sometimes drop σ and n to shorten notation.

Claim 1 *Take any strategy σ'. Take the sequence of best responses. Let $k \in [0, 2]$. Take any converging sequence $(y_n)_{n\in\mathbb{N}}$. If $x(y_n, k) > 0$ for all n,*

$$\frac{x_n(y_n, k)}{x_n(\Pr(\alpha|\text{piv}), k)} \approx 1. \tag{34}$$

Proof. Differentiating the first-order condition (18) implicitly, we show

$$\lim_{n\to\infty} \frac{\partial x^*(y, k; \sigma_n, n)}{\partial y} = 0. \tag{35}$$

in the Appendix. Together with Lemma 3, (35) implies (34). ■

We show that the interval of types acquiring information, $x(t) > 0$, is asymptotically symmetric around the critical type with $y(t) = \Pr(\alpha|\text{piv})$.

Claim 2 *Take any sequence σ'_n. Take the sequence of best responses $\sigma_n = (x_n, \mu_n)$. Then, for any $k \in (0, 2)$,*

$$\frac{x_n^{**}(\Pr(\alpha|\text{piv}), k)}{\phi_g^+(k) - \Pr(\alpha|\text{piv})} \approx \frac{x_n^{**}(\Pr(\alpha|\text{piv}), k)}{\Pr(\alpha|\text{piv}) - \phi_g^-(k)} \approx c_3, \tag{36}$$

for $x_n^{**}(y,k) = x_n(y,k)(1 - \frac{c(x_n(y,k))}{c'(x_n(y,k))x_n(y,k)})$*, and where* c_3 *is a constant that only depends on* $\lim_{n\to\infty} \Pr(\alpha|\text{piv})$*.*

Proof. The proof of Lemma 2 provides also an equivalent description of the boundary conditions (23) and (24): the information acquisition interval $[\phi_g^-(k), \phi_g^+(k)]$ is implicitly given by

$$\frac{1}{2} - x^{**}(\phi_g^-(k), k) = \chi(\phi_g^-(k)), \tag{37}$$

$$\frac{1}{2} + x^{**}(\phi_g^+(k), k) = \chi(\phi_g^+(k)), \tag{38}$$

for $\chi(y) = \frac{\Pr(\beta|\text{piv})y}{\Pr(\alpha|\text{piv})(1-y)+\Pr(\beta|\text{piv})y}$.[20] Since $\phi_g^-(k) \to \Pr(\alpha|\text{piv})$ and $\phi_g^+(k) \to \Pr(\alpha|\text{piv})$ (see Lemma 3) and since $\chi(\Pr(\alpha|\text{piv})) = \frac{1}{2}$, Taylor approximations of $\chi(\phi_g^-(k))$ and $\chi(\phi_g^-(k))$ give

$$\chi'(\Pr(\alpha|\text{piv}))\Big[\phi_g^+(k) - \Pr(\alpha|\text{piv})\Big] \approx x^{**}(\phi_g^+(k)), \tag{39}$$

$$\chi'(\Pr(\alpha|\text{piv}))\Big[\Pr(\alpha|\text{piv}) - \phi_g^-(k)\Big] \approx x^{**}(\phi_g^-(k)). \tag{40}$$

Finally, (36) follows from (39), (40), (34), and the continuity of c. ■

4.2.2 Aggregate Information of a Voter Group

Denote by f the density of the cumulative distribution function of the threshold of doubt $y(t)$. Now, we finish the proof of Lemma 4. For this, we show that, fixing the total intensity $k = k(t)$, the average precision of citizen types is proportional to the likelihood of the critical type and the weighted intensity $k(t)^{\frac{2}{d-1}}$,

$$\begin{aligned} &\mathrm{E}(x(t)|k(t) = k, t \in g) \\ \approx\ &\underbrace{f(\Pr(\alpha|\text{piv})|k(t) = k, t \in g)}_{\text{likelihood of critical type}} \quad \underbrace{k^{\frac{2}{d-1}}}_{\textit{weighted intensity}} \quad \Pr(\text{piv})^{\frac{2}{d-1}} c_1. \end{aligned} \tag{41}$$

for a constant $c_1 > 0$ that only depends on $\Pr(\alpha|\text{piv})$. Then, we aggregate over k to obtain (36). Details for this aggregation are in the Appendix.

[20] To see how (37) and (23) relate, rewrite (23), $\frac{\frac{1}{2}-x^{**}(t)}{\frac{1}{2}+x^{**}(t)} \leq \frac{\Pr(\beta|\text{piv})}{\Pr(\alpha|\text{piv})} \frac{y(t)}{1-y(t)}$, and rewrite further, $\frac{1}{2} - x^{**}(t) \leq \chi(y)$. Similarly, for (38) and (24).

First, given Lemma 3, Taylor approximations of the c.d.f yield

$$\frac{f(\Pr(\alpha|\text{piv}) \mid k(t)=k, t\in g)\left[\phi_g^-(k)-\phi_g^+(k)\right]}{\Pr(\{t : \phi_g^-(k) \le y(t) \le \phi_g^+(k)\} \mid k(t)=k, t\in g)} \approx 1. \tag{42}$$

Combining (34), (36), and (42), for any k,

$$\begin{aligned} & \mathrm{E}(x(t)|k(t)=k, t\in g) \\ \approx\ & x_n(\Pr(\alpha|\text{piv}), k)x_n^{**}(\Pr(\alpha|\text{piv}), k)f(\Pr(\alpha|\text{piv})|k(t)=k, t\in g)c_4 \\ \approx\ & x_n(\Pr(\alpha|\text{piv}), k)^2 f(\Pr(\alpha|\text{piv})|k(t)=k, t\in g)c_5. \end{aligned} \tag{43}$$

for constants $c_4 \neq 0$ and $c_5 \neq 0$ and where, for the last line, we used that $x_n^{**}(\Pr(\alpha|\text{piv}), k) \approx \frac{d-1}{d}x_n(\Pr(\alpha|\text{piv}), k)$ since $\frac{1}{d} = \lim_{x\to 0}\frac{c(x)}{c'(x)x}$. We see that what matters are the likelihood and the precision of the critical type. The precision of the critical type scales with the total intensity,

$$x(\Pr(\alpha|\text{piv}), k)^2 \approx k(t)^{\frac{2}{d-1}}\left[\Pr(\text{piv})c_1(\Pr(\alpha|\text{piv}))\right]^{\frac{2}{d-1}}, \tag{44}$$

given (18), so that (44) and (43) imply (41).

4.2.3 Power of a Voter Group

We call

$$W(g) = W(g, \hat{p}) \tag{45}$$

the *power* of a voter group, where $\hat{p}$ is the unique belief $\hat{p}$ for which the electorates preferences are split, $\Phi(\hat{p}) = \frac{1}{2}$. The next lemma shows that for any informative equilibrium sequence, the threshold of doubt of the critical types converges to $\hat{p}$, so that, given Lemma 4, $W(g)$ measures the amount of information acquired by the group in any such equilibrium sequence.

Lemma 5 *Let $\Phi(\hat{p}) \neq \frac{1}{2}$. Then, for any informative equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$,*

$$\lim_{n\to\infty}\Pr(\alpha|\text{piv}; \sigma_n^*, n) = \hat{p}. \tag{46}$$

The proof is provided in the Appendix. There, we show that when (46) does not hold, the vote shares $q(\omega;\sigma_n^*)$ do not converge to $\frac{1}{2}$, and as a consequence, the citizens choose exponentially low levels of precision. This, in turn, implies that the difference in the vote shares in α and β is exponentially small. Finally, we show that this implies that the distribution of the election outcome is asymptotically the same in both states as $n\to\infty$, which cannot be true in any informative equilibrium sequence.

4.3 Result

The main result shows that for all informative equilibrium sequences, the outcome preferred by the group with the larger power is elected as $n\to\infty$. Moreover, there exists an informative equilibrium sequence when information of low precision $x\approx 0$ is sufficiently cheap; this will be captured by a condition on the elasticity at zero, $\lim_{x\to 0}\frac{c'(x)x}{c(x)}$.[21] We call $W(L)\neq W(C)$, $W(L)\neq$, and $W(C)\neq 0$ the *genericity conditions.*

Theorem 1 *Let* $\lim_{x\to 0}\frac{c'(x)x}{c(x)}>3$. *Take any preference distribution* H *satisfying the genericity conditions and* $\Phi(\Pr(\alpha))\neq\frac{1}{2}$.

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

$$\begin{aligned}\lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*,n) &= \lim_{n\to\infty}\Pr(B|\beta;\sigma_n^*,n)\\ &= \begin{cases}0 & \text{if } W(L)<W(C),\\ 1 & \text{if } W(L)>W(C).\end{cases}\end{aligned}\tag{47}$$

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

4.4 Proof: Power Rule

This section proves the first item of Theorem 1, showing that the order of the power of aligned and contrarians determines election outcomes. First of all, $q(\alpha;\sigma_n^*)>q(\beta;\sigma_n^*)\Leftrightarrow\frac{\int_{t\in L}x(t)dH(t)}{\int_{t\in C}x(t)dH(t)}>1$,[22] so that, given Lemma 4, the order

[21]The same condition appears in Martinelli (2006)'s model as a sufficient condition for informative and determinate equilibrium outcomes.

[22]First, partisans vote the same in both states. Second, for aligned types the likelihood to vote A in α differs by $2x(t)$ from the likelihood to vote A in β. Third, for contrarian types the likelihood to vote A in α differs by $-2x(t)$ from the likelihood to vote A in β. Together, $q(\alpha;\sigma_n^*)-q(\beta;\sigma_n^*)=\int_{t\in L}2x(t)dH(t)-\int_{t\in C}2x(t)dH(t)$, which implies the equivalence stated.

of $W(g)$ pins down the order of the vote shares: for n large enough,

$$q(\alpha;\sigma_n^*) > q(\beta;\sigma_n^*) \quad\Leftrightarrow\quad \frac{W(L)}{W(C)} > 1, \tag{48}$$

The key step is to establish that when the elasticity of the cost function at zero is sufficiently large, $\lim_{x\to 0}\frac{xc'(x)}{c(x)} = d > 3$, then, for any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ outcomes are determinate as $n\to\infty$,

$$\lim_{n\to\infty} \Pr(A|\omega;\sigma_n^*, n) \in \{0,1\}. \tag{49}$$

in each state ω. For informative equilibrium sequences, this implies that A is elected in one state, and B in the other. When the vote share for policy A is higher in α than in β, A is elected in α, and B in β and vice versa when the vote share for policy A is higher in β than in α. We conclude that (48) and (49) together imply (47). The following section proves (49).

4.4.1 Determinate Outcomes

We show that, given $d > 3$, for any sequence of equilibria, the outcomes are determinate, as $n\to\infty$, that is, we prove (49).

For this, for any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and any n, let $\mathbf{q}(\sigma_n) = (q(\alpha;\sigma_n), q(\beta;\sigma_n))$, and denote by $s(\omega;\mathbf{q}(\sigma_n)) = \Big[q(\omega;\sigma_n)(1-q(\omega;\sigma_n)(2n+1)\Big]^{\frac{1}{2}}$ the standard deviation of the vote share in ω. Let

$$\delta(\omega) = \lim_{n\to\infty}\frac{2n+1}{s(\omega;\mathbf{q}(\sigma_n))}\left[q(\omega;\sigma_n) - \frac{1}{2}\right] \tag{50}$$

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

The proof of (49) proceeds in three steps. The first step shows that, as a consequence of the central limit theorem, as $n\to\infty$, the asymptotic distribution of the outcome policies only depends on the distance of the vote share to the majority threshold in terms of standard deviations, i.e. $\delta(\omega)$. The proof of this step is in the Appendix.

