Information Frictions and Opposed Political Interests *

Carl Heese †

September 12, 2022

This paper studies collective choices with information frictions. In a majority election, voters can acquire private information about policy consequences before voting though this requires costly effort. Information frictions alter the power relationships between opposed political interests by turning the election into an informational contest: There is an equilibrium in which the policy preferred by the interest group with the higher aggregate information acquisition effort is elected; outcomes therein represent voters with a minority interest if, they have comparably high utilities. Information advantages and internal conflicts of opinion matter: we characterize how information cost and the dispersion of priors modulate the influence of an interest group.

In many collective choices, there are information frictions. It is costly to pay attention to, filter, and process all of the relevant information. In particular, when there is uncertainty about who benefits and who loses from a given choice, voters often engage in costly activities in order to cast an informed vote, and thereby advance their interests. For example, in general elections, millions of citizens watch the presidential debates and make use of information websites that provide information about the candidates' positions.[1] Members of administrative committees—such as legislative committees, and hiring committees—invest substantial time and effort in evaluating policy positions and candidates.

---

*I am grateful for helpful discussions with Johannes Hörner, Daniel Krähmer, Stephan Lauermann, Philipp Strack, Karl Schlag, 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, the Young Economist Symposium 2020, and the World Congress of the Game Theory Society 2021. 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 Vienna, Department of Economics, Oskar-Morgenstern-Platz 1, 1020 Vienna, Austria; carl.heese@univie.ac.at

This paper asks if (and how) such information frictions and informational efforts affect whose interests find representation through the collective choice. We employ a model of a simple majority election.

Our baseline model modifies the canonical voting setting by Feddersen and Pesendorfer (1997) to include information frictions. We follow the standard approach to modeling information frictions, as in Martinelli (2006). There is a simple majority election with two policies, A and B. Before the election, voters choose the precision of a private, binary signal about a pay-off relevant state, α or β. An uninformative signal is costless and a more precise signal is more costly. Voter types may differ in their state-dependent preferences in a general way. In the leading scenario, in expectation, a majority of the voters prefers A only in α, and a minority prefers A only in β. That is, there are two *interest groups* that favor opposite policies in both states.

Such preferences arise naturally in situations in which voters are uncertain about who benefits and loses from a given collective choice. For instance, consider elections in which two candidates compete that are by and large centrists and voters are uncertain if candidate A is more left than candidate B (state α) or more right (state β).[2] Leftists prefer A only in α and rightists only in β. Similarly, "younger" parties in parliamentary systems are often not clearly positioned on the left-right spectrum.[3] Further examples that have been discussed in the literature include referenda on distributive reforms such as free-trade agreements, and open primaries with "crossover" voters of the opposing party.[4]

The seminal finding in the setting of Feddersen and Pesendorfer (1997) without information frictions is that the outcome in *all* equilibria of a large election is "as if" the state is known. Feddersen and Pesendorfer (1997) and later Bhattacharya (2013) have shown this full-information equivalence result for the broad class of "monotone preferences" and when citizens receive *exogenous* noisy information about the state in the form of conditionally i.i.d. signals. Full-information equivalence means that the majority-preferred outcomes are elected state-by-state.

[1]Popular sites from the 2020 US elections include `https://www.isidewith.com/elections/` and `https://2020election.procon.org/2020-election-quiz.php`. An example from Europe (Germany) is `https://www.bpb.de/politik/wahlen/wahl-o-mat/`.

[2]Such uncertainty may arise as a strategic choice of the candidates (Kartik *et al.*, 2017).

[3]This is because they do not originate around issues of the traditional left-right divide but around other topics. Two anecdotes: the green party of Canada once campaigned with the slogan "not left, not right, forward together", and the green party of Germany refused to being seated either on the left side or the right side of the parliament.

[4]See, e.g., Fernandez and Rodrik (1991); Meirowitz *et al.* (2006); Kim and Fey (2007); Bhattacharya (2013); Ali *et al.* (2018).

The main insight of this paper is that information frictions fundamentally alter the power relationships between political interests, by turning the election into an informational contest. Precisely, in a large election, there is a "tug-of-war"-equilibrium in which the policy preferred by the interest group with the higher aggregate precision of the signals is elected. The equilibrium is robust; that is, it exists for all prior beliefs. This result may be surprising: One may think that a group of voters with a majority interest could always exert its dominance since the one-person-one-vote principle grants them more formal voting power (and there is no voting cost). For this result, it matters that voters are uncertain about which policy benefits them. Without such uncertainty, the majority voters would coordinate perfectly on voting for the policy that benefits them and enforce this policy as the outcome.

The contest-like structure of the equilibrium yields several insights about *how* information frictions shape the power relationships between opposed political interests. We develop these insights within the baseline model and within a generalized version that allows for heterogeneity in the prior beliefs and information cost (Section 5). First, power shifts into the direction of the voters with high preference intensities. Specifically, outcomes represent the interests of a minority of the voters in all states if they have comparably high intensities. This speaks to the concern that majority elections may not be able to reflect preference intensities and may always lead to majority-preferred outcomes, even when this would entail large losses in social welfare.[5] Second, the information frictions shift power into the direction of the voters with comparably low information cost. If voters of an interest group face sufficiently low cost, ceteris paribus, the group's preferred policy is elected in the tug-of-war equilibrium. In light of the mentioned applications, this shows that there are "insider advantages" in collective committee decisions and that information websites in general elections may play a crucial role in "leveling the playing field". Third, for a certain class of type distributions, we illustrate that information frictions shift power into the direction of the interest group with more homogeneous prior beliefs. If voters of an interest group have sufficiently dispersed prior beliefs, ceteris paribus, the group's preferred policy is not elected in the tug-of-war equilibrium. This re-

[5] Several other streams of the literature have provided complementary arguments as to how elections can reflect intensities; see, for example, the literature on turnout and voting cost (see, for example, Palfrey and Rosenthal, 1985; Ledyard, 1984; Krishna and Morgan, 2011, 2015). See also the literature on public good provision; for example, Ledyard and Palfrey (2002) show that public good provision through simple majority voting schemes is approximately utilitarian efficient when there are many agents.

sult matches the intuition that political groups with (more) internal conflicts of opinion are less powerful. All of these insights are implications of a deeper structural result that characterizes which interest group dominates the informational contest. Namely, we provide an explicit formula for the ratio of the aggregate precision of the minority and the majority types as the electorate grows large, in terms of primitives.

Another insight of this paper is that the information frictions create strategic complementarities. While it is well-known that the possibility to "free-ride" on others reduces the incentives of the voters to acquire costly information, the observation of complementarities is novel. The logic is as follows: How severely the incentives to acquire information are impacted by the free-riding motive depends on the voters' expectations about the closeness of the election. The closer the voters expect the election to be, the more likely they believe their individual votes to affect the outcome. Then, information about policy consequences is more valuable to the voters. Critically, when the other citizens acquire information and vote in an informed manner, this may change a voter's belief about the closeness of the election. We describe how this may spur the information acquisition of the given voter; that is, information acquisition can be complementary.

The complementarities "modulate" the competitive forces of endogenous information acquisition. They act in such a way that there are three equilibria in a large election, ordered by the aggregate precision (or "effort") of the voter types.[6] In comparison to the tug-of-war-equilibrium—which is effort-maximal—outcomes are less strongly shaped by the competitive forces in the other equilibria. In the low effort equilibrium, outcomes are given by the prior beliefs and are "as if" the cost of any information acquisition is infinite. In the medium effort equilibrium, outcomes depend both on the prior beliefs and on which interest group's voters acquire a higher precision, in the aggregate.

The paper contributes to the understanding of the competition between opposed political interests. Much of the literature on interest groups has assumed that groups act as perfectly coordinated entities.[7] In our model, individual members of an interest group maximize their individual interest, taking as given the behaviour of others with shared interests. This way, our model reflects the competition between decentralized political interests such as in elections. It provides a game-theoretic analysis of how the coordination of behaviour within groups

[6]We use the terms "precision" and "effort" interchangeably.

[7]See, e.g., page 95ff in the review of Grossman and Helpman (2001).

of agents with ex-post aligned interest is shaped by differences in prior beliefs and preference parameters. These insights may be useful beyond political and election settings in the same way that the analysis of auctions provides insights useful to more general price setting environments.[8]

This paper is one of the first to study how the competition between political interests channels through costly information acquisition efforts. Previous work has analyzed a large variety of other factors and forms of political competition (see Grossman and Helpman, 2001). Complementary to our paper is work from the literature on electoral competition. There the question is, how the electoral competition between politicians is affected by the voters' limited attention to politics (Matějka and Tabellini, 2021; Yuksel, 2021).[9] By contrast, we focus on the competition between interest groups of voters. Our analysis is driven by the strategic interdependencies of the voters' behavior and the large heterogeneity of the voters, including differences in prior beliefs and cost. These features are not present in the electoral competition models. In Appendix J, we discuss a central result from Matějka and Tabellini (2021) and explain how predictions differ in our model.

A central question in the literature is if and how the competition of political interests affects the welfare properties of policy outcomes. To this end, it has been shown in settings with participation cost that turnout may adjust endogenously so that outcomes in large elections maximize utilitarian welfare (see, e.g., Krishna and Morgan, 2011, 2015). Recently, this question has sparked the interest in novel democratic mechanisms that endogenously distribute political power. Examples include "quadratic voting", in which quadratic prices are attached to voting rights. Such mechanisms garner increasing attention by organisations and political parties (see, e.g., Hardt and Lopes, 2015; Blum and Zuber, 2016), and have been shown to exhibit desirable welfare properties (Lalley and Weyl, 2018; Eguia and Xefteris, 2018).

Potentially surprisingly, our analysis uncovers parallels between costly information acquisition on the one hand and costly participation (Palfrey and Rosenthal, 1985; Krishna and Morgan, 2011, 2015) and vote-buying (Lalley and Weyl, 2018; Eguia and Xefteris, 2021) on the other hand. our results show that the effects of costly information acquisition are similar in spirit to the effects of costly participation or vote-buying. Specifically, when it is costly to acquire informa-

[8] See the literature on information aggregation in auctions (Wilson, 1977; Milgrom, 1981; Pesendorfer and Swinkels, 1997).

[9] See also the work in Grossman and Helpman (2001) on how differential *exogenous* knowledge of citizens affects the electoral competition.

tion, there is an equilibrium in which the outcomes reflect preference intensities in a way so that a minority preference is elected if the minority has comparably high utilities at stake. A priori, the economic forces of information and participation cost seem not alike. In particular, the mentioned welfare results are driven by how the distribution of the voting rights or who exercises them forms endogenously. However, information cost do not alter this distribution. Concretely, in our setting each citizen has one vote and voting is mandatory. A further parallel is that the election is like a contest in some equilibrium: the interest group with the higher aggregate informational effort "wins". In the settings with costly participation or vote-buying, the competition has a similar but exogenously given contest structure: There is an exogenous mapping from action profiles to outcomes, and the group with the higher aggregate costly action wins.

The rest of the paper is structured as follows: Section 1 illustrates central ideas and the tug-of-war equilibrium with an example. Section 2 presents the model. Section 3 contains preliminaries; in particular, a detailed analysis of the best response for the baseline setting with a common cost type and prior belief about the state. Section 4 presents the main result for the baseline setting, and Section 5 for the generalized setting. Section 6 discusses the relation to the literature on information aggregation in elections, including Bhattacharya (2013), and previous work with costly information (e.g., Martinelli, 2006, 2007; Triossi, 2013).

1 Example

There are $2n+1 \geq 3$ voters (or citizens). With probability $1 > \lambda > \frac{1}{2}$, a voter is *aligned* and prefers a reform A over the status quo B in α and B over A in β. With probability $1-\lambda$, a voter is *contrarian* and prefers A in β and B in α.

Aligned and contrarian voters are of three types: an "unbiased" type, a "reform leaning" type, and a "reform skeptical" type. These types differ in their willingness-to-pay for being able to change the outcome in a given state. Consider the aligned. The unbiased types have a willingness-to-pay of $2k_g$ to change the outcome in any state, for some $k_g > 0$. The reform-leaning aligned are willing to pay more to change the outcome in the state α in which they prefer the reform ($3k_g$), and less in β (k_g). Conversely, the reform-skeptical aligned are willing to pay less to change the outcome in the state β in which they prefer the reform (k_g), and more in α ($3k_g$). Conditional on being an aligned type, each voter is equally likely to be a reform-leaning or reform-skeptical. For the contrarian types, the analogous statements hold, where we switch the role of α and β. The voters hold a common, uniform prior about the state. Each voter receives a private, binary signal $s \in \{a, b\}$ about the state. Types are denoted by t and drawn from a commonly known distribution H, independently across voters and independently of the signals.

The timing is as follows: The voter types realize. Each voter chooses the precision $x \in [0, \frac{1}{2}]$ of her signal, that is $\frac{1}{2} + x = \Pr(a|\alpha) = \Pr(b|\beta)$. When choosing precision x, the voter bears a cost $c(x) = \frac{x^d}{d}$, with $d > 1$ so that $c'(0) = 0$. The state and private signals realize. After observing the private signals, all citizens vote simultaneously. Finally, the outcome is decided by simple majority rule.

The example is deliberately symmetric across types and states. Hence, it is immediate to show that there are strategy profiles in which the vote shares and outcomes are symmetric across states; that is, the expected vote share of A in α equals that of B in β. In particular, given such a symmetric strategy profile, the probability that a given citizen's vote affects the election outcome has the same likelihood in both states.[10] So, if a type votes A, she expects to tip the election outcome from B to A with the same probability in both states. Doing so benefits her in one state and comes with a utility loss in the other. Similarly, voting B tips the election outcome from A to B with the same probability in both states. In one state, the type gains from tipping the outcome from B to A, and in the other she loses from it. What matters for the voter‘s decision is the utility (willingness-to-pay) that she attaches to these two events.

The following illustrates two points that will also be central in the later analysis of the general model. First, each interest group faces internal conflicts of opinion; different types of the same interest group vote for opposed policies in equilibrium. Second, how "well" the voters of a group coordinate on voting for their preferred policy is determined by their informational efforts. This renders the election an "informational contest".

Conflicts of opinion. Take, for example, the aligned. For the reform-skeptical types, the utility gain from tipping the outcome from the reform A to the status quo B in β is higher than the utility loss from doing so in α. Given the uniform

[10] A single citizen's vote is decisive for the election outcome only in the event in which the votes of the other citizens split into n votes for A and n votes for B.

prior about the state, the reform-skeptical types thus strictly prefer to vote for the status quo without additional information about the state. Analogously, the reform-leaning types strictly prefer to vote for the reform without additional information. Any signal that could turn around the strict preference of these types would have to have sufficiently high precision $x > \bar{x}$ for some $\bar{x} > 0$. When the electorate size $2n + 1$ is large, the benefit from more information is small because a single citizen expects that her vote affects the outcome only with a probability close to zero. So, benefits do not outweigh the cost of a precision $x > \bar{x}$. Any informative signal with a smaller precision does not affect the type's voting decision and is also not worth the cost. In total, the reform-leaning and reform-skeptical types choose to receive an uninformative signal; that is, $x = 0$. Finally, since the reform-leaning and reform-skeptical types are equally likely, their votes split $50 - 50$ between both policies in expectation, effectively canceling each other out.