Observation 1 *Take 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*

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

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

What determines the equilibrium distance of the vote shares to each other, and thereby their distance to the majority treshold, is how much information the voters acquire in equilibrium,

$$q(\alpha; \sigma_n) - q(\beta; \sigma_n) = 2\Big[\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t)\Big]. \tag{51}$$

For the second step, suppose that the election is not determinate in a state ω, e.g. in α. Given Observation 1, $\delta(\alpha) \in \mathbb{R}$. We show that $q(\alpha; \sigma_n) - q(\beta; \sigma_n)$ is of of an order larger than inverse of the standard deviation of the vote share,

$$\lim_{n\to\infty} \Big[q(\alpha; \sigma_n) - q(\beta; \sigma_n)\Big] s(\omega; \mathbf{q}(\sigma_n)) \in \{\infty, -\infty\}. \tag{52}$$

if $d > 3$. To prove (52), first, we show that

$$\lim_{n\to\infty} \Big[q(\alpha; \sigma_n) - q(\beta; \sigma_n)\Big] \Pr(\text{piv}|\sigma_n, n)^{-1} = \infty. \tag{53}$$

if $d > 3$. To see why, note that

$$\begin{aligned} &\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t) \\ \approx\ &\Big[W(L) - W(C)\Big] \Pr(\text{piv}|\sigma_n^*, n)^{\frac{2}{d-1}} c_2, \end{aligned} \tag{54}$$

given Lemma 4. Using that the pivotal likelihood goes to zero as $n \to \infty$, (53) follows from (51), (54), and the genericity conditions. Second, using the local central limit theorem, we show that, for all strategies with vote shares close to the majority threshold as in the lemma, the pivotal likelihood is inversely proportional to the the standard deviation of the vote share.

Observation 2 *For any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$. If $\lim_{n\to\infty} q(\omega; \sigma_n) \in (0, 1)$, then*

$$\lim_{n\to\infty} \Pr(\text{piv}|\omega; \sigma_n) s(\omega; \mathbf{q}(\sigma_n)) \quad = \quad \phi(\delta(\omega)), \tag{55}$$

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

The proof is an application of the local central limit theorem, and provided in the Appendix.[23] Observation 2 and (53) together yield (52).

Finally, we prove (49). Note that we can write $\delta(\omega) = \lim_{n\to\infty} s(\omega; \mathbf{q}(\sigma_n))\Big[q(\omega;\sigma_n) - \frac{1}{2}\Big]$. Hence, (52) implies $\delta(\alpha) - \delta(\beta) \in \{\infty, -\infty\}$. Since $\delta(\alpha) \in \mathbb{R}$, we have $\delta(\beta) \in \{\infty, -\infty\}$. Then, Lemma 2 implies that the inference from the pivotal event is not bounded, and $\lim_{n\to\infty} \Pr(\alpha|\text{piv}, s; \sigma_n^*) = 0$ for $s \in \{a, b\}$. Hence, as $n \to \infty$, citizens vote as if they know that β holds, so that $\lim_{n\to\infty} q(\alpha; \sigma_n^*) = \Phi(0) < \frac{1}{2}$, which contradicts $\delta(\alpha) \in \mathbb{R}$. The assumption that the election outcome is determinate in β similarly leads to a contradiction.

4.5 Proof: Existence

This section proves existence of an informative equilibrium sequence when $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$. For this, first, we provide a convenient equilibrium representation.

4.5.1 Equilibrium Representation through 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 (11)-(14), with $\sigma' = \sigma$, (18) for all types t with $x(t) > 0$, and (22).
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.,

$$\mathbf{q}(\sigma) = (q(\alpha;\sigma), q(\beta;\sigma)). \tag{56}$$

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 (8). Given (11)-(14), (18), and (37)-(38), 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 $\mathbf{q}$. Then, σ^* is an equilibrium, if and only if, $\sigma^* = \sigma^{\mathbf{q}(\sigma^*)}$. Conversely, an

[23]See Gnedenko (1948), and Davis & McDonald (1995) for the local limit theorem for triangular arrays of integer-valued variables.

equilibrium can be described by a vector of vote shares $\mathbf{q}^*$ that is a fixed point of $\mathbf{q}(\sigma^-)$, i.e.,[24]

$$\mathbf{q}^* = \mathbf{q}\left(\sigma^{\mathbf{q}^*}\right). \tag{57}$$

In the following, I use the notation $\Pr(\alpha|\text{piv};\mathbf{q})$ to denote the posterior consistent with (8) and the vote shares $\mathbf{q}$, and also further analogous notation. The next two sections provide an analysis of the best response function $q(\sigma^-)$ in two steps. Section 4.5.2 describes the pivotal inference given vote shares $\mathbf{q}$. Section 4.5.3 describes the vote shares of the best response, given some pivotal inference about the state.

4.5.2 Inference in Large Elections

We 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 6 *Take any strategy σ for which $\Pr(\text{piv}|\beta;\sigma,n) \in (0,1)$. If*

$$\left|q(\alpha;\sigma) - \frac{1}{2}\right| < (\leq) \left|q(\beta;\sigma) - \frac{1}{2}\right|, \tag{58}$$

then

$$\frac{\Pr(\text{piv}|\alpha;\sigma,n)}{\Pr(\text{piv}|\beta;\sigma,n)} > (\geq) 1. \tag{59}$$

Proof. The function $q(1-q)$ has an inverse u-shape on $[0,1]$ and is symmetric around its peak at $q = \frac{1}{2}$. 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 (8) that (58) implies (59). ■

Moreover, Lemma 6 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.

[24] The ability to write an equilibrium as a finite-dimensional fixed point via (57) is a significant advantage. Similarly, a reduction to finite dimensional equilibrium beliefs has been useful in other settings; see Bhattacharya (2013b), Ahn & Oliveros (2012) and Heese & Lauermann (2017).

Lemma 7 *Consider any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$. If,*

$$\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right| < (>) \lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|, \tag{60}$$

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

$$\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}n^{-\kappa} = \infty(0). \tag{61}$$

Proof. Let

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

From (8), the left-hand side of (61) 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 in the Appendix. Therefore, (60) 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. ■

4.5.3 Vote Shares and the Citizen's Inference

We show that, as $n \to \infty$, under the best response, the expected vote share for policy A in ω is given by the mass of types preferring A given the pivotal belief $\Pr(\alpha|\text{piv};\sigma'.n)$, that is $\Phi(\Pr(\alpha|\text{piv};\sigma'_n))$.

Lemma 8 *Take any sequence of strategies $(\sigma'_n)_{n\in\mathbb{N}}$. Take the sequence of best responses σ_n. For any $\omega \in \{\alpha,\beta\}$,*

$$\lim_{n\to\infty} q(\omega;\sigma_n) = \lim_{n\to\infty}\Phi(\Pr(\alpha|\text{piv};\sigma'_n)). \tag{62}$$

The proof is provided in the Appendix. The basic intuition is that, as $n \to \infty$, the precision of all types signals goes to zero uniformly, (27), so that, given (11)- (14), "in the limit" voters decide simply according to the pivotal belief.

4.5.4 Intuition: Information Acquisition can be a Complement

Lemma 8 is key to get an intuition why informative equilibrium sequences exist. The relevant economic observation coming from the lemma is that information acquisition can be complementary as a result of the pivotal inference:

Take the case $\Phi(\Pr(\alpha)) > \frac{1}{2}$. Without pivotal inference, $\Pr(\alpha|\text{piv};\sigma_n, n) = \Pr(\alpha)$. Then, under the best response, A is elected with a positive margin as $n \to \infty$, $\lim_{n\to\infty} q(\omega;\sigma_n^*) > \frac{1}{2}$, using Lemma 8 and the weak law of large numbers. Since A is elected with a positive margin, the incentives to get informed are small, in fact, exponentially small, see (26). However, if citizens acquire more information, so that $q(\alpha;\sigma_n)$ and $q(\beta;\sigma_n)$ differ sufficiently much, voters may make an inference about the state when conditioning on the election being tied in a way, so that $\lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv};\sigma_n, n)) = \frac{1}{2}$. Then, under the best response, the election is close to being tied, thereby creating incentives to get informed. This illustrates how information acquisition can spur even more information acquisition, that is information acquisition may be complementary.

4.5.5 Fixed Point Argument

This section uses a fixed point argument to show that there is a sequence of equilibrium vote shares $(\mathbf{q}_n^*)_{n\in\mathbb{N}}$ such that the corresponding sequence of equilibrium strategies are informative. We provide the proof for the case when $\Phi(\Pr(\alpha)) < \frac{1}{2}$ and when the minority group has the higher power, $W(L) < W(C)$. The proof proceeds in two steps. First, we show that for any vote share $q(\alpha)$ in α close to $\frac{1}{2}$, we find a vote share $q_n^*(\beta)$ such that the best response to $(q(\alpha), q_n^*(\beta))$ has again the same vote share in α.

Step 1 *Let $\Phi(\Pr(\alpha)) < \frac{1}{2}$ and $W(L) < W(C)$. 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*

$$q(\alpha) \quad = \quad q(\alpha;\sigma^{(q(\alpha),q_n^*(\beta))}). \tag{63}$$

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

Take $\frac{1}{2} - \frac{\epsilon}{2} \leq q(\alpha) \leq \frac{1}{2}$, and let $\mathbf{q} = (q(\alpha), q(\beta))$ in the following.

Substep 1 *If $q(\beta) = \frac{1}{2} + \epsilon$, then, for $\epsilon > 0$ small enough and n large enough,*

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

The election is more close to being tied in α, and, by Lemma 7, voters become convinced that the state is α, i.e., $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\mathbf{q}, n) = 1$. It follows from

Lemma 8 that $\lim_{n\to\infty} q(\alpha;\sigma^{\mathbf{q}}) = \Phi(1)$. Finally, (64) follows when ϵ is small enough since $\Phi(1) > \frac{1}{2}$.

Substep 2 *If $q(\beta) = \frac{1}{2}$, then for $\epsilon > 0$ small enough and any n,*

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

The election is more close to being tied in β, and, by Lemma 6, voters update towards β, i.e. $\Pr(\alpha|\text{piv};\mathbf{q},n) \le \Pr(\alpha)$. Since $\Phi(\Pr(\alpha)) < \frac{1}{2}$, Lemma 8 implies that $\lim_{n\to\infty} q(\alpha;\sigma^{\mathbf{q}}) < \frac{1}{2}$. Finally, (65) follows when ϵ is small enough.

Since $q(\alpha;\sigma^{\mathbf{q}})$ is continuous in $q(\beta)$, it follows from Substep 1, Substep 2, and the intermediate value theorem that, for n large enough, there is $q_n^*(\beta)$ such that (63) holds. It follows from the implicit function theorem that $q_n^*(\beta)$ is continuous in $q(\alpha)$.

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

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

Substep 1 *For $q(\alpha) = \frac{1}{2}$, and any n large enough,*

$$q(\beta;\sigma^{(q(\alpha),q_n^*(\beta))}) > q_n^*(\beta), \tag{67}$$

Recall that Φ is strictly increasing. Lemma 8 together with (63) implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\mathbf{q}_n,n) = \hat{p} \in (0,1)$ for $\mathbf{q}_n = (\frac{1}{2}, q_n^*(\beta))$. We claim that

$$\delta(\beta)(\mathbf{q}_n) \in \mathbb{R}, \tag{68}$$

where the notation highlights that $\delta(\beta) = \lim_{n\to\infty}(q_n^*(\beta) - \frac{1}{2})\frac{2n+1}{s(\beta;\mathbf{q}_n)}$ depends on $\mathbf{q}_n$. Otherwise, since $\delta(\alpha)(\mathbf{q}_n) = \lim_{n\to\infty}(q(\alpha) - \frac{1}{2})\frac{2n+1}{s(\alpha;\mathbf{q}_n)} = 0$, Observation 2 implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\mathbf{q}_n,n) = 1$, which contradicts the earlier observation $\lim_{n\to\infty} \Pr(\alpha|\text{piv};\mathbf{q}_n,n) \in (0,1)$. We claim that

$$\lim_{n\to\infty} \Big[q(\beta;\sigma^{\mathbf{q}_n}) - q(\alpha;\sigma^{\mathbf{q}_n})\Big] s(\beta;\sigma^{\mathbf{q}_n}) \in \{\infty, -\infty\}. \tag{69}$$

For this, we show that

$$\lim_{n\to\infty} \Big[q(\beta;\sigma^{\mathbf{q}_n}) - q(\alpha;\sigma^{\mathbf{q}_n})\Big] \Pr(\text{piv}|\mathbf{q}_n,n)^{-1} = \infty. \tag{70}$$

To see why, note that

$$
\begin{aligned}
q(\beta;\sigma^{\mathbf{q}_n}) - q(\alpha;\sigma^{\mathbf{q}_n}) &= 2\Big[\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t)\Big] \\
&\approx \Big[W(L) - W(C)\Big] \Pr(\text{piv}|\mathbf{q}_n, n)^{\frac{2}{d-1}} c_2, \qquad (71)
\end{aligned}
$$

where the first equality restates (51), and the second line follows from Lemma 4. Using that the pivotal likelihood goes to zero as $n \to \infty$, (70) follows from $d > 3$, (71), and the genericity conditions. Then, (69) follows from (70), Observation 2 and $\delta(\omega)(\mathbf{q}_n) \in \mathbb{R}$. Note that $q(\alpha;\sigma^{\mathbf{q}_n} = \frac{1}{2}$, given (63), and that $q(\beta;\sigma^{\mathbf{q}_n}) > q(\alpha;\sigma^{\mathbf{q}_n})$ for n large, given (48) and $W(L) < W(C)$. Therefore, (68) and (69) together imply (67).

Substep 2 *For $q(\alpha) = \frac{1}{2} - \epsilon$, and any n large enough,*

$$
q(\beta;\sigma^{(q(\alpha),q_n^*(\beta))}) < q_n^*(\beta), \tag{72}
$$

Lemma 8 together with (63) implies $\lim_{n\to\infty} q(\beta;\sigma^{(q(\alpha),q^*(\beta))}) = \frac{1}{2} - \epsilon$. Since $q_n^*(\beta) > \frac{1}{2}$ by construction, (72) holds for n large enough.

Finally, using (72) and (67) and that $q(\beta;\sigma^{(q(\alpha),q_n^*(\beta))})$ is continuous in $q(\alpha)$, the claim of Step 2 follows from an application of the intermediate value theorem .

It follows from Step 1 and Step 2 that for any n large enough, there is a pair of vote shares $q_n^*(\alpha)$ such that $\mathbf{q}_n^* = (q_n^*(\alpha), q_n^*(\beta))$ is a fixed point of $\mathbf{q}(\sigma^-)$. Moreover $q_n^*(\alpha) \leq \frac{1}{2} \leq q_n^*(\beta)$ by construction, implying that $\lim_{n\to\infty} \Pr(A|\alpha;\sigma^{\mathbf{q}}, n) \leq \frac{1}{2} \leq \lim_{n\to\infty} \Pr(A|\beta;\sigma^{\mathbf{q}}, n)$. Recalling that limit equilibrium outcomes are determinate when $d > 3$ (see (49)), this implies that the equilibrium sequence is informative. This concludes the proof when $W(C) > W(L)$ and $\Phi(\Pr(\alpha)) < \frac{1}{2}$. The other cases are analogous.

4.6 Weighted Welfare Rules

This section shows that for a large class of settings, elections lead to outcomes that maximize a weighted welfare rule. Roughly speaking, the result holds under independence conditions which imply that the utilities of the critical types are representative of the whole population.

Independence Conditions. We consider preference distributions for which the conditional distribution of the threshold of doubt, $F(\cdot|t \in g)$, is independent of the voter group, i.e. for all $g \in \{L, C\}$,

$$F(\cdot|t \in g) = F. \tag{73}$$

The conditional distribution $J(\cdot|t \in g)$ of the total intensities of types $t \in g$ is independent from F, that is, for all $g \in \{C, L\}$ and all $y \in [0, 1]$

$$J(\cdot|t \in g, y(t) = y) = J(\cdot|t \in g). \tag{74}$$

Recall that partisans stay uninformed and simply vote for their preferred policy, so that the information cost cannot screen their intensities. Therefore, we consider settings without partisans,[25]

$$\Pr(\{t \in L\}) \cup \Pr(\{t \in C\}) = 1. \tag{75}$$

Weighted Welfare. For any $\kappa \in [0, 1]$, any state ω, the κ-weighted welfare of A is

$$\sum_{i=1,\ldots,2n+1} (t_\omega(i))^\kappa, \tag{76}$$

where we added the label i of each citizen to the notation. The κ-weighted welfare of B is zero. Given the independence assumptions (73) and (74), $W(L) > W(C) \Leftrightarrow \Pr(\{t : t \in L\})\mathrm{E}(k(t)^{\frac{2}{d-1}}|t \in L) > \Pr(\{t : t \in C\})\mathrm{E}(k(t)^{\frac{2}{d-1}}|t \in C)$. Using $t_\alpha = k(t)(1 - y(t))$ and $t_\beta = k(t)y(t)$ and the assumption that the total intensity $k(t)$ is independent of the threshold of doubt $y(t)$,

$$\begin{aligned} W(L) &> W(C) \\ \Leftrightarrow \Pr(\{t : t \in L\})\mathrm{E}(t_\omega^{\frac{2}{d-1}}|t \in L) &> \Pr(\{t : t \in C\})\mathrm{E}(t_\omega^{\frac{2}{d-1}}|t \in C). \end{aligned} \tag{77}$$

[25] In fact, all results hold under the weaker condition that the welfare at stake is, in expectation, the same for A-partisans and B-partisans,

$$\begin{aligned} & \Pr(\{t : t_\alpha > 0, t_\beta > 0\})\mathrm{E}_H(|t_\omega||\omega, \{t : t_\alpha > 0, t_\beta > 0\}) \\ = \; & \Pr(\{t : t_\alpha < 0, t_\beta <\})\mathrm{E}_H(|t_\omega||\omega, \{t : t_\alpha < 0, t_\beta < 0\}). \end{aligned}$$

for all $\omega \in \{\alpha, \beta\}$.

for any state ω. Therefore, Theorem 1 together with the weak law of large numbers yields:

Theorem 2 *Let* $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$. *Take any preference distribution* H *satisfying the genericity conditions and the independence conditions (73)- (75). For all informative equilibrium sequences, the elected policy maximizes* κ*-weighted welfare, for* $\kappa = \frac{2}{d-1}$ *with probability converging to* 1, *as* $n\to\infty$.

5 Non-Informative Equilibrium Sequences

This section shows that there are two types of non-informative equilibrium sequences, thereby finishing the complete characterization of equilibrium sequences.

5.1 Voting According to the Prior is a Limit Equilibrium

There is an equilibrium sequence where, as $n\to\infty$, all citizens vote according to the prior belief. Hence, A is elected when a majority prefers A given the prior belief, $\Phi(\Pr(\alpha)) > \frac{1}{2}$, and B is elected when a majority prefers B given the prior belief, $\Phi(\Pr(\alpha)) < \frac{1}{2}$. The proof is in the Appendix.

Theorem 3 *Let* $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$. *There exists an equilibrium sequence* $(\sigma_n^*)_{n\in\mathbb{N}}$ *for which*

$$\lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*,n) = \lim_{n\to\infty}\Pr(A|\beta;\sigma_n^*,n) = \begin{cases} 1 & \textit{if } \Phi(\Pr(\alpha)) > \frac{1}{2}, \\ 0 & \textit{if } \Phi(\Pr(\alpha)) < \frac{1}{2}, \end{cases} \tag{78}$$

Theorem 3 and Theorem 1 show that citizens may coordinate on acquiring much information, but they may also (mis)coordinate on acquiring very few information. The proof of Theore 3 highlights the role of the complementarity of information acquisition. Given the equilibrium sequence that converges to "voting according to the prior", citizens acquire very few information, so that the vote shares are approximately the same in each state. As a consequence, the pivotal event contains no information, $\lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma_n^*,n) = \Pr(\alpha)$, and $\lim_{n\to\infty} q(\omega;\sigma_n^*) = \Phi(\Pr(\alpha)) \neq \frac{1}{2}$ (see Lemma 8), so that either policy A or policy B wins by a clear margin. Anticipating this, citizens have in fact low incentives to get informed since the individual likelihood of affecting the outcome is exponentially small.

5.2 All Other Equilibria

We complete the characterization of equilibrium sequences. We show that when $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$, there is a third type of equilibrium sequence. This equilibrium sequence leads to the outcome that is preferred by the voter group with the larger power given the prior belief. The proof is in the Appendix.

Theorem 4 *Take any preference distribution H satisfying the genericity conditions.*

1. *If $\lim_{x\to 0} \frac{c'(x)x}{c(x)} < 3$, all equilibrium sequences satisfy (78).*
2. *If $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$, there are three types of equilibrium sequences. There is an informative equilibrium sequence satisfying (47). There is an equilibrium satisfying (78), and there is an equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ with*

$$\lim_{n\to\infty} \Pr(z(\omega)|\alpha; \sigma_n^*, n) = \lim_{n\to\infty} \Pr(A|\beta; \sigma_n^*, n) = 1 \tag{79}$$

   where $z(\omega)$ is the outcome preferred by the group g' with the larger power, $g' = \arg\max_{g\in\{L,C\}} W(g)$.

3. *Any equilibrium sequence satisfies either (47), (78), or (79).*

The basic intuition for why there is another equilibrium sequence comes again from the observation that information acquisition of citizens can be complementary, as discussed in Section 4.5.3.

For illustration, let $\hat{p} > \Pr(\alpha)$, so a majority prefers B given the prior. We argue that, when the contrarians have a larger power, $W(L) < W(C)$, then, there is an equilibrium sequences where A is elected in both states. To construct such an equilibrium sequence, we employ a fixed point argument similar to the one for the informative equilibrium sequence in Section 4.5.5, . We show that there are equilibrium vote shares $\mathbf{q}_n^* = (q(\alpha)_n, q(\beta)_n)$, satisfying

$$\Phi(\Pr(\alpha)) < \frac{1}{2} < q_n(\alpha) < q_n(\beta) \tag{80}$$

for n large such that $\Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n^*)) \to \frac{1}{2}$ as $n \to \infty$. Policy A is elected in both states cince equilibrium outcomes are determinate, as $n \to \infty$, when $\lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$ (see (49)). Information acquisition is complementary in the following sense: citizens acquire information such that the inference about the state implies $\Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n, n) \approx \frac{1}{2}$. Thus, the resulting vote shares are close to $\frac{1}{2}$ by (8), making the election close to being tied, and thereby creating incentives for all citizens to acquire information.

6 Discussion and Extensions

6.1 Heterogenous Information Access and Skills

Access to information sources and the ability to interpret information vary widely across citizens. We can capture this in the alternative model where the attention cost of the citizens depends on a private type $\gamma \in [\frac{1}{M}, M]$ for $M > 0$, and γ is drawn i.i.d. across voters from some absolutely continuous distribution with strictly positive density. 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.