Coordination through information. In contrast to the reform-leaning and reform-skeptical types, for any unbiased type t, the symmetry of the willingness-to-pay will imply that it is optimal to choose a non-zero precision $x(t) > 0$.[11] Any unbiased type t receives the "correct" signal with probability $\frac{1}{2} + x(t)$ and follows it, voting for the preferred policy with probability $\frac{1}{2} + x(t)$ in each state.

Since the votes of the reform-leaning and reform-skeptical cancel out each other, the difference in the expected vote shares of A and B is driven entirely by the informational efforts of the unbiased. Aggregating the behavior of all types, in each state, the expected vote shares differ by

$$\int_{t aligned} x(t)dH(t) - \int_{t contrarian} x(t)dH(t) \tag{1}$$

where $x(t)$ is the optimal precision chosen by a type t. Thus, the policy that is preferred by the interest group with the higher aggregate precision receives more votes in expectation. The election resembles an informational contest.

The endogenous precision choices naturally vary with the preference intensity k_C and the information cost, parametrized by d. Figure 1 illustrates how

[11]The symmetry of the willingness-to-pay and the symmetry of the prior imply that the unbiased type is indifferent between voting A and B without further information. A simple calculation shows that the benefit of choosing a precision x is positive and linear in x, given the indifference. Hence, for sufficiently small precision levels x, the marginal benefit is a positive constant and outweighs the marginal cost $c'(x) \approx 0$.

varying these parameters translates into equilibrium outcomes.[12] For the other parameters, we make a fixed choice. Going down the rows, the intensity k_C of the contrarians increases, as does the likelihood of their preferred policy being elected. Comparing the column for $d = 2$ and for $d = 3$, we see that the contrarians dominate the election with high intensities $k_C = 4$ when $d = 2$, but not when $d = 3$. This illustrates how intensities matter more when information is less "cheap"—as measured by the cost elasticity d.

k_C	$d = 2$	$d = 3$
0	0.79	0.94
1	0.65	0.86
2	0.5	0.77
3	0.35	0.75
4	0.21	0.75

Figure 1: This shows the likelihood of outcome A in α and B in β in equilibrium for different cost elasticities d and intensities k_C. We fix $2n + 1 = 31$, $k_L = 1$, $\lambda = \frac{1}{3}$, and the likelihood of the unbiased type to be 1 for both the aligned and contrarians.[13]

2 Model

The model generalizes the example from Section 1 by allowing for general type distributions. Besides that, the voting game is as per the example.

A voter type $t = (v, r, t_\alpha, t_\beta)$ is given by a prior belief, specifying the subjective likelihood $q \in (0, 1)$ of the state being α is, a cost type $r > 0$, and a preference type (t_α, t_β), where $t_\omega \in \mathbb{R}$ 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 and are drawn independently from a commonly known cumulative distribution function $H : [0, 1] \times \mathbb{R}_{>0} \times \mathbb{R}^2 \to [0, 1]$. A voter's type is her private information.

A strategy $\sigma = (x, \mu)$ of a voter consists of a function $x : [0, 1] \times \mathbb{R}_{>0} \times \mathbb{R}^2 \to [0, \frac{1}{2}]$ mapping types to signal precisions and of a function $\mu : [0, 1] \times \mathbb{R}_{>0} \times \mathbb{R}^2 \times \{a, b\} \to [0, 1]$ mapping types and signals to probabilities to vote A, that is, $\mu(t, s)$ is the probability that a voter of type t with signal s votes for A. We only

[12] There is a unique (non-trivial) Bayes-Nash equilibrium in this example. The uniqueness is driven by the symmetry between the reform-leaning and reform-skeptical types. Generically, there are multiple equilibria given the relevant conditions on the cost function.

consider non-degenerate strategies.[14] We analyze the Bayes-Nash equilibria of the Bayesian game of voters in symmetric strategies, henceforth called *equilibria*.

When choosing precision x, a voter with cost type r bears a cost $c(x) = \frac{r}{d}x^d$ for some $d > 0$. The cost type captures idiosyncratic differences. The parameter d is the elasticity of the cost function. We think of the elasticity of the cost function as varying the regime of how costly information is (up to idiosyncratic differences captured by r), where a higher d means that information of low precision is "cheaper".[15]

3 Baseline setting

For the main part of the analysis, we consider the setting in which all citizens share a common prior belief type $v = \Pr(\alpha) \in (0, 1)$ and a common cost type $r = 1$. This makes results particularly comparable to existing work and isolates the effects of the heterogeneity in preference intensities. In Section 5, we turn to the general setting.

Preference types. Slightly abusing the notation, we denote by H the distribution of (t_α, t_β). We assume in the following that H has a continuous density on its support. The support is the Cartesian product of $K_\alpha \subseteq \mathbb{R}$ and $K_\beta \subseteq \mathbb{R}$, which are connected, compact and contain 0 in their interior. Figure 2 shows the area of the 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*). All of the analysis also goes through when all voters share common interests; for example, when all types in the support are aligned, $K_\alpha \times K_\beta \subseteq \mathbb{R}_{\geq 0} \times \mathbb{R}_{\leq 0}$.

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.[16]

[14]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 $x(t) = 0$ for all t are trivial equilibria.

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

[16]In fact, for any non-degenerate strategy, the likelihood of the pivotal event is non-zero (see Section 3.1.1) such that not acquiring any information and voting for the preferred policy is the *unique* best response for all partisans.

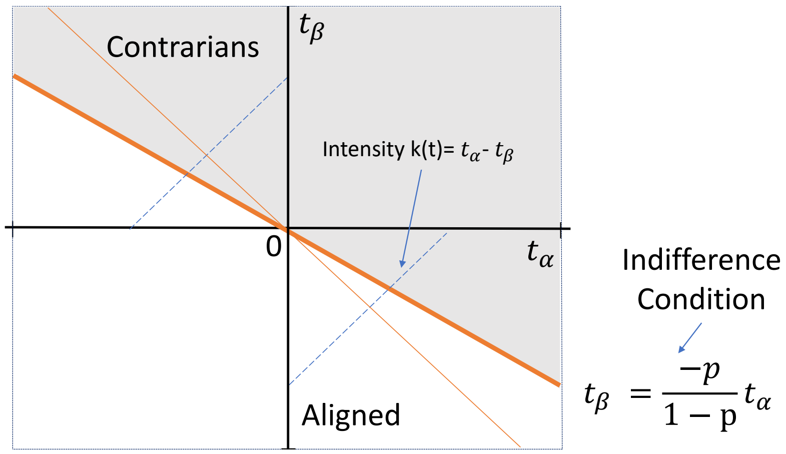


Figure 2: For any given belief $p = \Pr(\alpha) \in (0,1)$, the set of types t that are indifferent given p is given by $t_\beta = \frac{-p}{1-p} t_\alpha$. Voter types north-east of the indifference line (shaded area) prefer A given p. Contrarian and aligned types are uniquely identified by their *(total) intensity* $k(t) = |t_\alpha - t_\beta|$ (dashed lines) and their *threshold of doubt* $y(t) = \frac{-t_\beta}{t_\alpha - t_\beta}$ (straight lines).

Monotone preferences. A central object of the analysis is the *aggregate preference function*

$$\Psi(p) = \Pr_H(\{t : p \cdot t_\alpha + (1-p) \cdot t_\beta \geq 0\}), \tag{2}$$

which maps a belief $p \in [0,1]$ about the state to the probability that a random type t prefers A given p. Figure 2 illustrates Ψ. The (bold straight) 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 $\Psi(p)$. 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, Ψ is continuously differentiable in p.

We assume that

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

such that the median-voter preferred outcome is A in α and B in β. In particular, this excludes the (trivial) cases when there is a majority of partisans for one policy in expectation. We also assume that Ψ is strictly monotone.[17] The non-monotone

[17]The monotone case is the case for which the literature has established that equilibrium outcomes are full-information equivalent when information of the citizens is exogenous and conditionally i.i.d. (see Bhattacharya, 2013).

case is discussed in Appendix I. Henceforth, I will call distributions H for which Ψ is strictly increasing and satisfies (3) *monotone* preference distributions. The set of the aligned types is $L = \{t : t_\alpha > 0, t_\beta < 0\}$ and the set of the contrarian types is $C = \{t : t_\alpha < 0, t_\beta > 0\}$. Throughout, I use $g \in \{L, C\}$ as the generic symbol for a voter group, aligned or contrarians.

Threshold of doubt and total intensity. For the aligned and contrarians, 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*,

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

We call $y(t)$ the *threshold of doubt.* As Figure 2 illustrates, for any aligned type t, $y(t)$ and $k(t)$ together uniquely pin down t. Formally, $-y(t)k(t) = t_\beta$, and $(1 - y(t))k(t) = t_\alpha$. Similarly, for any contrarian type t, $y(t)$ and $k(t)$ together uniquely pin down t.

3.1 Best response

3.1.1 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

$$q\left(\alpha; \sigma\right) = \int_{t \in K_\alpha \times K_\beta} (\frac{1}{2} + x(t))\mu(t, a) + (\frac{1}{2} - x(t))\mu(t, b) dHt,$$

and

$$q\left(\beta; \sigma\right) = \int_{t \in K_\alpha \times K_\beta} (\frac{1}{2} - x(t))\mu(t, a) + (\frac{1}{2} + x(t))\mu(t, b) dHt.$$

We 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 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

$$\Pr\left(\text{piv}|\omega;\sigma',n\right)=\binom{2n}{n}\left(q\left(\omega;\sigma'\right)\right)^n\left(1-q\left(\omega';\sigma\right)\right)^n. \tag{6}$$

since conditional on the state the type and the signal of any 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 $x(t)$,

$$\Pr(\alpha|s,\text{piv};\sigma',n)\cdot t_\alpha+(1-\Pr(\alpha|s,\text{piv};\sigma',n))\cdot t_\beta>0\Rightarrow\mu\left(s,t\right)=1, \tag{7}$$

$$\Pr(\alpha|s,\text{piv};\sigma',n)\cdot t_\alpha+(1-\Pr(\alpha|s,\text{piv};\sigma',n))\cdot t_\beta<0\Rightarrow\mu\left(s,t\right)=0. \tag{8}$$

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 $t\in L$, (7) and (8) are equivalent to

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

$$\Pr(\alpha|s,\text{piv};\sigma',n)<y(t)\Rightarrow\mu(t,s)=0. \tag{10}$$

For all contrarian types $t\in C$, (7) and (8) are equivalent to

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

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

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.

3.1.2 Total intensity pins down signal precision

What is the marginal value of information to a citizen? Take an aligned voter, and fix the likelihood $x>0$ 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 vote. In the pivotal event, when she chooses precision x, her expected utility from the elected policy is

$$\Pr(\text{piv}|\sigma', n) \Pr(\alpha|\text{piv}; \sigma)(\frac{1}{2} + x)t_\alpha \tag{13}$$

in state α, and

$$\Pr(\text{piv}|\sigma', n) \Pr(\beta|\text{piv}; \sigma)(\frac{1}{2} - x)t_\beta \tag{14}$$

in state β. Here, we used Lemma 1 and that the utility from B is normalized to zero.[18] Therefore, summing (13) and (14) and taking the derivative, the marginal benefit of a higher precision x is

$$\begin{aligned} & MB\Big[t; \sigma', n\Big] && (15)\\ = \; & \Pr(\text{piv}|\sigma', n)(\Pr(\alpha|\text{piv}; \sigma)t_\alpha - \Pr(\beta|\text{piv}; \sigma)t_\beta) \\ = \; & \Pr(\text{piv}|\sigma', n)k(t)e(y(t)) \end{aligned}$$

for $e(y(t)) = \Pr(\alpha|\text{piv}; \sigma)(1 - y(t)) + \Pr(\beta|\text{piv}; \sigma)y(t)$. Here, 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, the precision is pinned down by equating marginal benefits and marginal cost,

$$c'(x) = MB\Big[t; \sigma', n\Big]. \tag{16}$$

when $d > 1$ and n is large enough. No type acquires full information ($x = \frac{1}{2}$). This is because $d > 1$ implies $c'(0) = 0$ and when n is sufficiently large, the pivotal likelihood is small enough so that (16) has an interior solution $x < \frac{1}{2}$ for all types, given the compactness of the type space. In the following, we maintain the standing assumption that $d > 1$ and that n is large enough so that (16) has an interior solution. The unique solution to (16) for t is

$$x^*(t; \sigma', n) = MB\Big[t; \sigma', n\Big]^{\frac{1}{d-1}}. \tag{17}$$

[18]Similarly, in the pivotal event, a contrarian's expected utility when choosing x is $\Pr(\text{piv}; \sigma', n) \Pr(\alpha|\text{piv}; \sigma)(\frac{1}{2} - x)t_\alpha$ in state α, and $\Pr(\text{piv}; \sigma', n) \Pr(\beta|\text{piv}; \sigma)(\frac{1}{2} + x)t_\beta$ in state β.

Here, we used that $c(x) = \frac{r}{d}x^d$ and that $r = 1$.

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

$$\forall t \in L : x(t) > 0 \Rightarrow \mu(t, a) = 1 \ \textit{and} \ \mu(t, b) = 0, \tag{18}$$

$$\forall t \in C : x(t) > 0 \Rightarrow \mu(t, a) = 0 \ \textit{and} \ \mu(t, b) = 1. \tag{19}$$

The proof is in the Appendix A.

3.1.3 Who acquires additional information?

The types t with $y(t) = \Pr(\alpha|\text{piv}; \sigma', n)$ are indifferent between A and B without further information, given (9) - (12), and called the *marginal types*. Lemma 2 shows that, for each total intensity $k = k(t)$, only types in a certain interval around the marginal types acquire information, as illustrated in Figure 6 in Appendix D.

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

$$x(t) > 0 \quad \Rightarrow \quad y(t) \in [y_g^-(k), y_g^+(k)], \tag{20}$$

$$y(t) \notin [y_g^-(k), y_g^+(k)] \quad \Rightarrow \quad x(t) = 0. \tag{21}$$

Note that, for $d \leq 1$, the marginal cost are bounded away from zero, $c'(0) > 0$. Thus, (16) has no solution when n is large and all types stay uninformed.

To get more intuition for the result, take, for example, the aligned boundary type t with $y(t) = y_L^-(k)$. This type is indifferent between voting A without additional information and choosing the precision $x = x^*(t; \sigma', n)$, as in (17). We show that the type's indifference condition can be rewritten as

$$\chi(y(t)) + \frac{1}{2} = \frac{(d-1)}{d} x^*(t; \sigma', n) \tag{22}$$

where $\chi(y) = \frac{-\Pr(\beta|\text{piv};\sigma,n)y(t)}{\Pr(\alpha|\text{piv};\sigma,n)(1-y(t)-\Pr(\beta|\text{piv};\sigma,n)y(t)}$. Details of the algebra are in Appendix E.

The function on the right hand side converges uniformly to zero as $n \to \infty$. This is because $x^*(t; \sigma, n)$ is proportional to the pivotal likelihood, given

(15) and (17). The left hand side captures the bias towards policy A without additional information. The bias is zero at the indifferent type's threshold $\bar{y} = \Pr(\alpha|\text{piv};\sigma,n)$, as a simple calculation verifies. Intuitively, for an aligned type, the lower the threshold of doubt $y(t)$, the higher the bias towards policy A. In fact, the left hand side strictly increases in the distance of $y(t) < \bar{y}$ to $\bar{y}$.[19] Altogether, we see that the left hand side crosses the right hand side exactly once when n is large. Thus, the indifference equation has a unique solution $y_L^-(k) < \bar{y}$.