It turns out that the previous analysis already captures this alternative model since cost and preference intensities are strategically equivalent: precisely, 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, (11)-(14), (18), (37) and (38). Therefore, it is without loss to treat the additional heterogeneity in terms of cost as part of the preference type distribution; for any distribution of γ and H, call $\hat{H}$ the *induced* preference distribution, capturing both types of heterogeneity.

When the effort type is independent of the preference types and signals of the voters, the previous welfare results (e.g. Theorem 2) carry over. This is for two reasons: first, independence implies that the policies maximizing κ-weighted welfare are the same under H and $\hat{H}$ as $n \to \infty$.[26] Second, if H satisfies the independence conditions (73)-(75), then so does $\hat{H}$.

More interesting are the situations where attention cost and preference types are correlated. It can happen that such correlation hinders welfare-

[26] Since γ and t are independent, $\mathrm{E}((\frac{1}{\gamma}k(t)y(t))^\kappa) = \mathrm{E}(\frac{1}{\gamma}^\kappa)\mathrm{E}(k(t)y(t))^\kappa)$. Thus $\mathrm{E}((\frac{1}{\gamma}k(t)y(t))^\kappa) > 0$ is equivalent to $\mathrm{E}(k(t)y(t))^\kappa) > 0$.

efficient outcomes. An example: suppose that elder people prefer policies aligned with the state and younger people do not. Empirically, elder people care a lot about healthcare issues. Thus, suppose that it is utilitarian to choose their preferred policy. Typically, elder people are also less educated in information technologies. One can show, that, when, ceteris paribus, effort cost are much higher for the elder, their electoral power $W(L)$ is relatively low, and their preferred policy is not elected in any informative equilibrium, given $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$.

6.2 Third-Party Manipulation: Obfuscation of Voters

From the entertainment of the arena in ancient rome to hollow media campaigns on social media platforms nowadays, diverting the attention of the people from important economic and political issues, is an ubiquitous tool of politicians for managing democracies. We ask: how manipulable are elections by hollow information provision of third-parties? To analyze this question, we consider the alternative model where a third party can send a signal to specific voters, and the signal is uninformative for the issue relevant to the election ("obfuscation"). The game is as before, except that voters of the targeted group draw an uninformative signal with a given probability $\tilde{q}$, and else the costly signal with the precision $x(t)$ as acquired. Obfuscation has two effects. First, there is a *direct effect* on the precision of targeted voters; the average precision of a targeted voter choosing $x(t)$ is

$$(1-\tilde{q})x(t). \tag{81}$$

There is also an *indirect effect* since the targeted voter anticipates drawing an uninformative signal. This reduces the excepted benefit as well as the expected marginal benefit of her private information. One can show, that, as a consequence, a voter of a given type t is less likely to acquire any information when targeted relative to when not, and also chooses a lower precision.

Now—similar to the analysis before— the decision weight of each individual voter is given by her average precision. The *obfuscated power* of a voter group g is

$$\tilde{W}(g,\tilde{q}) = (1-\tilde{q})W(g). \tag{82}$$

The analogue of Theorem 1 holds: when $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$, in any informative equilibrium sequence, the policy preferred by the voter group with the larger power $\tilde{W}(g,\tilde{q})$ is elected. This illustrates the effectiveness of the obfuscation of voters, and implies:

Theorem 5 *Let $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$. Take any preference distribution H satisfying the genericity conditions and $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$. There is $\bar{q} < 1$, so that, if the third-party obfuscates a group g with a likelihood $\tilde{q} > \bar{q}$, then, for all informative equilibrium sequences $(\sigma_n^*)_{n\in\mathbb{N}}$,*

$$\lim_{n\to\infty} \Pr(z(\omega)|\omega;\sigma_n^*,n) \quad = \quad 0$$

for all $\omega \in \{\alpha,\beta\}$, where $z(\omega)$ is the policy preferred by the obfuscated voter group in ω.

6.3 Polarized Preferences

This section shows 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 in the group.

First, Lemma 9 shows that the relative power of a voter group is smaller when the preference intensities are more dispersed within the group. A preference distribution H' is a *g-intensity spread of H* if, ceteris paribus,

$$J(-|t\in g;H) \quad <_{\text{mps}} \quad J(-|t\in g;H'), \tag{83}$$

where $J(-|t \in g; H)$ is the conditional distribution of the (total) intensities $k(t)$ of the types $t \in g$, and where (83) means that $J(-|t \in g; H')$ is a mean-preserving spread of $J(-|t \in g; H)$, and by ceteris paribus, we mean that the conditional distribution of the preference types $t \in g' \neq g$ is unchanged as well as the conditional distribution of the threshold of doubt $y(t)$ of the types $t \in g$ and also the likelihood of a type being aligned or contrarian.

Lemma 9 *Let $d = \lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$. Let $g \in \{C, L\}$. Take any preference distributions H, H' satisfying (73) - (75) and the genericity conditions.*

1. *If H' is an L-intensity spread of H,*

$$\frac{W_{H'}(L)}{W_{H'}(C)} < \frac{W_H(L)}{W_H(C)}. \tag{84}$$

2. If H' is a C-intensity spread of H,

$$\frac{W_{H'}(L)}{W_{H'}(C)} > \frac{W_H(L)}{W_H(C)}. \tag{85}$$

The proof is in the Appendix. The basic argurment is that, when $d > 3$, the power of the group, $W(g)$, is proportional to the mean of a concave function of the intensities, $\mathrm{E}(k(t)^{\frac{2}{d-1}})$ see the definition in (4). The result will follow from an application of Jensen's inequality.

We lift the restriction that $t \in [-1, 1]^2$, and allow for more extreme preference types $t \in [-M, M]^2$ for $M > 0$. When M is arbitrarily large, there can be arbitrarily large within-group preference dispersion. Theorem 1 still holds, and based on it, we show that, when, ceteris paribus, the intensities within a given voter group are sufficiently dispersed, for *all* informative equilibrium sequences, the outcome preferred by the voter group is elected with probability going to 0 as $n \to \infty$. The formal statement and the proof are in the Appendix in Section G.

6.4 Further Remarks

Median-Voter Outcomes. Whenever the contrarians have a larger power, $W(L) < W(C)$, then, the vote shares are ordered as $q(\alpha; \sigma_n^*) < q(\beta; \sigma_n^*)$ in any equilibrium when n is large, see (48). This implies, in particular, that the median voter-preferred policy is less likely to be elected in one of the states since the median voter prefers A only in α.

Median-Voter Theorem with Common Interests. Suppose that all voters share a common interest, $\Pr_H(\{t \in C\}) = 0$. For such situations, Theorem 1 implies that whenever information of low precision $x \approx 0$ is sufficiently cheap, $d > 3$, there is an equilibrium of the large election where the median-voter preferred outcome is elected state-by-state. In particular, outcomes are equivalent to the outcome with publicly known states ("full-information equivalence"). This has only been known for certain symmetric settings so far (Martinelli (2006), Oliveros (2013b)).

Aggregate Cost. We show that the sum of the voters' cost converges to zero in all equilibrium sequences when $\lim_{x \to 0} \frac{c'(x)x}{c(x)} \neq 3$. The proof is in the Appendix.

Lemma 10 *Let* $\lim_{x \to 0} \frac{c'(x)x}{c(x)} \neq 3$. *Take any preference distribution satisfying the genericity conditions. Take any equilibrium sequence* $(\sigma_n^*)_{n \in \mathbb{N}}$ *and let* x_i *be the realisation of the precision of voter* $i \in \{1, \ldots, 2n+1\}$. *Then,*

$$\lim_{n \to \infty} \Big[\sum_{i=1,\ldots,2n+1} c(x_i) \Big] = 0. \tag{86}$$

The lemma qualifies the discussion of welfare implications in Section 4.6 that does not take into account the costs of the voters.

Non-Monotone Preferences. So far, we provided the analysis assuming that the aggregate preference function Φ is strictly monotone. When Φ is non-motone, there may be multiple beliefs $\hat{p}$ for which $\Phi(\hat{p}) = \frac{1}{2}$. One can show that, for any such $\hat{p}$, there will be two equilibrium sequences, both satisfying $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n) = \hat{p}$. There is one informative equilibrium sequence, for which the outcome preferred by the group with the larger power $W(g, \hat{p})$ is elected state-by-state. And, there is one non-informative equilibrium sequence, for which the outcome preferred by the group with the larger power $W(g, \hat{p})$ given the prior belief, is elected; compare to Theorem 4. In particular, it may happen that different outcomes arise in different informative equilibria since the power of a voter gorup is a local notion when Φ is non-monotone.[27]

7 Literature

Information Aggregation Literature. This paper contributes to the literature on information aggregation in large elections. Condorcet's Jury Theorem (1785) states that if voters have common interests, but the information is dispersed throughout the electorate, then 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 Condorcet's Jury Theorem in a setting where citizens vote strategically. Their results show that election outcomes

[27]These results mirror known results for the model with exogenous information: if citizens were to receive a binary, conditionally i.i.d. signal about the state and Φ is non-monotone, it is known that there is a multiplicity of equilibrium sequences, some of which do not aggregate information (Bhattacharya (2013a)).

are " full-information equivalent", that is, as if citizens have no uncertainty about the state. However, full-information equivalent outcomes are not necessarily socially optimal when voters have conflicting interests: take a situation where 51% of citizens marginally benefit from a reform, while the other 49% are severely impacted by it. This paper points at an empirical observation that has been mostly overlooked in this context: namely, that the dispersion of the voters' information is endogenous. We show how, for a large class of settings, the information being endogenous leads to equilibria with outcomes that maximize a weighted welfare rule (Theorem 2).

This paper also contributes to the literature on elections with costly information acquisition by studying a general setup that allows the voters to have conflicting interests. Thereby, we capture many relevant economic applications; for example, distributive reforms. The previous literature has studied information aggregation in situations where all voters share a common interest.[28] For the common interest case, we generalize the result of the literature showing that information aggregation is possible under a condition on the cost function provided in Martinelli (2006). We show that the possibility result extends to general continuous preference distributions, see the discussion in Section 6.4. Also, we characterize all the equilibria of the voting game, revealing an equilibrium multiplicity, and establishing that, generically, information aggregation only occurs in one of three equilibria (Theorem 4).

Vote Buying and Costly Voting Literature. This paper is related to work on elections with voting cost and vote-buying. Krishna & Morgan (2011, 2015) have shown that elections yield first-best outcomes when voting is voluntary and costly. In a companion paper Heese (2020), we show that analogous results hold when voters have the binary choice between a costless uninformative signal and a given costly informative signal, similar to the binary choice between voting at a cost and not voting.

The model in this paper is more closely related to the literature on vote-buying. Lalley & Weyl (2018) have shown that equilibrium outcomes in large electorates are utilitarian when each voter can buy any number of votes at a total price that is quadratic in the number of votes bought. Similarly,

[28] See Martinelli (2006) and Oliveros (2013a)), and the more distantly related papers Triossi (2013) and Martinelli (2007) who study heterogeneous cost in common interest setups, and Oliveros (2013b) who studies the relationship of abstention and information cost.

this paper shows that when information is costly and cost are arbitrarily close to "cubic", e.g. $c(x) = x^{3+\epsilon}$, there are equilibrium sequences where limit outcomes maximize utilitarian welfare for a large class of preference distributions.[29]

Eguia & Xefteris (2018) show that vote-buying mechanisms with general price functions implement a set of weighted welfare rules. Similarly, we have shown that a subset of the same weighted welfare rules arises when political infomation is costly (Theorem 2).

Distributive Politics Literature. A rich literature in distributive politics seeks to understand if, and when, elections select policies that maximize social welfare. See e.g. Fernandez & Rodrik (1991), Alesina & Rodrik (1994) and Persson & Tabellini (1994). This paper introduces a novel aspect into this discussion; namely, endogenous attention to politics.[30] Fernandez & Rodrik (1991) study the effect of asymmetric information on distributive politics: there is a group of citizens who gain from a reform with certainty; however, for a majority, the individual consequences are uncertain, and given the prior, each majority voter's expected gain is negative. Without further information, this leads to rejection of the reform in a simple majority vote, even when the reform enhances the utilitarian welfare of the electorate as a whole. We would like to point out that these results may not carry over when citizens can acquire information about the distributive consequences. Future work may investigate the closer connection to this literature.

8 Conclusion

A modified version of the classical setting by Feddersen & Pesendorfer (1997) captures applications like distributive reforms, e.g. health care or education reforms. Election results are driven by how much demographic groups pay attention to politics. In all limit equilibria with state-dependent outcomes, the implicit decision weight of a voter is proportional to how much attention she

[29]Given the assumptions of Theorem 2, the informative equilibrium sequences lead to outcomes maximizing $\frac{1}{1+\epsilon}$-weighted welfare. Note that utilitarian welfare is what we call 1-weighted welfare. Hence, these equilibrium sequences are utilitarian except for the small set of preference distributions where the policy maximizing 1-and $\frac{1}{1+\epsilon}$-weighted welfare is not the same.

[30]Similar to this paper, Ali *et al.* (2018) transports the informational approach to elections to the literature on distributive politics.

pays to politics. This is a structural insight with wide-reaching consequences. Since citizens with higher utilities pay more attention, elections screen the voter's utilities, and the result implies strong welfare properties of elections for a large class of settings. Elections lead to policies maximizing a certain weighted welfare rule.

The results, albeit implying a positive welfare theorem when information cost are symmetric across voters, point at the scope of manipulability of elections through informational campaigns. Politicians and third parties may successfully affect elections by diverting attention of targeted groups, thereby reducing their effective electoral power. They may successfully affect elections by hampering the physical access to information, or by spreading confusion among target groups; in other words, by making it more costly to acquire knowledge about policies and their consequences. We believe that this paper can be a starting point for the analysis of many current topics concerning the role of information in elections.

Appendices

A Auxiliary Results

The auxiliary results are used in the proofs of this Appendix. Some of the auxiliary results will be restated as lemmas or observations in the main text when needed for the arguments there.

A.1 Pivotal Likelihood Ratio

For any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and any n, let $\mathbf{q}(\sigma_n) = (q(\alpha;\sigma_n), q(\beta;\sigma_n))$, and denote by $s(\omega;\mathbf{q}(\sigma_n)) = \Big[q(\omega;\sigma_n)(1-q(\omega;\sigma_n)(2n+1)\Big]^{\frac{1}{2}}$ the standard deviation of the vote share in ω. Let

$$\delta_n(\omega) = \frac{2n+1}{s(\omega;\mathbf{q}(\sigma_n))}\Big[q(\omega;\sigma_n) - \frac{1}{2}\Big] \tag{87}$$

be the normalized distance of the expected vote share to the majority threshold, and $\delta(\omega) = \lim_{n\to\infty}\delta_n(\omega)$.

Lemma 11 *For any sequence of strategies* $(\sigma_n)_{n\in\mathbb{N}}$*,*

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} = \Big[1-\frac{1}{2n+1}x_n\Big]^n. \tag{88}$$

for

$$x_n = \frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\delta_n(\alpha)^2-\delta_n(\beta)^2. \tag{89}$$

Proof. Recall the definitions of $\delta_n(\omega)$ and $s(\omega;\mathbf{q}(\sigma_n))$,

$$\begin{aligned}\delta_n(\omega) &= \frac{2n+1}{s(\omega;\mathbf{q}(\sigma_n))}(q(\omega;\sigma_n)-\frac{1}{2})\\ &= (2n+1)^{\frac{1}{2}}\frac{q(\omega;\sigma_n)-\frac{1}{2}}{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}.\end{aligned} \tag{90}$$

The ratio of the likelihoods of the pivotal event in the two states is

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

where we used (90) for the equality on the last line. Plugging in (89) yields (88). ■

Lemma 12 *Take any sequence of strategies* $(\sigma_n)_{n\in\mathbb{N}}$*. If* $\lim_{n\to\infty}\delta_n(\alpha)-\delta_n(\beta)=0$*, then*

$$\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\beta;\sigma_n^*,n)}=1, \tag{91}$$

Proof. Recalling Lemma 11, we rewrite (88),

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\beta;\sigma_n^*,n)} = (\Big[1-\frac{1}{2n+1}x_n\Big]^n - e^{-\frac{1}{2}x_n}) + e^{-\frac{1}{2}x_n} \tag{92}$$

with x_n given by (89). In the following, we analyse the two summands separately. Note that $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0$ implies $\lim_{n\to\infty} q(\alpha;\sigma_n) - q(\beta;\sigma_n) = 0$, and therefore

$$\lim_{n\to\infty} x_n = 0. \tag{93}$$

This yields

$$\lim_{n\to\infty} e^{-\frac{1}{2}x_n} = 1, \tag{94}$$

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

$$|(1 - \frac{x_n}{(2n+1)})^n - e^{-x_n}| \leq \frac{x_n^2}{(2n+1)^3} \tag{95}$$

Finally, (91) follows from (92) - (95). ■

A.2 Proof of Observation 1: Outcome Distribution

Observation 1 *Take 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*

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

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

Proof. Let $q_n = q(\omega, \sigma_n)$. By using the normal approximation[31]

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

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

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

[31] 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 sufficiently large the indicator function in the condition takes the value zero.

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 formula $s(\omega; \mathbf{q}(\sigma_n)) = \Big[q(\omega;\sigma_n)(1 - q(\omega;\sigma_n)(2n+1)\Big]^{\frac{1}{2}}$. ■

A.3 A Lemma on the Optimal Precision

Lemma 13

$$\lim_{n\to\infty} \frac{\partial x^*(y,k;\sigma_n,n))}{\partial y} = 0. \tag{96}$$

uniformly for all (y,k).

Proof. Implicit differentiation of the first-order condition (18) shows

$$\frac{\partial x^*(y,k;\sigma_n',n))}{\partial y} = \frac{MB'(y)}{c''(x^*(y,k;\sigma_n',n))}. \tag{97}$$

Using (17) and (18), $MB'(y) = \Pr(\text{piv}|\sigma',n)\Big[\Pr(\beta|\text{piv};\sigma',n) - \Pr(\alpha|\text{piv};\sigma',n)\Big] = c'(x^*(y,k;\sigma_n',n))c_2$ for some constant $c_2 \in \mathbb{R}$. Therefore, (27) together with $\lim_{x\to 0}\frac{c'(x)}{c''(x)} = \lim_{x\to 0}\frac{x}{d-1} = 0$ imply (35). ■

A.4 Proof of Lemma 8: Limit Vote Shares

Lemma 8 *Take any sequence of strategies* $(\sigma_n')_{n\in\mathbb{N}}$. *Take the sequence of best responses* σ_n. *For any* $\omega \in \{\alpha, \beta\}$,

$$\lim_{n\to\infty} q(\omega;\sigma_n) = \lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv};\sigma_n')). \tag{62}$$

Proof. Recall that $\Pr(\text{piv}|\sigma_n') \to 0$ as $n \to \infty$. Therefore, the first-order condition (18) implies that $x(t) \to 0$ uniformly. Hence, for any private signal realization s of a voter type, $\lim_{n\to\infty} \Pr(\alpha|\text{piv}, s; \sigma_n^*, n) - \Pr(\alpha|\text{piv}; \sigma_n^*, n) = 0$. Thus, (11)-(14) imply $\lim_{n\to\infty} q(\omega; \sigma_n^*) = \Phi(\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*))$. Finally, (62) follows since Φ is continuous. ■

B Proof of Lemma 1

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 $x > 0$, we show below that the posteriors are ordered as

$$\Pr(\alpha|b, \text{piv}; \sigma', n) < \Pr(\alpha|a, \text{piv}; \sigma', n). \tag{98}$$

We argue that $x(t) > 0$ implies

$$\Pr(\alpha|b, \text{piv}, \sigma', n) < y(t) < \Pr(\alpha|b, \text{piv}, \sigma', n). \tag{99}$$

Otherwise, given (11)-(14), there is a policy $z \in \{A, B\}$ that the voter weakly prefers, independent of her private signal $s \in \{a, b\}$. But then, she would be strictly better off by not paying for the information $x(t) > 0$ and simply voting the same after both signals. Finally, (11)-(14), and (99) together imply (20).

B.1 Proof of (98)

Note that the posterior likelihood ratio of the states conditional on a signal $s \in \{a, b\}$ with precision $x(t)$ and the event that the voter is pivotal is

$$\frac{\Pr(\alpha|s, \text{piv}; \sigma', n)}{\Pr(\beta|s, \text{piv}; \sigma', n)} = \frac{\Pr(\alpha)}{\Pr(\beta)} \frac{\Pr(\text{piv}|\alpha; \sigma', n)}{\Pr(\text{piv}|\beta; \sigma', n)} \frac{\Pr(s|\alpha; \sigma)}{\Pr(s|\beta; \sigma)}, \tag{100}$$

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. Then, the order of the likelihood ratios in (98) follows from $\Pr(a|\alpha; \sigma) = \frac{1}{2} + x$ and $\Pr(a|\beta; \sigma) = \frac{1}{2} - x$, and the analogous formula for $s = b$.

C Proof of Lemma 2

Step 1 *There is $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$: take any strategy σ'. For any t, $x(t) > 0$ if and only if*

$$\frac{1}{2} + x^{**}(t) \geq \chi(y(t)) \geq \frac{1}{2} - x^{**}(t) \tag{101}$$

for $\chi(y) = \frac{\Pr(\beta|\text{piv};\sigma',n)y}{\Pr(\alpha|\text{piv};\sigma',n)(1-y)+\Pr(\beta|\text{piv};\sigma',n)y}$ and $x^{**}(t;\sigma',n) = x^*(t;\sigma',n)(1 - \frac{c(x^*(t;\sigma',n))}{x^*(t;\sigma',n)c'(x^*(t;\sigma',n))})$, where $x^*(t;\sigma',n)$ is the unique solution to the first-order condition (18).