The following studies the equilibria of the election as the number of citizens $2n+1$ grows without bound. Considering a large number of citizens allows for a precise analysis.

3.2 Informative equilibrium sequences

3.2.1 Informativeness

For any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$ and any n, let

$$\delta_n(\omega;\sigma_n) = \frac{q(\omega;\sigma_n) - \frac{n}{2n+1}}{s(\omega;\sigma_n)}. \tag{23}$$

This measures the distance between the expected vote share and the majority threshold in multiples of the standard deviation $s(\omega;\sigma_n)$ of the vote share distribution for $\omega \in \{\alpha,\beta\}$, where $s(\omega;\sigma_n)^{-1} = \sqrt{\frac{(2n+1)}{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}}$.[20] Figure 3 illustrates a normal approximation of the distribution of the number of A-votes. This approximation shows that, as $n\to\infty$, the probability that A gets elected in ω converges to[21]

$$\lim_{n\to\infty} \Pr(A|\omega;\sigma_n) = \lim_{n\to\infty} 1 - \Phi(-\delta_n(\omega;\sigma_n)). \tag{24}$$

[19] For the calculation of the derivative, see Appendix E. The condition $\lim_{n\to\infty}\Pr(\alpha|\text{piv};\sigma',n) \in (0,1)$ of the lemma ensures that the derivative stays bounded away from zero as $n\to\infty$.

[20] Let $q_n = q(\omega;\sigma_n)$. The number v_n of A-votes follows a Binomial distribution with variance $(2n+1)q_n(1-q_n)$. So, the vote share $\frac{v_n}{2n+1}$ of A follows a distribution with standard deviation $s(\omega;\sigma_n)$.

[21] Let $q_n = q(\omega;\sigma_n)$. Take the normal approximation $\mathcal{B}(2n+1,q_n) \simeq \mathcal{N}((2n+1)q_n,(2n+1)q_n(1-q_n))$ of the distribution of the number of A-votes. It shows that the probability that there are more A-votes than B-votes converges to $\lim_{n\to\infty} 1-\Phi(\frac{(2n+1)(\frac{n}{2n+1}-q_n)}{((2n+1)q_n(1-q_n))^{\frac{1}{2}}}) = \lim_{n\to\infty} 1-\Phi(-\delta_n(\omega;\sigma_n))$. Note that we are applying the Lindeberg-Feller version of the central limit theorem for the normal approximation, which also applies to triangular arrays of random variables.

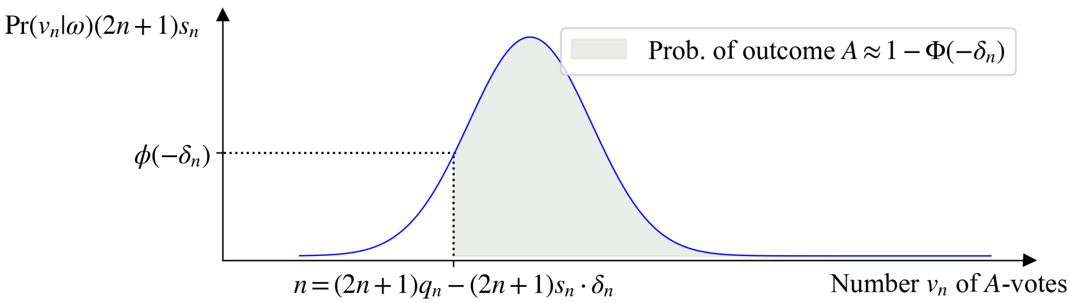


Figure 3: Illustration of the Normal approximation of the Binomial distribution of the number of A-votes v_n. The Binomial has mean $(2n+1)q_n$ for $q_n = q(\omega;\sigma_n)$ and standard deviation $(2n+1)s_n = ((2n+1)(q_n(1-q_n))^{\frac{1}{2}}$ for $s_n = s(\omega;\sigma_n)$. The outcome is A if there are more than n votes for A.

Here, $\Phi(\cdot)$ is the cumulative distribution of the standard normal distribution. So, the asymptotic distribution of the outcome policy only depends on $\lim_{n\to\infty}\delta_n(\omega;\sigma_n) \in \mathbb{R}\cup\{\infty,-\infty\}$.

An equilibrium sequence is *informative* if $\lim_{n\to\infty}\delta_n(\alpha;\sigma_n) - \delta_n(\beta;\sigma_n) \neq 0$. Informativeness captures that the aggregate effect of the voters' information acquisition on vote shares is large enough so as to impact outcomes. Given (24), it is a necessary condition for the outcome distribution to be different in the two states.

3.2.2 Close elections: An equilibrium outcome

For any informative equilibrium sequence, the outcome is close to being tied in *all* states ω,

$$\lim_{n\to\infty} q(\omega;\sigma_n) = \frac{1}{2} \tag{25}$$

Intuitively, the election must be close in at least *some* state since otherwise the incentives to acquire costly information are too small.[22]

Formally, a voters' individual incentives to acquire information depend on the pivotal likelihood; recall, for example, the cost-benefit analysis for the opti-

[22]This observation may be viewed as a rationalization of the frequent occurrence of close elections as an informational phenomenon. Historical examples of notoriously close elections include the 2000 US presidential election: George W Bush won the electoral college with 271 votes to Gore's 266 and lost the popular vote by some 500,000. Similarly, the 1960 election between Kennedy and Nixon was an extremely tight race, with the candidates tied at 47 percent in the Gallup polls. Kennedy won the popular vote by less than 120,000 votes. In Germany, chancellor Schröder won the 2002 federal election by a mere 6,000 out of more than 48 million votes.

mal (interior) precision, (15). A Stirling approximation of the pivotal likelihood yields[23]

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

This implies that the pivotal likelihood is exponentially small unless (25) holds for at least some state. This is because the function $q(1-q)$ takes the maximum $\frac{1}{4}$ at $q = \frac{1}{2}$ only. Therefore, if (25) does not hold in *any* state, voters acquire exponentially little information under the best response, given (15) and (17). Consequently, the difference of the vote shares in the two states—measured in standard deviations—goes to zero; that is, $\lim_{n\to\infty} \delta_n(\alpha; \sigma_n) - \delta_n(\beta; \sigma_n) = 0$. In other words, the equilibrium sequence is not informative.

The reason why the election is close in all and not just in one state (i.e. (25)) is that the likelihood that a random citizen votes A is asymptotically the same across states. This is because, given the best response to any strategy sequence, the signal precision of any voter type is of an order weakly smaller than $n^{-\frac{1}{2(d-1)}}$, given (15), (17), and (26). Thus $\Pr(\alpha|s, \text{piv}; \sigma_n, n)$ converges to $\Pr(\alpha|\text{piv}; \sigma_n, n)$ uniformly as $n \to \infty$. So, the definition (2) together with (7) and (8) implies

$$q(\omega; \sigma_n^*) \to \Psi(\Pr(\alpha|\text{piv}; \sigma_n, n)) \tag{27}$$

for both states $\omega \in \{\alpha, \beta\}$.

3.2.3 Limit marginal types

The closeness of the election, that is, (25), pins down the marginal types as $n \to \infty$. This is because the threshold of doubt $y(t) = \Pr(\alpha|\text{piv}; \sigma_n^*, n)$ of the marginal types necessarily satisfies

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

given (25) and (27). Since Ψ is continuous, this entails $\Psi(\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n^*, n)) = \frac{1}{2}$. Since Ψ is strictly increasing, this entails $\Pr(\alpha|\text{piv}; \sigma_n^*, n) \to \bar{y} \in (0, 1)$ where $\bar{y}$ is the unique belief for which $\Psi(\bar{y}) = \frac{1}{2}$.

[23] 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, $\binom{2n}{n} \approx (2\pi)^{-\frac{1}{2}} 2^{2n+\frac{1}{2}} n^{-\frac{1}{2}} = 4^n (n\pi)^{-\frac{1}{2}}$. Plugging this expression for the binomial coefficient into (6) yields $\Pr(\text{piv}|\omega; n) \approx 4^n (n\pi)^{-\frac{1}{2}} (q(1-q))^n$ for $q = q(\omega; \sigma_n)$.

4 Main result: Baseline setting

The literature on information aggregation in elections has established under fairly general conditions that large elections lead to full-information equivalent outcomes; that is, the policy preferred by the majority under full information is elected state-by-state.[24] This result has been established, in particular, for a setting identical to that of Section 3, but assuming that citizens receive an exogenous (costless) i.i.d. signal about the state (see Theorem 1 in Bhattacharya, 2013). In other words, with costless information, the competition between the interest groups (aligned and contrarians) is decided by the *size* of the interest groups. The majority group wins.

We show that information frictions alter the power relationships between the opposed groups fundamentally. Outcomes may not align with the majoritarian principle, but are driven by the endogenous informational efforts of the competing interest groups. Theorem 1 characterizes when an informative equilibrium exist. Further, it characterizes *all* informative equilibrium sequences, based on a measure of the type distribution (the index) that we will show to be proportional to the aggregate precision of the interest groups (Lemma 3).

Recall that d is the elasticity of the cost function, t_ω is the type's utility from policy A in ω, and $\bar{y}$ is the threshold of doubt of the limit marginal types. In the following, we denote by $\mathrm{E}(-|g)$ and $h(-|g)$ the conditional expectation and the conditional likelihood when conditioning on the set of types $\{t : t \in g\}$ of an interest group. Similarly, we use $h(g)$ for the unconditional likelihood and $\mathrm{E}(-|y)$ and $h(-|y)$ when conditioning on the set of types with threshold of doubt $y(t) = y$, et cetera. The *κ-index* of an interest group g in ω is

$$W(\kappa, g, \omega) = \underbrace{h(g)h(\bar{y}|g)}_{\text{likelihood of limit marginal types}} \underbrace{\mathrm{E}(||t_\omega||^\kappa|g, \bar{y}, \omega)}_{\kappa\text{-measured intensity}}, \tag{28}$$

for any $\kappa > 0$. Since all limit marginal types have the same *relative* intensities across the states, that is, $\frac{t_\alpha}{t_\beta} = -\frac{1-\bar{y}}{\bar{y}}$, the index differs only by a scalar across states, $W(\kappa, g, \alpha) = (\frac{1-\bar{y}}{\bar{y}})^\kappa W(\kappa, g, \beta)$. Hence, the order of the indices of the two interest groups does not depend on the state.

Theorem 1 *Let $d = \lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$ and $\kappa = \frac{2}{d-1}$. Take any preference distribution H such that Ψ is strictly increasing, the richness condition* (3) *holds and $W(\kappa, L, \alpha) \neq W(\kappa, C, \alpha)$.*

[24] See, for example, Feddersen and Pesendorfer (1997) and Austen-Smith and Banks (1996).

1. *There is an equilibrium sequence in which the policy preferred by the interest group (aligned or contrarians) with the higher κ-index is elected with probability converging to* 1 *as* $n \to \infty$.

2. *If* $\Psi(\Pr(\alpha)) \neq \frac{1}{2}$, *there is an equilibrium sequence in which the policy that is preferred by the majority of the citizens given the prior beliefs is elected with probability converging to* 1 *if the κ-index of the aligned is larger than that of the contrarians, and with probability converging to* 0 *if the κ-index of the aligned is smaller than that of the contrarians*

Welfare. The index W has a compelling interpretation in terms of welfare. A type's intensity t_ω is her willingness to pay for having the collective choice changed to the preferred policy. The index takes the willingness to pay of each type to the power $\kappa = \frac{2}{d-1}$ and then averages over the marginal types of the interest group. Hence, it interpolates between two extremes: When $\kappa = 0$, the index is purely ordinal. It is proportional to the likelihood of the marginal types. If $\kappa = 1$, the index is proportional to the utilitarian welfare of the marginal types. In general, preference intensities matter more when information of low precision is "cheaper"; that is, when d is lower and κ is higher.[25]

Full-Information Outcomes. Since Theorem 1 characterizes *all* informative equilibrium sequences, it implies that, when the contrarians have a higher index, there is no equilibrium sequence—informative or non-informative—in which the full-information outcome (A in α, B in β) is chosen in both states as the electorate grows large.

Factors of political power. In the equilibrium of the first item of Theorem 1, the election outcomes represent the political interests of the group with the higher index. Here, the parameter κ captures exactly how intensities substitute with the mass of the (marginal) types of a group in determining the political "power" of an interest group.

An informational tug-of-war. The next result, Lemma 3, shows that the κ-index of an interest group is proportional to the aggregate precision (or the informational effort) of the interest group, as $n \to \infty$. This result holds for the best response to any sequence of strategies with interior limit marginal types.

[25]For $c_d(x) = \frac{x^d}{d}$, we have $\lim_{x\to 0} \frac{c_d(x)}{c_{d'}(x)} = \infty$ if $d' > d$.

Lemma 3 *Let $d > 1$. For any strategy sequence $(\sigma'_n)_{n\in\mathbb{N}}$ for which* $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma'_n, n) = \bar{y} \in (0,1)$ *and any interest group $g \in \{L, C\}$, the best response $\sigma_n = (x_n, \mu_n)$ satisfies*

$$\lim_{n\to\infty} \frac{\int_{t\in g} x_n(t)dH(t)}{\Pr(\text{piv}|\sigma'_n, n)^{\frac{2}{d-1}} e(\bar{y}, d)} = W(g, \kappa, \alpha). \tag{29}$$

for a constant $e(\bar{y}, d) > 0$ that only depends on the threshold of doubt $\bar{y}$ of the limit marginal types and the cost elasticity $d > 0$.

Lemma 3 implies that the equilibrium of the first item of Theorem 1 resembles an informational tug-of war: The interest group with the higher aggregate precision wins the election. A sketch of the proof follows momentarily in Section 4.2 and the formal proof is in Appendix F.

4.1 Intuition for Theorem 1: Social inference and mis-coordination of the uninformed

With regard to the tug-of-war equilibrium, it is surprising that the voters who exert no informational effort seem to play no role—although they also vote. To understand the logic of this equilibrium, recall the example in Section 1. In the example, the types that stay uninformed mis-coordinate in an extreme way. Their votes cancel out each other completely. All other types t vote for their preferred policy with probability $\frac{1}{2} + x(t)$; compare to Lemma 1. As a consequence, the policy that is preferred by the interest group $g \in \{L, C\}$ with the higher aggregate precision $\int_{t\in g} x(t)dH(t)$ receives more votes in expectation.

In the example, the mis-coordination of the uninformed is driven by symmetry assumptions: All types hold a symmetric prior, and the reform-leaning and reform-skeptical types have the same likelihood. so that the votes of the uninformed split $50 - 50$ given their *prior beliefs*. In general, this is not true. However, we show that they *do* split close to $50 - 50$ given their *equilibrium beliefs*. This will be true in any informative equilibrium sequence.

Such mis-coordination is necessary in any informative equilibrium sequence. This is because the share of the uninformed goes to 1 and the election has to be close to $50 - 50$ when the electorate grows large, as observed in Section 3.2.2. Only this closeness creates incentives that are high enough so that sufficiently many voter types acquire costly information and consequently the equilibrium is informative.

Such mis-coordination is possible since the citizens do not only take into account their prior information when voting, but make an equilibrium inference from the behavior of the other voters. Namely, they update their beliefs conditional on the pivotal event. We will show that this "social inference" can be such that the votes of the uninformed split *close* to $50-50$ under the best response. Based on this, we construct the two equilibria of Theorem 1. They differ in how strongly the uninformed mis-coordinate. In the first equilibrium of Theorem 1, the mis-coordination is stronger and the outcomes are given by which group's κ-index is higher, or equivalently by which group's aggregate precision is higher, as in the example. In the second equilibrium, the asymmetry of the the preferences given the prior beliefs, $\Psi(\Pr(\alpha)) \neq \frac{1}{2}$, creates a bias towards one of the policies and the same policy is elected in both states.

By way of review, in Section 4.2, we sketch the proof of Lemma 3. In Section 4.3, we explain the relevant condition on the information cost from Theorem 1 ($d > 3$) that is necessary for the existence of informative equilibria.[26] In Section 4.4, we analyze the equilibrium inference of the voters. There, we will also illustrate how the information acquisition of the voters can be complementary, which gives an intuition for the equilibrium multiplicity. In Section 4.5, we prove Theorem 1.

4.2 The endogenous information of the interest groups

Sketch of the proof of Lemma 3. Fix an interest group $g \in \{L, C\}$; for example, the aligned types. To evaluate the integral $\int_{t\in g} x_n(t)dH(t) = h(g)\mathrm{E}(x_n(t)|g)$, we aggregate over all types that acquire information.[27] For this, recall that there is a one-to-one relation between types t and pairs of thresholds of doubt $y(t)$ and total intensities $k(t)$.[28] Moreover, for any fixed total intensity k, only the types with threshold of doubt $y(t)$ in the interval $[y_g^-(k), y_g^+(k)]$ choose a positive precision, as is illustrated in Figure 6.[29] We integrate iteratively, first along the y-dimension, then along the k-dimension.

[26] This condition is similar to a condition ($c'''(0) = 0$) identified by Martinelli (2007). For the power cost functions $c(x) = kx^d$, the conditions are equivalent.

[27] Here, we use the earlier short-hand notation $\mathrm{E}(-|g)$ for the expectation conditional on the set of types of the interest group g. Below we will use more of such short-hand notation; for example, $H(|g, k)$ for the distribution of the types conditional on the set of types of an interest group $g \in \{L, C\}$ with total intensity $k(t) = k$.

[28] For example, for the aligned types, $t_\alpha = k(t)(1 - y(t))$ and $t_\beta = -k(t)y(t)$, given (4) and (5).

[29] Note that we suppress the dependence on n in the notation for the interval boundaries.

Fix $k(t) = k$. In the proof, we make two key observations. Take any two sequences of types $(t_n)_{n\in\mathbb{N}}$, $(t'_n)_{n\in\mathbb{N}}$ in the information acquisition interval. First, we show that these types choose signals with asymptotically equivalent precision; that is, $x_n(t_n) \approx x_n(t'_n)$.[30] As a consequence, the integral $\int_{t\in g:k(t)=k} x_n(t)dH(t|g,k)$ is asymptotically equivalent to the mass of the types in the information acquisition interval times the precision of the marginal type $\bar{t}_n(k)$ around which the interval forms. Second, we show that the mass of the types in the information acquisition interval is asymptotically proportional to the precision times the likelihood of the marginal type. Formally, $\Pr(\{t : x(t) > 0\}|g,k) \approx h(\bar{t}_n(k)|g,k)x_n(\bar{t}_n(k))c(\bar{y},d)$ for some constant $c(\bar{y},d) > 0$. Importantly, this constant only depends on the cost elasticity d and the limit marginal type $\bar{y}$ from Section 3.2.3, but not on k.

We aggregate over k and use these two observations to show that the integral $\int_{t\in g} x_n(t)dH(t)$ is proportional to the likelihood of the marginal types of an interest group times the average of the square precision they choose, $\mathrm{E}(x_n(\bar{t}_n(k))^2|g)$.

Finally, what matters is that the intensities t_ω pin down the precision choices. Precisely, combining (15), (17), and the indifference condition $t_\alpha \Pr(\alpha|\mathrm{piv};\sigma'_n,n) = -t_\beta \Pr(\beta|\mathrm{piv};\sigma'_n,n)$ of a given marginal type yields

$$\frac{x_n(\bar{t}_n(k))}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{1}{d-1}}} = t_\alpha^{\frac{1}{d-1}} e_1(\bar{y}_n,d), \tag{30}$$

for $\bar{y}_n = \Pr(\alpha|\mathrm{piv};\sigma'_n,n)$ the threshold of doubt of the marginal types and $e_1(\bar{y}_n,d) = (2\bar{y}_n)^{\frac{1}{d-1}}$. This explains that the mean $\mathrm{E}(t_\alpha^{\frac{2}{d-1}}|g,\bar{y}_n)$ of the exponentiated intensities of the marginal types together with the likelihood of the marginal types is what pins down the aggregate precision $\int_{t\in g} x_n(t)dH(t)$. Precisely, we show that $\int_{t\in g} x_n(t)dH(t)$ is asymptotically proportional the κ-index $W(g,\kappa,\alpha)$ for $\kappa = \frac{2}{d-1}$, in the way claimed in Lemma 1.

4.3 Existence: Free-riding and information cost

The voters face a free-rider problem. If a voter acquires information, she is bearing the cost privately, while all voters with the same interest benefit from her casting a more informed ballot. In the following, we explain why the condition $d > 3$ from Theorem 1 is the critical condition for the severity of the free-rider problem in a large electorate. In particular, we sketch an argument based on two

[30] Formally, two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are asymptotically equivalent if $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$.

observations, showing that, if $d < 3$, no informative equilibrium sequence exists.

Take a candidate informative equilibrium sequence. The election is necessarily close in both states; that is, (25) holds. Given (24), what matters for the "informativeness" of the aggregate voting behavior is the distance between the expected vote share in the two states in terms of standard deviations,

$$\begin{aligned}\lim_{n\to\infty} \delta_n(\alpha;\sigma_n) - \delta_n(\beta;\sigma_n) &= \lim_{n\to\infty} \frac{q(\alpha;\sigma_n) - q(\beta;\sigma_n)}{s(\alpha;\sigma_n)} \\ &= \lim_{n\to\infty} \frac{2\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t)}{s(\alpha;\sigma_n)}. \quad (31)\end{aligned}$$

Here, we used the definition (23) of $s(\omega;\sigma_n)$ and that $\lim_{n\to\infty} \frac{s(\alpha;\sigma_n)}{s(\beta;\sigma_n)} = 1$, given (25). Hence, the relevant comparison is how fast the aggregate precision $\int_t x_n(t)dH(t)$ decreases relative to how fast the standard deviation of the vote share increases.

We make two observations. The first observation is that, depending on whether $d < 3$ *or* $d > 3$, the aggregate precision acquired by the voters of any given interest group is of an order smaller *or* larger than the pivotal likelihood. This is a direct consequence of Lemma 3; see (29).

The second observation is that the normal approximation (24) also holds locally,[31]

$$\lim_{n\to\infty} \Pr(\text{piv}|\omega;\sigma_n)(2n+1)s(\omega;\mathbf{q}(\sigma_n)) = \lim_{n\to\infty} \phi(\delta_n(\omega;\sigma_n)), \quad (32)$$

where ϕ the density of the standard normal distribution. This approximation is illustrated in Figure 3. Let $s_n = s(\omega;\mathbf{q}(\sigma_n))$ and $q_n = q(\omega_n;\sigma_n)$. Given (32), the pivotal likelihood is a finite multiple of $((2n+1)s_n)^{-1}$. Since $((2n+1)s_n)^{-1} = s_n(q_n(1-q_n))^{-1}$, it is a finite multiple of the standard deviation.[32] Combining this with the first observation, we see that the aggregate precision vanishes relative to the standard deviation if $d < 3$. Hence, the candidate sequence cannot be informative, given (31).

[31] The local central limit theorem is due to Gnedenko (1948). The version that we apply is the one for triangular arrays of integer-valued variables as in Davis and McDonald (1995), Theorem 1.2. Compare also to the equation (11) therein.

[32] Recall that $((2n+1)s_n)^{-1}$ is the standard deviation of the Binomial distribution of the number of vote shares. Note that $((2n+1)s_n)^{-1} = \left[(2n+1)(q_n(1-q_n))\right]^{-\frac{1}{2}} = s_n(q_n(1-q_n))^{-1}$ since $s_n = (\frac{(2n+1)}{q_n(1-q_n)})^{-\frac{1}{2}}$; see (23) and thereafter.

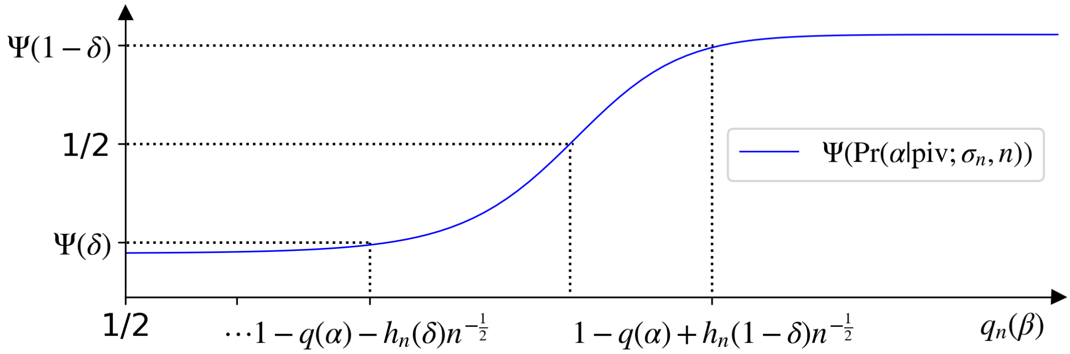


Figure 4: Fix $q(\alpha) < \frac{1}{2}$. The figure shows the limit vote share for policy A under the best response as $n \to \infty$, that is, $\Psi(\Pr(\alpha|\text{piv};\sigma_n'))$, as a function of the expected vote share in β, for $q_n(\beta) > \frac{1}{2}$. The function $h_n(x)$ is so that, given $(q_n(\beta) - \frac{1}{2}) - (\frac{1}{2} - q(\alpha)) = h_n(x)n^{-\frac{1}{2}}$, the limit vote share is x.

4.4 Existence: Social Inference and Information complementarities

What drives the existence of the two informative equilibrium sequences of Theorem 1—besides information of low precision being sufficiently cheap ($d > 3$)—is that the voters' information acquisition exhibits complementarities. Below, we sketch how these complementarities act.

Fix a vote share $q(\alpha) < \frac{1}{2}$. We can vary the informativeness of a voter strategy σ_n with $q(\alpha;\sigma_n') = q(\alpha)$ and $q(\beta;\sigma_n') = q_n(\beta)$ by varying $q_n(\beta)$; see Section 3.2.1. Figure 4 shows the limit vote share ($n \to \infty$) for policy A under the best response σ_n', as a function of $q_n(\beta) > \frac{1}{2}$. The limit vote share only depends on the prior belief and the inference from the pivotal event given the behavior of the others (the "social inference"). It is given by $\lim_{n\to\infty} q(\omega;\sigma_n) = \Psi(\Pr(\alpha|\text{piv};\sigma_n'))$; compare to (27).

As more people vote A in β, the vote share in β is less close to the majority threshold. Then, voters believe the state α to be more likely conditional on the election being tied. The support for A increases since preferences are "monotone"; that is, Ψ is strictly increasing. Importantly, as we will show, there are vote shares in β so that the election becomes close to being tied under the best response, $\Psi(\Pr(\alpha|\text{piv};\sigma_n')) \approx \frac{1}{2}$. Hence, certain levels of informative voting induce a close election and thereby high incentives to acquire information. This way, information acquisition can be complementary.

4.5 Proof of Theorem 1

We represent informative equilibrium sequences in a compact way as sequences of roots of one-dimensional auxiliary maps. First, we show that equilibrium can be alternatively characterized in terms of the vector of the expected vote shares of A in state α and β; that is,

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

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 (6). Given (9)-(12), (17), (20)- (21), the vector of the pivotal likelihoods is a sufficient statistic for the best response. Thus, $\mathbf{q}(\sigma)$ is a sufficient statistic 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, given $\mathbf{q}$. Then, σ^* is an equilibrium, if and only if, $\sigma^* = \sigma^{\mathbf{q}(\sigma^*)}$. Conversely, an equilibrium can be described by a vector of vote shares $\mathbf{q}^* = (q^*(\alpha), q^*(\beta))$ that is a fixed point of $\mathbf{q}(\sigma^{-})$, i.e.,

$$q^*(\alpha) = q\left(\alpha;\sigma^{\mathbf{q}^*}\right), \tag{33}$$

$$q^*(\beta) = q\left(\beta;\sigma^{\mathbf{q}^*}\right), \tag{34}$$

4.5.1 The one-dimensional auxiliary maps

We use the insights from Section 4.4 to select curves of vote share vectors that solve (34). Let us sketch the argument. For example, take any $\frac{1}{2} - \epsilon < q(\alpha) < \frac{1}{2}$ with $\epsilon > 0$. Figure 4 shows the limit vote share $\lim_{n\to\infty} q(\beta;\sigma^{(q(\alpha),q_n(\beta))}) = \Psi(\Pr(\alpha|\text{piv};\sigma^{(q(\alpha),q_n(\beta))})$ as a function of $q_n(\beta)$. It is smaller than $\frac{1}{2}$ for $q_n(\beta) = \frac{1}{2}$ and it is close to $\Psi(1) > \frac{1}{2}$ for $q_n(\beta) = \frac{1}{2} + 2\epsilon$. Thus,

$$q(\beta;\sigma^{(q(\alpha),q_n(\beta))}) < q(\beta) \text{ for } q_n(\beta) = \frac{1}{2}, \text{ and}$$
$$q(\beta;\sigma^{(q(\alpha),q_n(\beta))}) > q(\beta) \text{ for } q_n(\beta) = \frac{1}{2} + 2\epsilon,$$

for $\epsilon > 0$ small enough. An application of the intermediate value theorem yields that for any n large enough, there are vote shares $q_n(\beta)$ so that $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ solves (34). In Appendix B, we provide a topological argument to make a continuous selection of such vote share vectors $\mathbf{q}_n$. Further, we do not only construct one continuous curve, but four. They differ in whether the vote share in a given state is larger or smaller than the majority threshold. Formally, we prove the following lemma.

Lemma 4 *Let $d > 1$. For any $(x(\alpha), x(\beta)) \in \{0,1\}^2$, there are $\epsilon > 0, \Delta > 0, \bar{n} \in \mathbb{N}$ so that for all $n \geq \bar{n}$, there is a continuous map*

$$\begin{aligned} v_n : [0,1] &\rightarrow [\Psi(0)+\epsilon, \Psi(1)-\epsilon]^2 \\ t &\mapsto \mathbf{q}_n^t = (q_n^t(\alpha), q_n^t(\beta)), \end{aligned}$$

so that $\mathbf{q}_n^t$ solves (34), $\operatorname{sgn}(q^t(\omega) - \frac{1}{2}) = x(\omega)$ for all $t \in [0,1]$ and $\omega \in \{\alpha, \beta\}$, $(q_n^0(\alpha), q_n^1(\alpha)) = (\frac{1}{2} - \epsilon, \frac{1}{2} - \Delta n^{-\frac{1}{2}})$ if $x(\alpha) = 0$, and $(q_n^0(\alpha), q_n^1(\alpha)) = (\frac{1}{2} + \Delta n^{-\frac{1}{2}}, \frac{1}{2} + \epsilon)$ if $x(\alpha) = 1$.