Proof. Take an aligned type. Recall that, if $x(t) > 0$, then, $x(t) = x^*(t;\sigma',n)$, and her expected utility from the policy elected in the pivotal event is given by (15) in α and by (16) in β. Hence, an aligned type prefers choosing precision $x = x^*(t;\sigma',n)$ over voting A without further information if

$$\begin{aligned} & \Pr(\text{piv}|\sigma',n)\Big[\Pr(\alpha|\text{piv};\sigma',n)(\frac{1}{2}+x)t_\alpha + \Pr(\beta|\text{piv};\sigma',n)(\frac{1}{2}-x)t_\beta\Big] - c(x) \\ \geq \; & \Pr(\text{piv}|\sigma',n)\Big[\Pr(\alpha|\text{piv};x,\sigma',n)t_\alpha + \Pr(\beta|\text{piv};\sigma',n)t_\beta\Big]. \end{aligned} \tag{102}$$

Rearranging,

$$\begin{aligned} & \Pr(\text{piv}|\sigma',n)\Big[(\frac{1}{2}+x)\Big[\Pr(\alpha|\text{piv};\sigma',n)t_\alpha - \Pr(\beta|\text{piv};\sigma',n)t_\beta\Big] + \Pr(\beta|\text{piv};\sigma',n)t_\beta\Big] - c(x) \\ \geq \; & \Pr(\text{piv}|\sigma',n)\Big[\Pr(\alpha|\text{piv};\sigma',n)t_\alpha - \Pr(\beta|\text{piv};\sigma',n)t_\beta + 2\Pr(\beta|\text{piv};\sigma',n)t_\beta\Big] \end{aligned} \tag{103}$$

Plugging (17) and (18) into (103),

$$\begin{aligned} & (\frac{1}{2}+x)c'(x) - c(x) + \Pr(\text{piv}|\sigma',n)\Pr(\beta|\text{piv};\sigma')t_\beta \\ \geq \; & c'(x) + 2\Pr(\text{piv}|\sigma',n)\Pr(\beta|\text{piv};\sigma',n)t_\beta. \end{aligned} \tag{104}$$

We divide by $c'(x)$ rearrange, and use (18) and (17) again,

$$(\frac{1}{2}+x) - \frac{c(x)}{c'(x)} \geq 1 + \frac{\Pr(\beta|\text{piv};\sigma',n)t_\beta}{\Pr(\alpha|\text{piv};\sigma',n)t_\alpha + \Pr(\beta|\text{piv};\sigma',n)(-t_\beta)}. \tag{105}$$

Using $t_\alpha = k(t)(1-y(t))$ and $t_\beta = k(t)y(t)$,

$$(\frac{1}{2}+x) - \frac{c(x)}{c'(x)} \geq 1 + \frac{-\Pr(\beta|\text{piv};\sigma',n)y(t)}{\Pr(\alpha|\text{piv};\sigma',n)(1-y(t)) + \Pr(\beta|\text{piv};\sigma',n)y(t)}. \tag{106}$$

Rearranging gives the right inequality of (101). 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 the left inequality of (101) holds. The argument for the contrarian types is analogous. ■

Step 2 *For any $g \in \{L, C\}$, any $k > 0$ and any $\epsilon > 0$, there is $\delta > 0$ such that the derivatives of*

$$\frac{1}{2} + x^{**}(y,k) - \chi(y), \quad \textit{and,} \tag{107}$$

$$\frac{1}{2} - x^{**}(y,k) - \chi(y) \tag{108}$$

are negative and bounded above by $-\delta$.

Proof. Since $\lim_{x\to 0} \frac{c'(x)x}{c(x)} = d$, Lemma 13 implies that the derivative of $x^{**}(y,k;\sigma',n)$ with respect to y converges to zero uniformly as $n \to \infty$. Not 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.[32] For n large, enough, (107) and (108) follow. ■

Now, we finish the proof of Lemma 2. Note that $\chi(\hat{y}_n) = \frac{1}{2}$ and $x^{**}(\hat{y}_n, k) > 0$ for $\hat{y}_n = \Pr(\alpha|\text{piv}; \sigma', n)$. Thus, $\chi(\hat{y}_n) < \frac{1}{2} + x^{**}(\hat{y}_n, k)$ and $\chi(\hat{y}_n) > \frac{1}{2} - x^{**}(\hat{y}_n, k)$. It follows from Step 1 and Step 2 and since $\lim_{n\to\infty} x^{**}(\hat{y}_n, k) = 0$, that, for any n large enough, there are $\phi_g^-(k), \phi_g^+(k)$ with $\phi_g^-(k) < \Pr(\alpha|\text{piv}; \sigma_n, n) < \phi_g^+(k)$ such that it is optimal to acquire information if and only if $y(t) \in [\phi_g^-(k), \phi_g^+(k)]$.

[32] 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 large enough.

D Proof of Lemma 4: Aggregation over $k = k(t)$

Here, we finish the proof of Lemma 4. We have

$$\begin{aligned}\int_{t\in g} x(t)dH(t) &= \Pr(t\in g)\mathrm{E}(x(t)|t\in g)\\ &= \Pr(t\in g)\mathrm{E}(\mathrm{E}(x(t)|t\in g, k(t)=k))\\ &= \Pr(t\in g)\int_{k=k(t)} \mathrm{E}(x(t)|t\in g, k(t)=k)dH(k(t)|t\in g),\end{aligned}$$

where we used the law of iterated expectations for the second equality and where $H(k(t)|t\in g)$ is the conditional distribution of the total intensity of the types $t\in g$. Using (41),

$$\begin{aligned}&\int_{t\in g} x(t)dH(t)\\ \approx\ &\Pr(t\in g)\int_{k=k(t)} f(\Pr(\alpha|\text{piv})|k(t)=k, t\in g)k^{\frac{2}{d-1}}dH(k(t)|t\in g)\\ &\Pr(\text{piv})^{\frac{2}{d-1}}c_2\end{aligned}$$

for a constant $c_2\neq 0$ that only depends on $\Pr(\alpha|\text{piv})$. Rewriting,

$$\begin{aligned}&\int_{t\in g} x(t)dH(t)\\ \approx\ &\Pr(t\in g)f(\Pr(\alpha|\text{piv})|t\in g)\mathrm{E}\Big[k(t)^{\frac{2}{d-1}}|t\in g, y(t)=\Pr(\alpha|\text{piv})\Big]\end{aligned}$$

Taking limits $n\to\infty$,

$$\int_{t\in g} x(t)dH(t)\approx W(g,\hat{p})\Pr(\text{piv})^{\frac{2}{d-1}}c_2,$$

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

E Proof of Lemma 5

Suppose that $\lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma_n^*)\neq\hat{p}$. Then, Lemma 8 implies $\lim_{n\to\infty}q(\omega;\sigma_n)\neq\frac{1}{2}$ for $\omega\in\{\alpha,\beta\}$. Then, (26) implies that the pivotal likelihood is exponentially small, which in turn implies that $x(t)$ is exponentially small for all t, given (19) and (17). Therefore, the difference in the vote shares is exponentially small, which implies $\left[q(\alpha;\sigma_n^*)-q(\beta;\sigma_n^*\right]s(\omega;\sigma_n^*)=0$ for $\omega\in\{\alpha,\beta\}$ since the standard deviation of the realized votes is of order $n^{\frac{1}{2}}$. Hence $\delta(\alpha)=\delta(\beta)$. Finally, Lemma 1 implies $\lim_{n\to\infty}\Pr(A|\alpha;\sigma_n^*)=\lim_{n\to\infty}\Pr(A|\beta;\sigma_n^*)$. But this contradicts with the assumption that $(\sigma_n^*)_{n\in\mathbb{N}}$ is an informative equilibrium sequence.

F Proof of Observation 2

Fix a voter and a state ω. The number of realized A-votes among the votes of the other citizens is the sum of $2n$ i.i.d. Bernoulli variables with mean $q(\omega;\sigma_n)$. Let $X_{k,n}=\mathcal{B}(1,q(\omega;\sigma_n))$ for any $1\le k\le 2n$ and $n\in\mathbb{N}$. Recall the assumption $\lim_{n\to\infty}q(\omega;\sigma_n)\in(0,1)$, and check that the conditions of Theorem 2 in Davis & McDonald (1995) are satisfied for $X_{k,n}$, $a_n=2nq(\omega;\sigma_n)$, and $b_n=\left[q(\omega;\sigma_n)(1-q(\omega;\sigma_n)\right]^{\frac{1}{2}})(2n)^{\frac{1}{2}}$. Note that $b_n\approx s(\omega;\mathbf{q}(\sigma_n))$. Further note that $\Pr(\text{piv}|\omega;\sigma_n,n)=\Pr(T_n=n)$ for $T_n=\sum_{i=1,\dots,2n}X_{k,n}$. Application of Theorem 2 in Davis & McDonald (1995) gives

$$\lim_{n\to\infty}\Pr(\text{piv}|\omega;\sigma_n,n)s(\omega,\sigma_n)=\phi(\delta(\omega)). \tag{109}$$

G Proof of Lemma 9

Proof. Recall $\lim_{x\to 0}\frac{c'(x)x}{c(x)}=d$. Since $k^{\frac{2}{d-1}}$ is strictly concave when $d>3$, an application of Jensen's inequality shows that for any g-intensity spread H' of H,

$$\mathrm{E}(k^{\frac{2}{d-1}}|t\in g;H')<\mathrm{E}(k^{\frac{2}{d-1}}|t\in g;H) \tag{110}$$

It follows from the definition of a g-intensity spread that for $g\neq g'\in\{L,C\}$,

$$\mathrm{E}(k^{\frac{2}{d-1}}|t\in g';H')=\mathrm{E}(k^{\frac{2}{d-1}}|t\in g';H). \tag{111}$$

Since H and H' satisfy (74), $\mathrm{E}(k^{\frac{2}{d-1}}|t\in g';H')=\mathrm{E}(k^{\frac{2}{d-1}}|t\in g',y(t)=\hat{p};H)$ and $\mathrm{E}(k^{\frac{2}{d-1}}|t\in g';H')=\mathrm{E}(k^{\frac{2}{d-1}}|t\in g',y(t)=\hat{p};H)$ for all $g'\in\{L,C\}$. Therefore (110), (111), the definition of $W(g)$ (see (45)) and the definition a g-intensity spread together imply (84) for $g=L$ and (85) for $g=C$, which finishes the proof of the lemma. ■

Theorem 6 *Let* $\lim_{x\to 0}\frac{c'(x)x}{c(x)} > 3$. *Let* $g \in \{L, C\}$. *Take any preference distribution* H *satisfying the genericity conditions and the independence conditions (73) - (75). When* M *is large enough, there is a* g*-intensity spread* H' *of* H *such that*

$$\lim_{n\to\infty} \Pr(z(\omega)|\omega;\sigma_n^*, n) \quad = \quad 0 \tag{112}$$

for all $\omega \in \{\alpha, \beta\}$, *where* $z(\omega)$ *is the policy preferred by the voter group* g *in* ω.

Consider the case $g = L$. Given Lemma 1 and Lemma 4, it remains to show that for any H there is an L-intensity spread H', so that

$$\frac{W_{H'}(L)}{W_{H'}(C)} < 1. \tag{113}$$

For this, it suffices to show that for any ϵ, we can choose H', so that

$$\mathrm{E}_{H'}(k(t)^{\frac{2}{d-1}}|t \in g) < \epsilon \tag{114}$$

since the genericity conditions ensure that $W_H(C) = W_{H'}(C) > 0$. Take L-intensity spreads $H'(\kappa)$ of H, so that

$$\begin{aligned}&\Pr(\{t : \kappa \le k(t) \le \kappa + \delta\}|t \in L; H'(\kappa))\\ &+\Pr(\{t : 0 \le k(t) \le \delta\}|t \in L; H'(\kappa)) \ge 1 - \delta\end{aligned} \tag{115}$$

for some $\kappa > 0$ and $\delta > 0$. Since the mean of the intensities is preserved under the L-intensity spread, the iterated law of expectation gives $\lim_{\delta\to 0}\Pr(\{t : \kappa \le k(t) \le \kappa + \delta\}|t \in L; H'(\kappa))\kappa = \mathrm{E}(k(t)|t \in L; H'(\kappa))$. Hence,

$$\begin{aligned}\lim_{\delta\to 0} \mathrm{E}(k(t)^{\frac{2}{d-1}}|t \in L; H'(\kappa)) &= \lim_{\delta\to 0} \Pr(\{t : \kappa \le k(t) \le \kappa + \delta\}|t \in L; H'(\kappa))\kappa^{\frac{2}{d-1}}\\ &= \frac{\mathrm{E}(k(t)|t \in L; H'(\kappa))}{\kappa}\kappa^{\frac{2}{d-1}} \overset{\kappa\to\infty}{\to} 0,\end{aligned} \tag{116}$$

where I used that $d > 3$ and hence $\frac{2}{d-1} < 1$. We conclude that for κ large enough and $\kappa < M$, we find an L-intensity spread of H, so that (113) holds. This finishes the proof for $g = L$. The proof for $g = C$ is analogous.

H Proof of Theorem 3

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, (56). Let

$$\mathrm{Q}_{\epsilon,n} = \{\mathbf{q} = (q(\alpha), q(\beta)) : |\mathbf{q} - (\frac{1}{2}, \frac{1}{2})| > \epsilon \text{ and } |q(\alpha) - q(\beta)| < \frac{1}{n^2}\} \quad (117)$$

We claim that when δ is small enough and n large enough, the best response is a self-map on $\mathrm{B}_{\delta,n}$,

$$\mathbf{q} \in \mathrm{Q}_{\epsilon,n} \Rightarrow \mathbf{q}(\sigma^{\mathbf{q}}) \in \mathrm{Q}_{\epsilon,n}. \quad (118)$$

The proof consists of three steps: Take $q \in \mathrm{Q}_{\epsilon,n}$. First, the vote shares in the two states are almost identical; in particular, the probability of a tie is also almost the same in the two states. Therefore, the pivotal event contains no information as $n \to \infty$,

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma^{\mathbf{q}}, n) = \Pr(\alpha), \quad (119)$$

To see why, recall that for any $\mathbf{q} \in \mathrm{Q}_{\epsilon,n}$, $q(\alpha) - q(\beta)| \leq \frac{1}{n^2}$. Recalling (90), this implies $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0$. Then, (119) follows from Lemma 12. Using Lemma 8, (119) implies $q(\sigma^{\mathbf{q}}) - (\frac{1}{2}, \frac{1}{2})| > \delta$ when ϵ is small enough and n large enough.

Second, the likelihood of the pivotal event is exponentially small, given (26). Thus, also $q(\alpha; \sigma^{\mathbf{q}}) - q(\beta; \sigma^{\mathbf{q}})$ is exponentially small, given Lemma 4.

Finally, an application of Kakutani's fixed point theorem shows that there is a sequence of equilibrium vote shares $(\mathbf{q}_n^*)_{n\in\mathbb{N}}$, that is, vote shares satisfying (56), and, given (119) and Lemma 8,

$$\lim_{n\to\infty} q_n^*(\omega) = \Phi(\Pr(\alpha)). \quad (120)$$

for all states ω. The theorem follows from the weak law of large numbers and $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$.

I Proof of Theorem 4

I.1 Third Item of Theorem 4

Take any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$. with $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n)$. Given (48), the order of the vote shares is pinned down by the order of the voter groups' power $W(g)$. Inspection of the cases shows that it is sufficient to show that when $W(L) > W(C)$, there is no equilibrium sequence for which, in both states, the outcome preferred by the contrarians is elected given the prior belief.

Case 1 $\Phi(\Pr(\alpha)) > \frac{1}{2}$.

Suppose $\lim_{n\to\infty} \Pr(A|\alpha; \sigma_n^*, n) = \lim_{n\to\infty} \Pr(A|\alpha; \sigma_n^*, n) = 0$. Hence, $q(\omega; \sigma_n^*) \leq \frac{1}{2}$ for n large. The order $W(L) > W(C)$ pins down the order of the vote shares, $q(\alpha; \sigma_n^*) > q(\beta; \sigma_n^*)$ for n large. Thus, $\Pr(\text{piv}|\alpha; \sigma_n^*, n) \geq \Pr(\text{piv}|\beta; \sigma_n^*, n)$ for n large enough. Since Φ is strictly increasing, $\lim_{n\to\infty} \Phi(\Pr(\text{piv}|\alpha; \sigma_n^*, n)) > \Phi(\Pr(\alpha))$. Lemma 8 implies $\lim_{n\to\infty} q(\omega; \sigma_n^*) > \frac{1}{2}$. The weak law of large numbers implies $\lim_{n\to\infty} \Pr(A|\alpha; \sigma_n^*, n) = \lim_{n\to\infty} \Pr(A|\alpha; \sigma_n^*, n) = 1$, contradicting the initial assumption.

Case 2 $\Phi(\Pr(\alpha)) < \frac{1}{2}$.

The proof is analogous to the case $\Phi(\Pr(\alpha)) > \frac{1}{2}$.

I.2 First Item of Theorem 4

Take any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$.

Case 1 $\lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n^*, n)) \neq \frac{1}{2}$

Given (26), the likelihood of the pivotal event is exponentially small. As a consequence, the difference of the vote shares $q(\alpha; \sigma_n^*) - q(\beta; \sigma_n^*)$ is exponentially small, given Lemma 4 and (51). This implies $\delta_n(\alpha) - \delta_n(\beta) \to 0$ (see the definition (87)). It follows from Lemma 12 that $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n) = \Pr(\alpha)$. Then, it follows from the weak law of large numbers that the equilibrium sequence satisfies (78). This was to be shown.

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

Recall that $\hat{p}$ is the unique belief with $\Phi(\hat{p}) = \frac{1}{2}$, thus $\hat{p} = \lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n)$. Recall the definition of $\delta_n(\omega)$, that is (90). We show that

$$\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0. \tag{121}$$

For this, first, we show

$$\lim_{n\to\infty} \Big[q(\alpha; \sigma_n) - q(\beta; \sigma_n)\Big] \Pr(\text{piv}|\sigma_n, n)^{-1} = 0. \tag{122}$$

if $d < 3$. To see why, note that

$$\begin{aligned} &\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t) \\ \approx \; &\Big[W(L) - W(C)\Big] \Pr(\text{piv}|\sigma_n^*, n)^{\frac{2}{d-1}} c_2, \end{aligned} \tag{123}$$

given Lemma 4. Using that the pivotal likelihood goes to zero as $n \to \infty$, (122) follows from (51), (123) and $d < 3$. Given Observation 2, the pivotal likelihood is of an order weakly smaller than $s(\omega; \sigma_n^*)^{-1}$. Hence, (122) implies $\lim_{n\to\infty} \Big[q(\alpha; \sigma_n) - q(\beta; \sigma_n)\Big] s(\omega; \sigma_n^*) = 0$, and thereby (121).

Now, Lemma 12 and (121) imply $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*) = \Pr(\alpha)$. However, this yields a contradiction to $\lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n^*) = \frac{1}{2}$ since $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$ by assumption. Hence, all equilibrium sequences satisfy the condition of Case (1), and we have already shown that this condition implies (78), which was to be shown.

I.3 Second Item of Theorem 4

This section uses a fixed point argument to show that there is a sequence of equilibrium vote shares $(\mathbf{q}_n^*)_{n\in\mathbb{N}}$ such that the corresponding sequence of equilibrium strategies satisfies (79). We provide the proof for the case when $\Phi(\Pr(\alpha)) < \frac{1}{2}$ and when the minority group has the higher power, $W(L) < W(C)$. The proof proceeds in two steps. First, we show that for any vote share $q(\alpha)$ in α close to $\frac{1}{2}$, we find a vote share $q_n^*(\beta)$ such that the best response to $\mathbf{q} = (q(\alpha), q_n^*(\beta))$ has again the same vote share in α.

Step 1 *Let $\Phi(\Pr(\alpha)) < \frac{1}{2}$ and $W(L) < W(C)$. For any $\epsilon > 0$ small enough,* *any* $\frac{1}{2} \leq q(\alpha) \leq \frac{1}{2} + \frac{\epsilon}{2}$*, and any* n *large enough, there is* $q_n^*(\beta) \geq \frac{1}{2}$ *such that*

$$q(\alpha) \quad = \quad q(\alpha; \sigma^{(q(\alpha), q_n^*(\beta))}). \tag{124}$$

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

Let $\mathbf{q} = (q(\alpha), q(\beta))$ in the following.

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

$$q(\alpha; \sigma^{\mathbf{q}}) > q(\alpha). \tag{125}$$

The election is more close to being tied in α, and, by Lemma 7, voters become convinced that the state is α, i.e., $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}, n) = 1$. It follows from Lemma 8 that $\lim_{n\to\infty} q(\alpha; \sigma^{\mathbf{q}}) = \Phi(1)$. Finally, (64) follows when ϵ is small enough since $\Phi(1) > \frac{1}{2}$.

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

$$q(\alpha; \sigma^{\mathbf{q}}) < q(\alpha). \tag{126}$$

The election is more close to being tied in β, and, by Lemma 6, voters update towards β, i.e. $\Pr(\alpha|\text{piv}; \mathbf{q}, n) \leq \Pr(\alpha)$. Since $\Phi(\Pr(\alpha)) < \frac{1}{2}$, Lemma 8 implies that $\lim_{n\to\infty} q(\alpha; \sigma^{\mathbf{q}}) < \frac{1}{2}$. Finally, (65) follows when ϵ is small enough.

Since $q(\alpha; \sigma^{\mathbf{q}})$ is continuous in $q(\beta)$, it follows from Step 1, Step 2, and the intermediate value theorem that, for n large enough, there is $q_n^*(\beta)$ such that (124) holds. It follows from the implicit function theorem that $q_n^*(\beta)$ is continuous in $q(\alpha)$.

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

$$q_n^*(\beta) = q(\beta; \sigma^{(q_n^*(\alpha), q_n^*(\beta))}). \tag{127}$$

Substep 1 *For* $q(\alpha) = \frac{1}{2}$*, and any* n *large enough,*

$$q(\beta; \sigma^{(q(\alpha), q_n^*(\beta))}) > q_n^*(\beta), \tag{128}$$

Recall that Φ is strictly increasing. Lemma 8 together with (124) implies

$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) = \hat{p} \in (0,1)$ for $\mathbf{q}_n = (\frac{1}{2}, q_n^*(\beta))$. We claim that

$$\delta(\beta)(\sigma^{\mathbf{q}_n}) = \lim_{n\to\infty} (q_n^*(\beta) - \frac{1}{2}) s(\beta; \sigma^{\mathbf{q}_n}) \in \mathbb{R}. \tag{129}$$

Otherwise, since $\delta(\alpha)(\sigma^{\mathbf{q}_n}) = \lim_{n\to\infty}(q(\alpha) - \frac{1}{2}) s(\beta; \sigma^{\mathbf{q}_n}) = 0$, Observation 2 implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) = 1$, which contradicts the earlier observation $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) \in (0,1)$. Recall (53); together with Observation 2 and $\delta(\omega)(\sigma^{\mathbf{q}_n}) \in \mathbb{R}$ for $\omega \in \{\alpha, \beta\}$,

$$\lim_{n\to\infty} \left[q(\beta; \sigma^{(q(\alpha), q_n^*(\beta))}) - q(\alpha; \sigma^{(q(\alpha), q_n^*(\beta))} \right] s(\beta; \sigma^{\mathbf{q}_n}) \in \{\infty, -\infty\}. \tag{130}$$

Since $q(\alpha; \sigma^{(q(\alpha), q_n^*(\beta))}) = \frac{1}{2}$, given (63), and since $q(\beta; \sigma^{(q(\alpha), q_n^*(\beta))} > q(\alpha; \sigma^{(q(\alpha), q_n^*(\beta))}$ for n large, given (48) and $W(L) < W(C)$, (129) and (130) together imply (128).