The vote shares pairs of the lemma are "similarly" far away from the majority threshold. Precisely, we claim that the distance only differs by finitely many standard deviations of the vote share; that is,[33]

$$\begin{aligned} \lim_{n\to\infty} \frac{|q_n^t(\alpha) - \frac{1}{2}| - |q_n^t(\beta) - \frac{1}{2}|}{s(\alpha; \mathbf{q}_n^t)} &\notin \{-\infty, \infty\} \\ \Leftrightarrow \lim_{n\to\infty} |\delta_n(\alpha; \mathbf{q}_n^t)| - |\delta_n(\beta; \mathbf{q}_n^t)| &\notin \{-\infty, \infty\} \end{aligned} \tag{35}$$

Note that we slightly abuse the previous notation here by treating the vote share pair $\mathbf{q}_n^t$ as a strategy. If (35) would not hold, the inference from conditioning on the election being tied would be unbounded; that is, the posteriors $\Pr(\alpha|\text{piv}; \mathbf{q}_n^t)$ would converge to 0 or 1.[34] Thus, the vote shares of the best response to $\Phi(0)$ and $\Phi(1)$, given (27). However, the vote share $q_n^t(\beta)$ from the lemma solves (34), that is, it is a fixed point under the best response. Since it is bounded away from $\Phi(0)$ and $\Phi(1)$, so is the vote share under the bet response. We arrive at a contradiction and conclude that (35) holds.

It follows from Lemma 4 that any root of the maps[35]

$$\hat{v}_n : t \mapsto q_n^t(\alpha) - q(\alpha; \sigma^{\mathbf{q}_n^t}), \tag{36}$$

[33]For the equivalence on the second line, we use the definition (23) and that $\lim_{n\to\infty} |\delta_n(\alpha; \mathbf{q}_n^t)| - |\delta_n(\beta; \mathbf{q}_n^t)| \notin \{-\infty, \infty\}$ implies that $\lim_{n\to\infty} \frac{s(\alpha;\mathbf{q}_n^t)}{s(\beta;\mathbf{q}_n^t)} = 1$.

[34]Formally, this follows since $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha;\mathbf{q}_n^t))}{\phi(\delta_n(\beta;\mathbf{q}_n^t))} = \lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n^t,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n^t,n)}$, given (32) (see also Appendix K for a comprehensive lemma on the voter's pivotal inference). Thus, $\lim_{n\to\infty} |\delta_n(\alpha; \mathbf{q}_n^t)| - |\delta_n(\beta; \mathbf{q}_n^t)| \in \{-\infty, \infty\}$ implies $\lim_{n\to\infty} \delta_n(\alpha; \mathbf{q}_n^t)^2 - \delta_n(\beta; \mathbf{q}_n^t)^2 \in \{-\infty, \infty\}$, so that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n^t,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n^t,n)} = \lim_{n\to\infty} e^{-\frac{1}{2}(\delta_n(\alpha;\mathbf{q}_n^t)^2 - \delta_n(\beta;\mathbf{q}_n^t)^2)} \in \{0, \infty\}$, and $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n^t) \in \{0, 1\}$.

[35]Note that we omit the dependence of $\hat{v}_n$ on $(x(\alpha), x(\beta)) \in \{0,1\}^2$ in the notation.

with $\mathbf{q}_n^t = (q_n^t(\alpha), q_n^t(\beta)) = v_n(t)$, satisfies (33) - (34). So, such a vote share pair corresponds to an equilibrium of the voting game. In Section 4.5.2, we construct the informative equilibrium sequences of Theorem 1 as roots of the maps $\hat{v}_n$.

4.5.2 Proof: Minority-preferred outcomes in all states

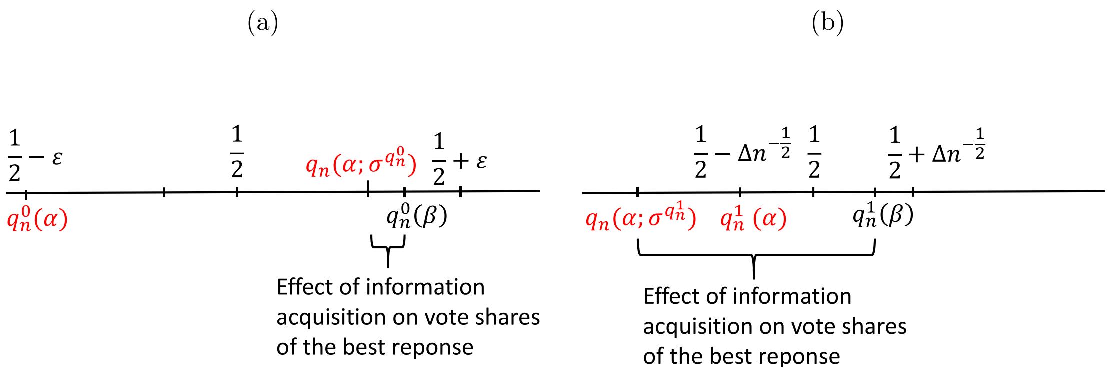
In the following, we provide the argument for the equilibrium of the first item of Theorem 1. Here, in the main text, we consider the case in which the contrarians have a higher index, $W(L, \kappa, \alpha) < W(C, \kappa, \alpha)$ for $\kappa = \frac{2}{d-1}$. In this case, the equilibrium leads to the outcomes preferred by the contrarians, which are a minority in expectation. The argument for the other cases and for the equilibrium of the second item is analogous and provided in Appendix G.

The argument relies on a precise analysis of the voters' incentives to acquire costly information. Further, we will use the observation that if $W(L, \kappa, \alpha) < W(C, \kappa, \alpha)$, Lemma 3 and (31) together imply that for any $\mathbf{q}_n$ and n large enough, the vote shares of the best response are ordered as

$$q(\alpha; \sigma^{\mathbf{q}_n}) < q(\beta; \sigma^{\mathbf{q}_n}). \tag{37}$$

Take the function v_n and consider the case $(x(\alpha), x(\beta)) = (0, 1)$ so that the function maps to vote shares $q_n^t(\alpha) < \frac{1}{2}$ and $q_n^t(\beta) > \frac{1}{2}$. Figures 5a and 5b illustrate the argument, which establishes that $\hat{v}_n$ has a root when n is large enough.

Figure 5: Fixed point argument



Panel (b) illustrates $\hat{v}_n(t)$ for $t = 1$. Recall from Lemma 4 that $q_n^1(\alpha)$ is finitely many standard deviations away from the majority threshold. Given (35), $q_n^1(\beta)$ is finitely many standard deviations above the majority threshold as $n \to \infty$.

The expectation of a close election in α and β creates relative large incentives to acquire information. Lemma 5 at the end of this section shows that, given the condition on the information cost, $d > 3$, these incentives are large enough so that the vote shares of the best response to $\mathbf{q}_n^1$ differ by arbitrarily many standard deviations in the two states when n grows large. The effect of the information acquisition is indicated in the figure as the distance $q(\alpha; \sigma^{\mathbf{q}_n^1}) - q(\beta; \sigma^{\mathbf{q}_n^1})$ between the vote shares of the best response. Since $q_n^1(\beta) = q(\beta; \sigma^{\mathbf{q}_n^1})$ and given (37),

$$\hat{v}_n(1) = q_n^1(\alpha) - q_n(\alpha; \sigma^{\mathbf{q}_n^1}) > 0 \tag{38}$$

for n large enough.

Panel (a) shows the map $\hat{v}_n(t)$ for $t = 0$. Recall from Lemma 4 that $q_n^0(\alpha) = \frac{1}{2} - \epsilon$. Hence, $q_n^0(\alpha)$ is bounded away from $\frac{1}{2}$ by some constant. Given (35), the same is true for $q_n^0(\beta)$ as $n \to \infty$. As a consequence, the incentives to acquire information are small. In fact, the pivotal likelihood is exponentially small for large n, given (26), and so is the precision of any voter type under the best response; see (17). Given exponentially little information acquisition, the vote shares of the best response do not differ by a standard deviation, as $n \to \infty$.[36] The effect of the information acquisition is indicated in the figure as the distance $q(\alpha; \sigma^{\mathbf{q}_n^0}) - q(\beta; \sigma^{\mathbf{q}_n^0})$ between the vote shares of the best response. Since $q_n^0(\beta) = q(\beta; \sigma^{\mathbf{q}_n^0})$ and $q_n^0(\beta) > \frac{1}{2}$ by construction,

$$\hat{v}_n(0) = q_n^0(\alpha) - q_n(\alpha; \sigma^{\mathbf{q}_n^0}) < 0 \tag{39}$$

for n large enough.

Finally, using (38), (39), and that $\hat{v}_n$ is continuous, an application of Kakutani's fixed point theorem shows that there is $t \in (0, 1)$ so that $\mathbf{q}_n^t = (q_n^t(\alpha), q_n^t(\beta))$ solves (33) and (34).

Further, it must be that $\lim_{n\to\infty} \frac{\frac{1}{2} - q_n^t(\alpha)}{s(\alpha; \mathbf{q}_n^t)} = \infty$ since otherwise (39) holds by the same argument as just given when discussing panel (b). Hence, also $\lim_{n\to\infty} \frac{q_n^t(\beta) - \frac{1}{2}}{s(\beta; \mathbf{q}_n^t)} = \infty$, given (35). The distance of the vote shares to the majority threshold becomes arbitrarily large in terms of standard deviations. This implies that B gets elected in α and A in β with probability converging to 1 as $n \to \infty$, given (24). The proof of Lemma 5 is in Appendix C.

Lemma 5 *Let $d > 3$. Take a monotone preference distribution H for which*

[36] Here, recall that the standard deviation of the vote share is of the order of $\frac{1}{\sqrt{n}}$, $s(\omega; \mathbf{q}_n) = (2n+1)^{-\frac{1}{2}} (q_n(\omega)(1 - q_n(\omega))^{\frac{1}{2}}$; see Section 3.2.1.

$W(\kappa, L, \alpha) \neq W(\kappa, C, \alpha)$ *for* $\kappa = \frac{2}{d-1}$. *Let* $t \in [0,1]$. *Take* $\mathbf{q}_n^t$ *as in Lemma 4. If* $\lim_{n\to\infty} \frac{|q_n^t(\alpha) - \frac{1}{2}|}{s(\alpha;\mathbf{q}_n^t)} \in \mathbb{R}$, *then,*

$$\lim_{n\to\infty} \frac{|q(\alpha; \sigma^{\mathbf{q}_n^t}) - q_n^t(\beta)|}{s(\alpha; \mathbf{q}_n^t)} = \infty. \tag{40}$$

5 Additional results and discussion

General setting: How differences in cost and priors affect outcomes. We return to the setting of Section 2 in which types are heterogeneous not only in the state-dependent intensities t_ω, but also in the cost type and the prior belief. In the working paper, we show that a generalization of Theorem 1 holds for this setting (Theorem 4).[37]

The proof is based on one main insight: We show that for any joint distribution of types, there is an auxiliary distribution that only admits heterogeneity in the state-dependent intensities and is *outcome-equivalent*; that is, it leads to the same set of equilibrium outcome distributions. This insight allows to leverage the previous analysis. Two observations lead to the auxiliary distribution. First, as in the baseline setting, equilibrium can be characterized as an *equilibrium vote share pair* $\mathbf{q} = (q(\alpha), q(\beta))$ that is a fixed point of the best response mapping on the level of vote shares, compare to (33)-(34). Second, we show that, for any type $t = (v, r, t_\alpha, t_\beta)$, the type $\zeta(t) = (\frac{1}{2}, 1, 2t'_\alpha, 2t'_\beta)$ with

$$t'_\alpha = \frac{v t_\alpha}{r}, \text{ and} \tag{41}$$

$$t'_\beta = \frac{(1-v)t_\beta}{r} \tag{42}$$

best responds in the same way to any given vote share pair $\mathbf{q}$. As a consequence, for any type distribution H, the push-forward distribution $\zeta^*(H)$ has the same equilibrium vote share pairs, but does not exhibit heterogeneity in priors and cost types.

The formal statement of Theorem 4 is almost identical to Theorem 1, except that Ψ and W are replaced by their generalizations $\mathcal{X}(p) = \Pr_H(\{t : pt'_\alpha + (1-p)t'_\beta \geq 0\})$ and $I(\kappa, g, \omega) = h(M^g)\mathrm{E}(\|\, t'_\omega \,\|^\kappa \mid M^g, \omega)$, in which t'_ω takes the the role of t_ω. Here, M^g is the set of the limit marginal types of interest group $g \in$

[37] The working paper is available here: https://www.researchgate.net/publication/358742924_Elections_with_information_frictions_and_distributive_uncertainty.

$\{L, C\}$ and $h(M^g)$ their likelihood.[38] In particular, Theorem 3 shows that there is a tug-of-war equilibrium in the general setting. In this equilibrium, the policy preferred by the interest group with the higher aggregation precision is elected, and the aggregate precision of an interest group $g \in \{L, C\}$ is proportional to $I(g, \kappa, \omega)$ for $\kappa = \frac{2}{d-1}$, analogous to the Lemma 3.

Theorem 4 yields important insights about the effect of differences in cost and prior beliefs on outcomes. First, consider the situation in which the types of a a given interest group have comparably high cost r. Intuitively, this will depress their information acquisition. Formally, the "weights" t'_ω become small when r is large; see (41)-(42). So, the index $I(\kappa, g, \omega)$ of the group will be smaller than the other interest group's index if r is sufficiently high, ceteris paribus. Then, the group's preferred policy is not elected in any state, given the tug-of-war equilibrium. Second, if an interest group has more dispersed prior beliefs, this may depress information acquisition similarly. We illustrate this for a class of symmetric distributions in the working paper.[39] We show that the likelihood of the marginal types decreases with the dispersion. Based on this, we show that if an interest group has sufficiently dispersed prior beliefs, there is an equilibrium in which the group's preferred policy is not elected in any state in the tug-of-war equilibrium.

Non-informative equilibrium sequences. Generically, there exist equilibrium sequences that are not informative, and in any non-informative limit equilibrium, all voters vote according to their prior belief. Thus, the policy that is preferred by a majority given the prior beliefs will be elected: the outcome is A if $\Psi(\Pr(\alpha)) > \frac{1}{2}$ and B if $\Psi(\Pr(\alpha)) < \frac{1}{2}$. The proof of Theorem 2 is in Appendix H.

Theorem 2 *Let* $\Psi(\Pr(\alpha)) \neq \frac{1}{2}$.

1. *There exists an equilibrium sequence that is not informative.*

2. *All equilibrium sequences* $(\sigma_n^*)_{n \in \mathbb{N}}$ *that are not informative satisfy* $\lim_{n \to \infty} q(\omega; \sigma_n^*) = \Psi(\Pr(\alpha))$ *for all states* $\omega \in \{\alpha, \beta\}$. *Hence,* $\lim_{n \to \infty} \Pr(A | \sigma_n^*, n) = 1$ *if* $\Psi(\Pr(\alpha)) > \frac{1}{2}$ *and* $\lim_{n \to \infty} \Pr(B | \sigma_n^*, n) = 1$ *if* $\Psi(\Pr(\alpha)) < \frac{1}{2}$.

[38] In the working paper, we derive the set M^g in terms of primitives, in a similar way as in the baseline setting; compare to Section 3.2.3.

[39] The working paper is available here: `https://www.researchgate.net/publication/358742924_Elections_with_information_frictions_and_distributive_uncertainty`.

Ordering the equilibrium sequences along their informativeness or the aggregate precision. Theorem 1 and Theorem 2 show that there exist three types of equilibrium sequences when $d > 3$ and $\Psi(\Pr(\alpha)) \neq \frac{1}{2}$. We show that the three types of equilibrium sequences can be ordered by their (absolute) informativeness, that is, by $\lim_{n\to\infty} |\delta_n(\alpha; \sigma_n) - \delta_n(\beta; \sigma_n)|$.

In any non-informative equilibrium sequence, by definition,

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

For the tug-of-war equilibrium sequence in which the policy preferred by the interest group with the higher index is elected as $n \to \infty$, the distribution of the limit outcomes is degenerate and varies with the state. Thus, (24) implies that

$$\lim_{n\to\infty} |\delta_n(\alpha; \sigma_n) - \delta_n(\beta; \sigma_n)| = \infty.$$

Take the other informative equilibrium sequence of Theorem 1 in which the limit outcome is the same in both states. This implies that the sign of $\delta_n(\alpha; \sigma_n)$ and $\delta_n(\beta; \sigma_n)$ is the same for n large enough, given (24). Therefore, (35) implies

$$\lim_{n\to\infty} |\delta_n(\alpha; \sigma_n) - \delta_n(\beta; \sigma_n)| \in (0, \infty).$$

We conclude that the informativeness lies in between that of the other two types of equilibrium sequences.

We can also think of these results as an ordering by the aggregate precision $\int_t x_n(t)dH(t)$ of the voters. This is because the informativeness is asymptotically proportional to the aggregate precision. This can be seen as follows: Given Lemma 3, the aggregate precision of the types of an interest group $g \in \{L, C\}$ compares as follows to the aggregate precision of all citizens, $\lim_{n\to\infty} \frac{\int_{t\in g} x_n(t)dH(t)}{\int_t x_n(t)dH(t)} = \frac{W(g,\kappa,\alpha)}{W(L,\kappa,\alpha)+W(C,\kappa,\alpha)}$. Combining this with (31),

$$\lim_{n\to\infty} |\delta_n(\alpha; \sigma_n) - \delta_n(\beta; \sigma_n)| = \lim_{n\to\infty} \frac{\frac{|2(W(L,\kappa,\alpha)-W(C,\kappa,\alpha))|}{W(L)+W(C)} \int_t x_n(t)dH(t)}{s(\alpha; \sigma_n)}.$$

6 Literature: Information aggregation in elections

The literature on information aggregation has shown that elections effectively aggregate *exogenous* information that is dispersed among many voters, so that outcomes in all equilibria are "as if" there is no uncertainty about the state (Austen-Smith and Banks, 1996; Feddersen and Pesendorfer, 1997).[40] Our analysis provides a succinct comparison by modifying a setting (as in Bhattacharya (2013)), in which information aggregates, to account for the voters' information being endogenous.[41] We show that information aggregation fails in the low and medium effort equilibrium and that it can even fail in all equilibria. The previous literature has identified other reasons for a failure of information aggregation, which are not present in our model.[42]

Our analysis also applies to situations of common interest, that is, when there is only one interest group. For these situations, the information aggregation result of the literature means that outcomes are (ex-post) utilitarian efficient. An open question is to which extent this efficiency result (widely known as the "Condorcet Jury Theorem") also holds when the voters' information is endogenous. Previous work on this question has found in a variety of settings that information aggregates, given appropriate conditions on information cost.[43] These findings are important and surprising possibility results, demonstrating that the intuition from Downs "rational ignorance hypothesis" does not necessarily imply inefficient election outcomes. [44] The analysis in this paper is considerably more general relative to previous work. We consider all (continuous) voter type distributions.[45] We find that the Condorcet Jury Theorem does not hold. The low and medium effort equilibrium are inefficient and exist for almost all type distributions. This equilibrium multiplicity is a novel finding. In contrast, in the settings of the literature, information aggregates in *all* equilibria when information cost are sufficiently low.

[40]See also Myerson (1998), Wit (1998), and Duggan and Martinelli (2001).

[41]To be precise, Bhattacharya (2013) shows that a sufficient and necessary condition for information aggregation is that preferences are "monotone". See also Acharya (2016) and Bhattacharya (2018). We maintain the appropriate monotonicity conditions for all results. We discuss the effect of non-monotonicities in our setting in Appendix I.

[42]Several failures due to an "invertibility problem" have been observed in settings in which the effective state is multi-dimensional (Feddersen and Pesendorfer, 1997; Mandler, 2012; Barelli *et al.*, 2019). Other mechanisms that lead to a failure are signaling motives (Razin, 2003), policy uncertainty (Gul and Pesendorfer, 2009), divided majorities (Bouton and Castanheira, 2012), and adverse selection problems (Ali *et al.*, 2018). A recent stream of literature considers "extended" election games in which biased third-parties inflict the failure (Bond and Eraslan, 2010; Ekmekci and Lauermann, 2020; Heese and Lauermann, 2019).

[43]See Theorem 2 and 6 in Martinelli (2006), Theorem 3(ii) in Martinelli (2007), Theorem 3 in Triossi (2013), and Proposition 5 in Oliveros (2013).

[44]See Downs *et al.* (1957); in particular, p.246.

[45]The previous work has considered settings in which all types share common preference intensities $t = (t_\alpha, t_\beta)$ or assumed symmetry conditions for the type distribution so that $\Phi(\Pr(\alpha)) = 1/2$.

7 Conclusion

We analyzed a model of collective choice with information frictions, that is, information cost. The model reflects situations in which decentralized, i.e. not centrally organized, political interests compete for influence. It is one of the first models to formalize how the competition of political interests channels through information frictions and informational efforts. Applications include general elections and referenda on distributive reforms.

In the model, absent information cost, outcomes in all equilibria are full-information equivalent (Bhattacharya, 2013); thus, the outcomes preferred by the majority interest group are elected. The main insight is that the information frictions fundamentally alter the power relationships between opposed interest groups, by turning the election into an informational contest. There is a "tug-of-war" equilibrium in which the policy preferred by the interest group with the higher aggregate informational effort is elected. This equilibrium is cardinal in the sense that outcomes represent voters with a minority interest if they have sufficiently high utilities at stake. Information advantages and internal conflicts of opinion matter for political influence: members of an interest group cast their votes in a less coordinated way, i.e., their binary voting choices are less correlated with the state, when they face higher information cost or if the distribution of prior beliefs about the state is more dispersed.

Another insight is that the information frictions create strategic complementarities. These complementarities modulate the competitive forces of endogenous information acquisition. They act in a way so that there are three equilibria, ordered by the aggregate informational effort of the electorate. The tug-of-war equilibrium is effort-maximal and resembles an informational contest. In the other equilibria, the competitive forces are less strongly shaped by informational efforts and also prior beliefs matter for outcomes.

We have provided some discussion and initial observations on similarities to and differences from classical models of decentralized political competition of the literature, such as Palfrey and Rosenthal (1985), Krishna and Morgan (2011, 2015) and more recent contributions (Lalley and Weyl, 2018; Eguia and Xefteris, 2021). We believe that several features of our model may lead to interesting observations when integrated into these models. For example, the policy uncertainty plays a crucial role in our setting and implies that changes in the prior belief distribution may upset the election outcome. Similar results may obtain in variants of the mentioned models.

Further, we discussed the implications of our results with respect to information aggregation in elections. Importantly, when information is endogenous due to information cost, information aggregation fails in (non-trivial) equilibria even when information cost are arbitrarily low. This contrasts findings in previous settings, in which no equilibrium multiplicity has been found and information aggregates in all equilibria when information cost are sufficiently low.

Appendix

A 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

$$\Pr(\alpha|b, \text{piv}; \sigma', n) < \Pr(\alpha|a, \text{piv}; \sigma', n). \tag{43}$$

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

$$\Pr(\alpha|b, \text{piv}, \sigma', n) < y(t) < \Pr(\alpha|b, \text{piv}, \sigma', n). \tag{44}$$

Otherwise, given (9)-(12), 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, (9)-(12), and (44) together imply (18) and (19)

Proof of (43)**.** Note that the posterior likelihood ratio of the states conditional on a signal $s \in \{a, b\}$ with precision $x(t)$ and conditional on the event that the voter is pivotal is

$$\frac{\Pr(\alpha|s, \text{piv}; \sigma', n)}{\Pr(\beta|s, \text{piv}; \sigma', n)} = \frac{\Pr(\alpha)}{\Pr(\beta)} \frac{\Pr(\text{piv}|\alpha; \sigma', n)}{\Pr(\text{piv}|\beta; \sigma', n)} \frac{\Pr(s|\alpha; \sigma)}{\Pr(s|\beta; \sigma)}, \tag{45}$$

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 (43) follows from $\Pr(a|\alpha; \sigma) = \Pr(b|\beta; \sigma) = \frac{1}{2} + x$ and $\Pr(a|\beta; \sigma) = \Pr(b|\alpha; \sigma) = \frac{1}{2} - x$.

B Proof of Lemma 4

Lemma 6 *Let $d > 1$. There are $\Delta, \epsilon > 0$, and $\bar{n} \in \mathbb{N}$, so that for any $q_n(\alpha) \in D_n$ with $D_n = [\frac{1}{2} - \epsilon, \frac{1}{2} - \frac{\Delta}{\sqrt{n}}] \cup [\frac{1}{2} + \frac{\Delta}{\sqrt{n}}, \frac{1}{2} + \epsilon,]$ and for any $n \geq \bar{n}$,*

$$q(\beta; \sigma^{(q_n(\alpha), q_n(\beta))}) - q_n(\beta) \quad > \quad 0 \quad \textit{for} \quad q_n(\beta) = \frac{1}{2} + 2\epsilon, \tag{46}$$

$$q(\beta; \sigma^{(q_n(\alpha), q_n(\beta))}) - q_n(\beta) \quad < \quad 0 \quad \textit{for} \quad q_n(\beta) = \frac{1}{2}. \tag{47}$$

Proof. First, we analyze the voter's posteriors about the state when conditioning on the pivotal event, given a strategy with vote shares $q_n(\beta) = \frac{1}{2}$ or $q_n(\beta) = \frac{1}{2} + 2\epsilon$, and $q_n(\alpha) \in D_n$. We slightly abuse the notation by identifying vote share pairs $\mathbf{q}_n$ with strategies.

Take $q_n(\beta) = \frac{1}{2}$. Take $\Delta' > 0$. When the distance of the vote share in α to $\frac{1}{2}$ is at least Δ' multiples of the standard deviation $s_n(\alpha; \mathbf{q}_n) = \frac{\sqrt{q_n(\alpha)(1-q_n(\alpha))}}{\sqrt{2n+1}}$ of the (empirical) vote share distribution, that is, when $\delta_n(\alpha; \mathbf{q}_n) \geq \Delta'$, it follows from (32) that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)} \leq \frac{\phi(\Delta')}{\phi(0)}$.[46] Hence, for any prior $\Pr(\alpha) \in (0,1)$, there is $\Delta > 0$ large enough, so that for any $q_n(\alpha) \in D_n$

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) < \Phi^{-1}(\frac{1}{2}). \tag{48}$$

Take $q_n(\beta) = \frac{1}{2} + 2\epsilon$. For any $q_n(\alpha) \in D_n$, the election is closer to being tied in α, and, given (32), voters become convinced that the state is α; that is,

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) = 1. \tag{49}$$

Now, we analyze the vote share in β under the best response, and establish (46) and (47). Recall (27), which states that

$$\lim_{n\to\infty} q(\beta_n; \sigma_n) = \lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \mathbf{q}_n, n)). \tag{50}$$

We see that (48) and (50) imply that for $q_n(\beta) = \frac{1}{2}$,

$$q_n(\beta) > q(\beta_n; \sigma^{\mathbf{q}_n}) \tag{51}$$

when n is large enough; that is (46) holds. Recall the richness condition (3). Let $\epsilon > 0$ be small enough so that $\Phi(1) > \frac{1}{2} + 3\epsilon$. Then, (49) and (50) imply that for $q_n(\beta) = \frac{1}{2} + 2\epsilon$,

$$q_n(\beta) < q(\beta_n; \sigma^{\mathbf{q}_n}) \tag{52}$$

when n is large enough; that is (47) holds. ■

Now, we state an analogue of the implicit function theorem that does not require any assumptions on partial derivatives.

Lemma 7 *Suppose $f : [0,1] \times [0,1] \to [-1,1]$ is a continuous function with*

$$f(r,0) < 0 \quad \textit{for all } r, \tag{53}$$

$$f(r,1) > 0 \quad \textit{for all } r. \tag{54}$$

Then, there exist continuous functions $\hat{r}, \hat{x} : [0,1] \to [0,1]$ such that $\hat{r}(0) = 0$, $\hat{r}(1) = 1$, and

$$f(\hat{x}(t), \hat{r}(t)) = 0 \quad \textit{for all } r. \tag{55}$$

A proof can be found in Ekmekci *et al.* (2022).

[46] See also Appendix K for a comprehensive lemma on the voter's pivotal inference.

Now, we prove Lemma 4 by an application of Lemma 7. We provide the proof for the case $(x(\alpha), x(\beta)) = (0, 1)$. The proofs for the other cases are analogous.

Consider the function $g_n : \hat{D}_n \to [0,1]$ for $\hat{D}_n = [\frac{1}{2} - \epsilon, \frac{1}{2} - \Delta n^{-\frac{1}{2}}] \times [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$ which maps pairs of vote shares $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ to $q_n(\beta) - q(\beta; \sigma^{\mathbf{q}_n})$. Take a homeomorphism $h_n : [0,1]^2 \to \hat{D}_n$ that maps the left edge to the left edge; that is, $f_n(\{0\} \times [0,1]) = \{\frac{1}{2} - \epsilon\} \times [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$. Further, it maps the lower edge to the lower edge, etc..