Substep 2 *For $q(\alpha) = \frac{1}{2} + \epsilon$, and any n large enough,*

$$q(\beta; \sigma^{(q(\alpha), q_n^*(\beta))}) < q_n^*(\beta), \tag{131}$$

Recall Lemma 8, which states $\lim_{n\to\infty} q(\omega; \sigma^{(q(\alpha), q^*(\beta))}) = \lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n, n))$. Given (124), $lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n, n)) = \frac{1}{2} + \epsilon$. Since Φ is strictly increasing, this implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n) > \Pr(\alpha)$, given that $\Phi(\Pr(\alpha)) < \frac{1}{2}$. Recalling Lemma 11, this implies that

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} \frac{q(\alpha)(1 - q(\alpha)}{(q(\beta)(1 - q(\beta))} \delta_n(\alpha; \mathbf{q}_n)^2 - \delta_n(\beta; \mathbf{q}_n)^2 \in (0,1) \tag{132}$$

Note that, in particular, this implies $q_n^*(\beta) \to \frac{1}{2} + \epsilon$. Now, we study the best response $\sigma^{\mathbf{q}_n}$. The pivotal likelihood given $\mathbf{q}_n$ is exponentially small since $\lim_{n\to\infty} q_n^*(\beta) = q(\alpha) = \frac{1}{2}$ and (26). Hence, given Lemma 4 and (51),

$$q(\alpha; \sigma^{\mathbf{q}_n}) - q(\beta; \sigma^{\mathbf{q}_n}) \leq y^n \tag{133}$$

for some $0 < y < 1$. This together with (132) implies (131) for n large enough.

Finally, using (128) and (131) and that $q(\beta; \sigma^{(q(\alpha), q_n^*(\beta))})$ is continuous in $q(\alpha)$, the intermediate value theorem implies Step 2.

It follows from Step 1 and Step 2 that for any n large enough, there is a pair of vote shares $q_n^*(\alpha)$ such that $\mathbf{q}_n^* = (q_n^*(\alpha), q_n^*(\beta))$ is a fixed point of $\mathbf{q}(\sigma^-)$. Moreover $q_n^*(\alpha) \leq \frac{1}{2} \leq q_n^*(\beta)$ by construction, implying that $\lim_{n\to\infty} \Pr(A|\alpha; \sigma^{\mathbf{q}}, n) \leq \frac{1}{2} \leq \lim_{n\to\infty} \Pr(A|\beta; \sigma^{\mathbf{q}}, n)$. Recalling from (49) that limit equilibrium outcomes are determinate when $d > 3$, this implies that the equilibrium sequence is informative. This concludes the proof of existence of informative equilibrium sequences when $W(C) > W(L)$ and $\Phi(\Pr(\alpha)) < \frac{1}{2}$. The proof for the other cases is analogous.

J Proof of Lemma 10

Case 1 $\lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n^*)) \neq \frac{1}{2}$.

Then, (26) implies that the likelihood of the pivotal event is exponentially small,

$$\Pr(\text{piv}|\alpha; \sigma_n^*) < z^n \tag{134}$$

for some $0 < z < 1$ and for n large enough. Hence, for all t,

$$x^*(t) < z^n c_2 \tag{135}$$

for some $c_2 \in \mathbb{R}$, given (19). Finally, this implies that $(2n+1)\mathrm{E}(c(x(t)) \to 0$ as $n \to \infty$ since $c(x)$ is approximately polynomial for x small enough, given (1). An application of the weak law of large number shows that the realized sum of the votes converges to 0 as $n \to \infty$.

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

Recall that $\Phi(\hat{p}) = \frac{1}{2}$. Fix $g \in \{\ell, s\}$. We use the notation $(y, k) = (y(t), k(t))$ for types $t \in g$, noting that (y, k) pin down the type uniquely. Let $\alpha = \arg\max\left(|q(\alpha; \sigma_n^*) - \frac{1}{2}|, |q(\alpha; \sigma_n^*) - \frac{1}{2}|\right)$. The other case will be analogous. First,

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha; \sigma_n^*)}{\Pr(\text{piv}|\beta; \sigma_n^*)} = \frac{\Pr(\beta)}{\Pr(\alpha)} \frac{\hat{p}}{1-\hat{p}}. \tag{136}$$

Multiplication of the first-order condition (18) by $n^{\frac{1}{2}}$ together with (136) yields

$$\begin{aligned} & n^{\frac{1}{2}} c'(x^*(\Pr(\alpha|\text{piv}; \sigma_n^*, n), 1)) \\ = \quad & n^{\frac{1}{2}} \Pr(\alpha) \Pr(\text{piv}|\alpha; \sigma_n^*, n) k(t) \left[(1 - y(t)) - y(t) \frac{\Pr(\text{piv}|\beta; \sigma_n^*, n)}{\Pr(\text{piv}|\alpha; \sigma_n^*, n)}\right]. \end{aligned} \tag{137}$$

Note that

$$\begin{aligned} 4^n(q(1-q))^n &= 4^n\Big[(\frac{1}{2}-(\frac{1}{2}-q))(\frac{1}{2}+(\frac{1}{2}-q)\Big]^n \\ &= 4^n(\frac{1}{4}-(\frac{1}{2}-q)^2)^n \\ &= (1-4\frac{(n^{\frac{1}{2}}(\frac{1}{2}-q))^2}{n})^n. \end{aligned} \tag{138}$$

for all $q \in (0,1)$. Combining (26) with (137) and (138) gives

$$n^{\frac{1}{2}}c'(x^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)) \approx c_2(1-4\frac{(n^{\frac{1}{2}}(\frac{1}{2}-q(\alpha;\sigma_n^*))^2}{n})^n$$

for some constant $c_2 > 0$. Multiplication of both sides with $\delta_n = n^{\frac{1}{2}}|q(\alpha;\sigma_n^*)-\frac{1}{2}|$ yields

$$\delta_n n^{\frac{1}{2}}c'(x^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)) \approx c_3\delta_n e^{-4\delta_n^2} + \delta_n\Big[(1-4\frac{\delta_n^2}{n})^n - e^{-4\delta_n^2}\Big]. \tag{139}$$

for some constant $c_3 > 0$. Using Lemmas 4.3 and 4.3 in Durrett (1991),

$$(1-4\frac{\delta_n^2}{n})^n - e^{-4\delta_n^2} \leq \frac{16\delta_n^4}{n^3}. \tag{140}$$

Therefore, $\lim_{n\to\infty}\delta_n\Big[(1-4\frac{\delta_n^2}{n})^n - e^{4\delta_n}\Big] = 0$. Since, given $d > 3$, all equilibrium sequences are determinate by (49), Observation 1 implies $\lim_{n\to\infty}\delta_n = \infty$, which in turn implies $\lim_{n\to\infty}\delta_n e^{-4\delta_n} = 0$. I conclude,

$$\lim_{n\to\infty}\delta_n n^{\frac{1}{2}}c'(x^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)) = 0. \tag{141}$$

Recall (51),

$$q(\alpha;\sigma_n^*) - q(\beta;\sigma_n^*) = 2\Big[\int_{t\in\ell} x(t)dH(t) - \int_{t\in s} x(t)dH(t)\Big]. \tag{142}$$

Recall (109) and (44), which imply $\int_{t\in g} x(t)dH(t) \approx c_4 x^*(\hat{p},1)^2 W(g)$ for some constant $c_4 \neq 0$. Hence,

$$x^*(\hat{p},1)^2 \approx c_5\Big[q(\alpha;\sigma_n^*) - q(\beta;\sigma_n^*)\Big]$$

for some constant $c_5 \neq 0$. Then,

$$\begin{aligned} x^*(\hat{p},1)^2 &\leq 2c_5\Big[|q(\alpha;\sigma_n^*) - \frac{1}{2}| + |q(\beta;\sigma_n^*) - \frac{1}{2}|\Big] \\ &\leq 4c_5 \frac{\delta_n}{n^{-\frac{1}{2}}}, \end{aligned} \tag{143}$$

where I used the triangle equality on the first inequality and $\alpha = \arg\max\left(|q(\alpha;\sigma_n^*) - \frac{1}{2}|, |q(\alpha;\sigma_n^*) -\right.$ for the second inequality. Hence, (141) implies

$$\lim_{n\to\infty} nx^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)^2 c'(x^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)) = 0 \tag{144}$$

Using (1),

$$\lim_{n\to 0} \frac{x^2 c'(x)}{xc(x)} = d. \tag{145}$$

Recall (27), hence $x^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1) \to 0$ as $n \to \infty$. Thus, combining (144) and (145),

$$\lim_{n\to\infty} nx^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)c(x^*(\Pr(\alpha|\text{piv};\sigma_n^*,n),1)) = 0. \tag{146}$$

We claim that any equilibrium sequence $(\sigma_n^*)_{n\in\mathbb{N}}$ satisfies

$$\int_{t\in g} c(x(t))dH(t) \approx \frac{2(d-1)}{d}\frac{c(x^*(p^*,1))x^*(p^*,1)}{\chi'(p^*)}W(g). \tag{147}$$

The proof follows from previous arguments: the proof is a verbatim to the calculations in section 4.2.2, except that we need to replace $x(\Pr(\alpha|\text{piv};\sigma_n^*,n),k)$ with $c(x(\Pr(\alpha|\text{piv};\sigma_n^*,n),k))$ at the appropriate places. Then, (146) and (147) imply

$$\lim_{n\to\infty}(2n+1)\Big[\int_{t\in[-1,1]^2} c(x(t))dH(t)\Big] = 0. \tag{148}$$

Finally, the lemma follows from the weak law of large numbers.

9 Additional Figures

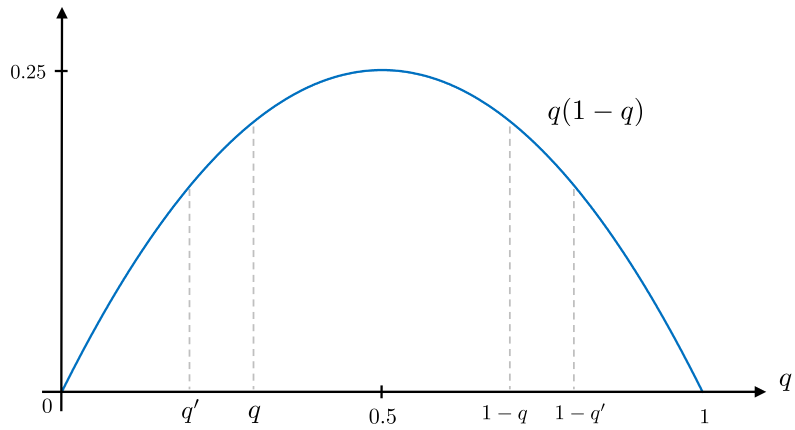

Figure 3: The function $q(1-q)$ for $q \in [0,1]$. If $|q - \frac{1}{2}| < |q' - \frac{1}{2}|$, then $q(1-q) > q'(1-q')$.

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.

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

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

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

Carpini, Michael X Delli, & Keeter, Scott. 1996. *What Americans know about politics and why it matters*. Yale University Press.

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

Davis, Burgess, & McDonald, David. 1995. An elementary proof of the local central limit theorem. *Journal of Theoretical Probability*, **8**(3), 693–702.

Downs, Anthony. 1957. An economic theory of political action in a democracy. *Journal of political economy*, **65**(2), 135–150.

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

Eguia, Jon X, & Xefteris, Dimitrios. 2018. Implementation by vote-buying mechanisms.

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. 2020. *Information Cost and Utilitarian Welfare in Elections*. Tech. rept. mimeo.

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

Henderson, Michael. 2014. Issue publics, campaigns, and political knowledge. *Political Behavior*, **36**(3), 631–657.

Hu, Li, & Li, Anqi. 2018. The Politics of Attention. *arXiv preprint arXiv:1810.11449*.

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.

Krosnick, Jon A. 1990. Government policy and citizen passion: A study of issue publics in contemporary America. *Political behavior*, **12**(1), 59–92.

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.

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

Matějka, Filip, & Tabellini, Guido. 2017. Electoral competition with rationally inattentive voters. *Available at SSRN 3070204*.

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

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

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

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