Note that g_n is continuous, given the characterization of the best response through (6), (9)- (12), (17), the indifference equations of Lemma 2 (see, e.g., (22)), and since the type distribution H has a continuous density. Lemma 6 implies that the functions $f_n = g_n \circ h_n$ satisfy the conditions of Lemma 7; precisely, the conditions (46) and (47) correspond to (53) and (54). Hence, applying the lemma yields continuous functions $\hat{x}_n, \hat{r}_n : [0,1] \to [0,1]$ so that $h_n(\hat{x}_n, \hat{r}_n) = 0$. In other words, $v_n = h_n \circ (\hat{x}_n, \hat{r}_n)$ maps $t \in [0,1]$ to vote share pairs $\mathbf{q}_n^t \in \hat{D}_n$ that solve (34). Note that $\operatorname{sgn}(q_n^t(\omega) - \frac{1}{2}) = x(\omega)$ for $\omega \in \{\alpha, \beta\}$ and $t \in [0,1]$ since $\mathbf{q}_n^t \in \hat{D}_n$. Further, v_n is continuous as the composition of continuous maps. Finally, note that $q_n^0(\alpha) = \frac{1}{2} - \epsilon$ since $\hat{r}(0) = 0$ and since h_n maps the edge $\{0\} \times [0,1]$ to the edge $\{\frac{1}{2} - \epsilon\} \times [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$. Similarly, $q_n^1(\alpha) = \frac{1}{2} - \Delta n^{-\frac{1}{2}}$. Taken together, these observations finish the proof of Lemma 4 for the case $(x(\alpha), x(\beta)) = (0, 1)$.

C Proof of Lemma 5

Fix $t \in [0,1]$. First, we note that the sequence of vote share pairs $\mathbf{q}_n^t = (q_n^t(\alpha), q_n^t(\beta))$ satisfies the condition $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n^t, n) \in (0,1)$ of Lemma 3: This is because, by construction, the implied vote share under the best response, $q(\beta; \sigma_n^{\mathbf{q}_n^t})$, lies in $[\Psi(0) + \epsilon, \Psi(1) - \epsilon]$; see Lemma 4. Given (27), $\lim_{n\to\infty} q(\beta; \sigma_n^{\mathbf{q}_n^t}) = \Psi(\Pr(\alpha|\text{piv}; \mathbf{q}_n^t, n))$. The continuity and monotonicity of Ψ imply $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n^t, n) \in (0,1)$.

The remainder of the proof follows arguments similar to those in Section 4.3. There, we discussed why the condition $d > 3$ is the critical condition for the severity of the free-rider problem in a large electorate. Much of the proof restates the observations from Section 4.3.

The first observation is that, if $d > 3$, the average precision of a random voter of the interest group is of an order larger than the pivotal likelihood,

$$\lim_{n\to\infty} \frac{\mathrm{E}(x(t)|g)}{\Pr(\text{piv}|\mathbf{q}_n^t, n)} = \infty. \tag{56}$$

for $g \in \{L, C\}$. To see why, recall from Lemma 3 that $\lim_{n\to\infty} \frac{\mathrm{E}(x(t)|g)}{\Pr(\text{piv}|\mathbf{q}_n^t, n)^{\frac{1}{d-1}}} \in \mathbb{R}$. For $d > 3$, it holds $\frac{2}{d-1} < 1$, so that this implies (56) since the pivotal likelihood converges to zero as $n \to \infty$.

The second observation is that the approximation (24) also holds locally,[47]

$$\lim_{n\to\infty} \Pr(\text{piv}|\omega; \mathbf{q}_n^t)(2n+1)s(\alpha; \mathbf{q}_n) \quad = \quad \lim_{n\to\infty} \phi(\delta_n(\omega; \mathbf{q}_n)), \tag{57}$$

where ϕ the density of the standard normal distribution and $\omega \in \{\alpha, \beta\}$. This local approxi-

[47]The local central limit theorem is due to Gnedenko (1948). The version that we apply is the one for triangular arrays of integer-valued variables as in Davis and McDonald (1995), Theorem 1.2.

mation is illustrated in Figure 3.

The assumption $\lim_{n\to\infty} \delta_n(\alpha; \mathbf{q}_n^t) = \lim_{n\to\infty} \frac{|q_n^t(\alpha)-\frac{1}{2}|}{s(\alpha;\mathbf{q}_n^t}$ $\in \mathbb{R}$ of Lemma 5 implies

$$\lim_{n\to\infty} \phi(\delta_n(\omega; \mathbf{q}_n^t)) \in (0, \infty) \tag{58}$$

Let $s_n = s(\alpha; \mathbf{q}_n^t)$ and $q_n = q_n^t(\alpha)$. Note that $((2n+1)s_n)^{-1} = s_n(q_n(1-q_n)^{-1}$.[48] Consequently, (57) together with (58) yields $\lim_{n\to\infty} \frac{\Pr(\mathrm{piv}|\omega;\mathbf{q}_n^t)}{s_n} \in \mathbb{R}_{>0}$. Combining this with (56),

$$\lim_{n\to\infty} \frac{\mathrm{E}(x_n(t)|g)}{s(\alpha; \mathbf{q}_n^t)} = \infty. \tag{59}$$

Recall (31),

$$\begin{aligned} &\lim_{n\to\infty} \frac{q(\alpha; \sigma^{\mathbf{q}_n}) - q(\beta; \sigma^{\mathbf{q}_n^t})}{s(\alpha; \mathbf{q}_n^t)} \\ = &\lim_{n\to\infty} \frac{2(\Pr(L)\mathrm{E}(x_n(t)|L) - \Pr(C)\mathrm{E}(x_n(t)|C))}{s(\alpha; \mathbf{q}_n^t)}. \end{aligned} \tag{60}$$

Lemma 3 implies that $\lim_{n\to\infty} \frac{\Pr(L)\mathrm{E}(x_n(t)|L)}{\Pr(C)\mathrm{E}(x_n(t)|C)} = \frac{W(L,\kappa,\alpha)}{W(C,\kappa,\alpha)}$. The genericity condition $W(L, \kappa, \alpha) \neq W(C, \kappa, \alpha)$ together with (59) and (60) shows

$$\lim_{n\to\infty} \frac{q(\alpha; \sigma^{\mathbf{q}_n^t}) - q(\beta; \sigma^{\mathbf{q}_n^t})}{s(\alpha; \sigma^{\mathbf{q}_n^t})} \in \{\infty, -\infty\}, \tag{61}$$

which is equivalent to (40). Finally, the claim (40) of Lemma 5 follows from (61) since $q(\beta; \sigma^{\mathbf{q}_n^t}) = q_n^t(\beta)$, by construction of $q_n^t(\beta)$; see Lemma 4.

References

Acharya, A. (2016). Information aggregation failure in a model of social mobility. *Games and Economic Behavior*, **100**, 257–272.

Ali, S. N., Mihm, M. and Siga, L. (2018). *Adverse Selection in Distributive Politics*. Working paper.

Austen-Smith, D. and Banks, J. S. (1996). Information aggregation, rationality, and the condorcet jury theorem. *American political science review*, **90** (1), 34–45.

Barelli, P., Bhattacharya, S. and Siga, L. (2019). Full information equivalence in large elections.

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

— (2018). *Condorcet Jury theorem in a spatial model of elections*. Working paper.

Blum, C. and Zuber, C. I. (2016). Liquid democracy: Potentials, problems, and perspectives. *Journal of Political Philosophy*, **24** (2), 162–182.

[48] Recall that $((2n+1)s_n)^{-1}$ the standard deviation of the Binomial distribution of the number of vote shares. Note that $((2n+1)s_n)^{-1} = \Big[(2n+1)(q_n(1-q_n))\Big]^{-\frac{1}{2}} = s_n(q_n(1-q_n))^{-1}$ since $s_n = (\frac{(2n+1)}{q_n(1-q_n)})^{-\frac{1}{2}}$; see (23) and thereafter.

Bond, P. and Eraslan, H. (2010). Strategic voting over strategic proposals. *The Review of Economic Studies*, **77** (2), 459–490.

Bouton, L. and Castanheira, M. (2012). One person, many votes: Divided majority and information aggregation. *Econometrica*, **80** (1), 43–87.

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

Downs, A. *et al.* (1957). An economic theory of democracy.

Duggan, J. and Martinelli, C. (2001). A bayesian model of voting in juries. *Games and Economic Behavior*, **37** (2), 259–294.

Durrett, R. (1991). Probability: Theory and examples, wadsworth & brooks/cole, pacific grove. *MR1068527*.

Eguia, J. X. and Xefteris, D. (2021). Implementation by vote-buying mechanisms. *American Economic Review*, **111** (9), 2811–28.

Ekmekci, M., Heese, C. and Lauermann, S. (2022). *A Generalized Intermediate Value Theorem*. Working paper.

— and Lauermann, S. (2020). Manipulated electorates and information aggregation. *The Review of Economic Studies*, **87** (2), 997–1033.

Feddersen, T. and Pesendorfer, W. (1997). Voting behavior and information aggregation in elections with private information. *Econometrica: Journal of the Econometric Society*, pp. 1029–1058.

Fernandez, R. and Rodrik, D. (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, pp. 1146–1155.

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

Grossman, G. M. and Helpman, E. (2001). *Special interest politics*. MIT press.

Gul, F. and Pesendorfer, W. (2009). Partisan politics and election failure with ignorant voters. *Journal of Economic Theory*, **144** (1), 146–174.

Hardt, S. and Lopes, L. C. (2015). Google votes: A liquid democracy experiment on a corporate social network.

Heese, C. and Lauermann, S. (2019). *Persuasion and Information Aggregation in Elections*. Working paper.

Kartik, N., Van Weelden, R. and Wolton, S. (2017). Electoral ambiguity and political representation. *American Journal of Political Science*, **61** (4), 958–970.

Kim, J. and Fey, M. (2007). The swing voter's curse with adversarial preferences. *Journal of Economic Theory*, **135** (1), 236–252.

Krishna, V. and Morgan, J. (2011). Overcoming ideological bias in elections. *Journal of Political Economy*, **119** (2), 183–211.

— and — (2015). Majority rule and utilitarian welfare. *American Economic Journal: Microeconomics*, **7** (4), 339–375.

Lalley, S. and Weyl, E. G. (2018). *Nash equilibria for quadratic voting*. Working paper.

Ledyard, J. O. (1984). The pure theory of large two-candidate elections. *Public choice*, **44** (1), 7–41.

— and Palfrey, T. R. (2002). The approximation of efficient public good mechanisms by simple voting schemes. *Journal of Public Economics*, **83** (2), 153–171.

Mandler, M. (2012). The fragility of information aggregation in large elections. *Games and Economic Behavior*, **74** (1), 257–268.

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

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

Matějka, F. and Tabellini, G. (2021). Electoral competition with rationally inattentive voters. *Journal of the European Economic Association*.

Meirowitz, A. *et al.* (2006). Designing institutions to aggregate preferences and information. *Quarterly Journal of political science*, **1** (4), 373–392.

Milgrom, P. R. (1981). Rational expectations, information acquisition, and competitive bidding. *Econometrica: Journal of the Econometric Society*, pp. 921–943.

Myerson, R. B. (1998). Extended poisson games and the condorcet jury theorem. *Games and Economic Behavior*, **25** (1), 111–131.

Oliveros, S. (2013). *Aggregation of endogenous information in large elections*. Working paper.

Palfrey, T. R. and Rosenthal, H. (1985). Voter participation and strategic uncertainty. *American political science review*, **79** (1), 62–78.

Pesendorfer, W. and Swinkels, J. M. (1997). The loser's curse and information aggregation in common value auctions. *Econometrica: Journal of the Econometric Society*, pp. 1247–1281.

Razin, R. (2003). Signaling and election motivations in a voting model with common values and responsive candidates. *Econometrica*, **71** (4), 1083–1119.

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

Wilson, R. (1977). A bidding model of perfect competition. *The Review of Economic Studies*, **44** (3), 511–518.

Wit, J. (1998). Rational choice and the condorcet jury theorem. *Games and Economic Behavior*, **22** (2), 364–376.

Yuksel, S. (2021). Specialized learning and political polarization. *International Economic Review*.

Online appendix

D Figure: Information acquisition region

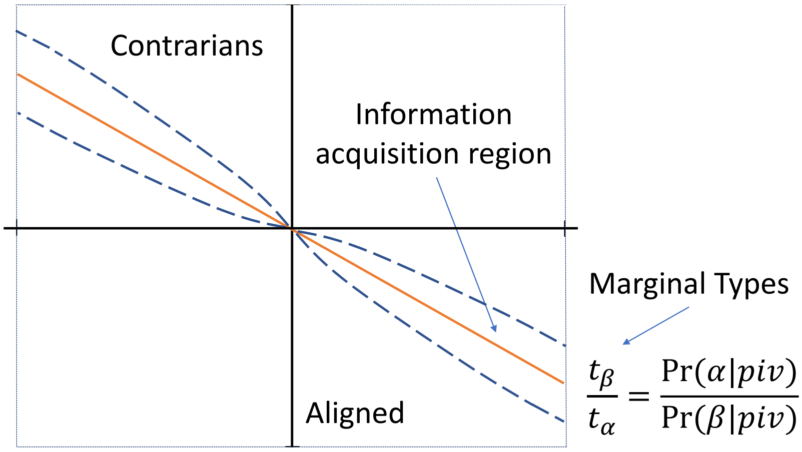


Figure 6: Types in the area between the dashed lines acquire information. Types outside that area stay uninformed.

E Proof of Lemma 2

In the following, for the ease of presentation, we drop the dependence on n and σ' in the notation. Fix a total intensity $k \in (0, \max_t (k(t)]$. The lemma claims that for each interest group $g \in \{L, C\}$, there is an information acquisition interval, given by the boundary types $y_g^-(k)$ and $y_g^+(k)$. We start with the argument for the type $y_L^-(k)$.

First, we characterize when a type t is indifferent between voting A without further information on the one hand and choosing the precision $x = x^*(t; \sigma', n)$ on the other hand. When choosing $x = x^*(t; \sigma', n)$ the expected utility from the policy elected in the pivotal event is given by (13) in α and by (14) in β. Hence, the indifference condition is

$$\begin{aligned} & \Pr(\text{piv})\Big[\Pr(\alpha|\text{piv})(\frac{1}{2}+x)t_\alpha + \Pr(\beta)(\frac{1}{2}-x)t_\beta\Big] - c(x) \\ = \; & \Pr(\text{piv})\Big[\, \Pr(\alpha|\text{piv})t_\alpha + \Pr(\beta|\text{piv})t_\beta\Big]. \end{aligned} \tag{62}$$

Rearranging,

$$\begin{aligned} & \Pr(\text{piv})\Big[(\frac{1}{2}+x)\Big[\Pr(\alpha|\text{piv})t_\alpha - \Pr(\beta|\text{piv})t_\beta\Big] + \Pr(\beta|\text{piv})t_\beta\Big] - c(x) \\ = \; & \Pr(\text{piv})\Big[\Pr(\alpha|\text{piv})t_\alpha - \Pr(\beta|\text{piv})t_\beta + 2\Pr(\beta|\text{piv})t_\beta\Big] \end{aligned} \tag{63}$$

Plugging (15) and (16) into (63),

$$\begin{aligned}
&(\frac{1}{2}+x)c'(x)-c(x)+\Pr(\text{piv})\Pr(\beta|\text{piv})t_\beta \\
= \quad & c'(x)+2\Pr(\text{piv})\Pr(\beta|\text{piv};)t_\beta. \qquad (64)
\end{aligned}$$

We divide by $c'(x)$, rearrange, and use (15) and (16) again,

$$(\frac{1}{2}+x)-\frac{c(x)}{c'(x)}=1+\frac{\Pr(\beta|\text{piv})t_\beta}{\Pr(\alpha|\text{piv})t_\alpha+\Pr(\beta|\text{piv})(-t_\beta)}. \qquad (65)$$

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

$$(\frac{1}{2}+x)-\frac{c(x)}{c'(x)}=1+\frac{-\Pr(\beta|\text{piv})y(t)}{\Pr(\alpha|\text{piv})(1-y(t))+\Pr(\beta|\text{piv})y(t)}. \qquad (66)$$

Since $c(x)=\frac{x^d}{d}$, we have $\frac{c(x)}{xc'(x)}=\frac{1}{d}$ and $x(1-\frac{c(x)}{xc'(x)})=x\frac{d-1}{d}$. Plugging this into (66) and rearranging gives (22), that is,

$$x\frac{d-1}{d}=\frac{1}{2}+\chi(y(t)) \qquad (67)$$

for $\chi(y)=\frac{-\Pr(\beta|\text{piv})y}{\Pr(\alpha|\text{piv})(1-y)+\Pr(\beta|\text{piv})y}$.

Second, the argument from the main text shows that, when n is sufficiently large, there is a unique solution to the indifference equation (67), denoted $y_L^-(k)$ and satisfying $y_g^-(k) < \Pr(\alpha|\text{piv})$. Here, we just fill in the left out algebra. We show that the derivative $\frac{\chi(y_n)}{\partial y_n}$ at $y_n = \Pr(\alpha|\text{piv};\sigma',n)$ stays bounded away from zero, as $n\to\infty$:

$$\begin{aligned}
(\frac{\chi(y_n)}{\partial y_n})_{y_n=\Pr(\alpha|\text{piv})} &= -\frac{1-y_n}{2y_n(1-y_n)}-\frac{2y_n^2(1-y_n)}{(2y_n(1-y_n))^2} \qquad (68)\\
&= \frac{-1}{2y_n(1-y_n)}. \qquad (69)
\end{aligned}$$

The assumption $\lim_{n\in\infty}\Pr(\alpha|\text{piv};\sigma',n)\in(0,1)$ of Lemma 2 implies that the derivative (68) stays indeed bounded away from zero.

Third, the argument analogous to the first two steps shows that there is a unique type $y_L^+(k) > \Pr(\alpha|\text{piv})$ that is indifferent between voting B without further information on the one hand and acquiring information on the other hand, when n is large enough. Putting things together, we see that the types $y_L^-(k)$ and $y_L^+(k)$ mark the boundaries of the interval of all the aligned types with intensity k that acquire information under the best response. The argument for the contrarian types is analogous.

F Proof of Lemma 3

Take the interest group of the aligned types; that is, fix $g=L$ in the following. The proof for the interest group of the contrarian types is analogous. We use that, for the aligned types, there is a one-to-one relation between types t and pairs of thresholds $y(t)$ and total intensities $k(t)$: $t_\alpha = k(t)(1-y(t))$ and $t_\beta = -k(t)y(t)$, given (4) and (5). In the following, we write $t(y,k)$ for the type with $y(t)=y$ and $k(t)=k$, $H(y,k)$ for the joint distribution of y and k, and $H(y)$ and $H(k)$ for the marginal distributions. We evaluate the mean precision

$$\mathrm{E}(x_n(t)|g) = \mathrm{E}(\mathrm{E}(x_n(t)|g,k)|g) \tag{70}$$

iteratively. We start by analyzing $\mathrm{E}(x_n(t)|g,k)$ for a fixed intensity $k = k(t)$.

First, we consider the "intensive margin". Take a type $t = t(y',k)$ who chooses a non-zero precision $x > 0$ under the best response. We show that the type must be arbitrarily close to the marginal type $\bar{y}_n = \Pr(\alpha|\text{piv}; \sigma'_n, n)$ as $n \to \infty$.

Step 1 $\lim_{n\to\infty} y' - \bar{y}_n = 0$.

Proof. Take the interval of types with intensity k that acquire information, $[y_g^-(k), y_g^+(k)]$. It is sufficient to show that the boundary types with $y(t) \in \{y_g^-(k), y_g^+(k)\}$ converge to $\bar{y}_n$ as $n \to \infty$. Take the indifference condition (65) that pins down the threshold of doubt of the boundary type, $y_g^-(k)$. The proof for the other boundary type is analogous. It follows from (15) and (17) that the right hand side of (65) goes to 0 as $n \to \infty$. This implies that $\chi(y_g^-(k)) \to \frac{1}{2}$ for the threshold of doubt $y(t)$ of the boundary type and for $\chi(y) = \frac{-(1-\bar{y}_n)y}{\bar{y}_n(1-y)-(1-\bar{y}_n)y}$ However, this is equivalent to $y_g^-(k) \to \bar{y}_n$. ■

Next, we show that the precision of $t(y',k)$ is asymptotically equivalent to that of the marginal type with the same total intensity k.

Step 2 $x(t(y',k)) \approx x(t(\bar{y}_n,k))$.

Proof. Recall that all types that choose a non-zero precision $x_n(t(y',k)) > 0$, choose the precision $x_n(t(y',k)) = x^*(t(y',k); \sigma'_n, n)$ that solves their first-order condition (17). Using a Taylor approximation of $x^*(t(y',k); \sigma'_n, n)$,

$$x_n(t(y',k)) - x_n(t(\bar{y}_n,k)) = (\bar{y}_n - y') \frac{d}{dy}_{|y=\hat{y}_n(y')} x^*(t(y,k); \sigma'_n, n) \tag{71}$$

for some $\hat{y}_n(y') \in [y', \bar{y}_n]$. Given (15) and (17),

$$\frac{d}{dy}_{|y=\hat{y}_n(y')} x^*(t(y,k); \sigma'_n, n) = x_n(t(\bar{y}_n,k)) M_n(y') \tag{72}$$

for $M_n(y') = \frac{\frac{d}{dy}_{|y=\hat{y}_n(y')} \Big[e(y)\Big]^{\frac{1}{d-1}}}{e(\hat{y}_n(y'))^{\frac{1}{d-1}}}$ and $e(y) = \bar{y}_n(1-y) + (1-\bar{y}_n)y$. By the chain rule of differentiation, $\frac{d}{dy}_{|y=\hat{y}_n(y')} \Big[e(y)\Big]^{\frac{1}{d-1}} = (1-2\bar{y}_n)e(\hat{y}_n(y'))^{\frac{1}{d-1}-1}$. Hence,

$$M_n(y') = \frac{(1-2\bar{y}_n)}{e(\hat{y}_n(y'))}. \tag{73}$$

It follows from Step 1 that $\hat{y}_n(y') \to \bar{y}_n$ as $n \to \infty$ for all y'. Thus,

$$\lim_{n\to\infty} \max_{y':x(t(y',k))>0} |M_n(y')| \quad = \quad \lim_{n\to\infty} |\frac{(1-2\bar{y}_n)}{e(\bar{y}_n)}| = \lim_{n\to\infty} |\frac{(1-2\bar{y}_n)}{2\bar{y}_n(1-\bar{y}_n)}|. \tag{74}$$

Since $\lim_{n\to\infty} \bar{y}_n = \bar{y} \in (0,1)$ by assumption, we have $\lim_{n\to\infty} M_n(y') \in \mathbb{R}$ for all y'. Combining (71) and (72),

$$\begin{aligned} x(t(y',k)) &= x(t(\bar{y}_n,k)) + x(t(\bar{y}_n,k))M_n(y')(\bar{y}_n - y'), \\ \Leftrightarrow \frac{x(t(y',k))}{x(t(\bar{y}_n,k))} &= 1 + M_n(y')(\bar{y}_n - y'). \end{aligned} \tag{75}$$

Finally, (75), the observation that $\lim_{n\to\infty} M_n(y') \in \mathbb{R}$ together with Step 1 implies Step 2. ■

Second, we consider the "extensive margin". We show that the likelihood that a random type with intensity k acquires some information $x > 0$ is asymptotically proportional to the product of precision and likelihood of the marginal type. Denote by $h(t|g,k)$ the density of a type t conditional on $t \in g$ and $k(t) = k$.

Step 3

$$\Pr(\{t : x_n(t) > 0\}|g,k) \approx h(t(\bar{y}_n,k)|g,k)x_n(t(\bar{y}_n,k))e_2(\bar{y}_n,d)$$

for $e_2(y,d) = \frac{4(d-1)}{d}(1-y)y$.

Proof. Using Taylor approximations of the conditional distribution of the threshold of doubt at the threshold $\bar{y}_n$ of the marginal type,

$$\Pr(\{t : x_n(t) > 0\}|g,k) \quad \approx \quad h(t(\bar{y}_n,k)|g.k)(y_g^+(k) - y_g^-(k)), \tag{76}$$

where the types with threshold of doubt $y(t) \in \{y_g^-(k), y_g^+(k)\}$ are the boundary types that are indifferent between no information and choosing the precision $x^*(t;\sigma'_n,n)$ that solves the first-order condition (17). Recall the indifference conditions

$$\frac{1}{2} + \chi(y_g^-(k)) = x^*(t(y_g^-(k),k);\sigma'_n,n)\frac{d-1}{d}, \tag{77}$$

$$\frac{1}{2} + \chi(y_g^+(k)) = -x^*(t(y_g^+(k),k);\sigma'_n,n)\frac{d-1}{d}; \tag{78}$$

see, for example, (22). Taylor approximations of the function χ yield $\chi(y) \approx \chi(\bar{y}_n) + \chi'(\bar{y}_n)(y - \bar{y})$ for $y \in \{y_g^-(k), y_g^+(k)\}$. Since $\chi(\bar{y}_n) = -\frac{1}{2}$, these approximations together with the indifference conditions yield

$$\chi'(\bar{y}_n)\Big[y_g^-(k) - \bar{y}_n\Big] \approx \frac{(d-1)}{d}x^*(t(y_g^-(k),k);\sigma'_n,n), \tag{79}$$

$$\chi'(\bar{y}_n)\Big[\bar{y}_n - y_g^+(k)\Big] \approx \frac{(d-1)}{d}x^*(t(y_g^+(k),k);\sigma'_n,n). \tag{80}$$

Recall (68), that is, $\chi'(\bar{y}_n) = -\frac{1}{2\bar{y}_n(1-\bar{y}_n)}$. Hence, (76)-(80) and Step 2 together imply Step 3. ■

We combine Step 2 and Step 3 to prove the next step.

Step 4 $\mathrm{E}(x_n(t(y,k))|g,k) \approx h(t(\bar{y}_n,k)|g,k)x_n(t(\bar{y}_n,k))^2 e_2(\bar{y},d)$.

Proof. We rewrite the conditional expectation in integral form,

$$\mathrm{E}(x_n(t(y,k))|g,k) = \int_{t:x_n(t)>0} x_n(t)dH(t|g,k). \tag{81}$$

Given Step 2, for any t with $x_n(t) > 0$, we have $x_n(t) = (1 + \epsilon_n(t))x_n(t(\bar{y}_n, k))$ for some sequence $\epsilon_n(t)$ that converges to zero as $n \to \infty$. Hence,

$$\begin{aligned} & \mathrm{E}(x_n(t(y,k))|g,k) \\ = \quad & x_n(t(\bar{y}_n,k)) \Pr(\{t : x_n(t) > 0\}|g,k) + x_n(t(\bar{y}_n,k)) \int_{t:x_n(t)>0} \epsilon_n(t) dH(t|g,k). \end{aligned} \quad (82)$$

Further,

$$\begin{aligned} & \left| \int_{t:x_n(t)>0} \epsilon_n(t) dH(t|g,k) \right| \\ \leq \quad & \int_{t:x_n(t)>0} |\epsilon_n(t)| dH(t|g,k) \\ \leq \quad & \Pr(\{t : x(t) > 0\}|g,k) M_n(y_g^+(k) - y_g^-(k)), \end{aligned} \quad (83)$$

for $M_n = \max_{y' \in [y_g^-(k), y_g^+(k)]} |M_n(y')|$. The first inequality follows from an application of the triangle inequality. For the second inequality, we use that $\epsilon_n(t) = M_n(y')(\bar{y}_n - y')$ given (75). Further, we use that y' and $\bar{y}_n$ lie in the interval $[y_g^-(k), y_g^+(k)]$ of types that choose to acquire information. Step 1 implies $y_g^+(k) - y_g^-(k) \to 0$, as $n \to \infty$. Since $\lim_{n\to\infty} M_n \in \mathbb{R}$ (recall (74) and the observation thereafter), $M_n(y_g^+(k) - y_g^-(k)) \to 0$ as $n \to \infty$. So, $|\int_{t:x_n(t)>0} \epsilon_n(t) dH(t|g,k)| \to 0$, given (83). Combining this with (82),

$$\mathrm{E}(x_n(t(y,k))|g,k) \approx x_n(t(\bar{y}_n,k)) \Pr(\{t : x(t) > 0\}|g,k). \quad (84)$$

Using Step 3,

$$\mathrm{E}(x_n(t(y,k))|g,k) \approx x_n^2(t(\bar{y}_n,k)) h(t(\bar{y}_n,k)|g,k) e_2(\bar{y}_n,d). \quad (85)$$

■

Recall (30) for $t = t(\bar{y}_n, k)$, which states that the marginal type's precision is proportional to a power of the pivotal likelihood and the power $k^{\frac{1}{d-1}}$ of the total intensity. Combining (30) and Step 4,

$$\frac{\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}} \approx \left[h(t(\bar{y}_n,k)|g,k) k^{\frac{2}{d-1}} \right] e_3(\bar{y}_n,d). \quad (86)$$

for $e_3(\bar{y}_n, d) = e_2(\bar{y}, d)(2\bar{y}_n)^{\frac{2}{d-1}}$. In other words, fixing k, the mean precision of a type in the interest group is proportional to the likelihood of the marginal type and the intensity to the power $\kappa = \frac{2}{d-1}$. We integrate over k:

$$\begin{aligned}
\lim_{n\to\infty}\frac{\mathrm{E}(x_n(t(y,k))|g)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}} &= \lim_{n\to\infty}\frac{\mathrm{E}(\mathrm{E}(x_n(t(y,k))|g,k)|g)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}} \\
&= \lim_{n\to\infty}\int_k \frac{h(k|g)\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}}dk \\
&= \int_k \lim_{n\to\infty}\frac{h(k|g)\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}}dk \\
&= \int_k \lim_{n\to\infty} h(k|g)h(t(\bar{y}_n,k)|g,k)k^{\frac{2}{d-1}}e_3(\bar{y}_n,d)dk \\
&= e_3(\bar{y},d)\int_k h(k|g)h(t(\bar{y},k)|g,k)k^{\frac{2}{d-1}}dk \\
&= e_3(\bar{y},d)\int_k h(\bar{y}|g)h(t(\bar{y},k)|g,\bar{y})k^{\frac{2}{d-1}}dk \\
&= e_3(\bar{y},d)h(\bar{y}|g)\mathrm{E}(k^{\frac{2}{d-1}}|g,\bar{y}). \qquad (87)
\end{aligned}$$

The first equality follows from the iterated law of expectations. The second equality restates the conditional expectation as an integral. The third equality follows from an application of the dominated convergence theorem. For the fourth equality, we use (86). The fifth equality follows from $\bar{y}_n \to \bar{y}$ as $n \to \infty$ and since $h(-|g,k)$ is continuous. The sixth equality follows since Bayes law implies $h(k|g)h(t(\bar{y},k)|g,k) = h(t(\bar{y},k)|g,\bar{y})h(\bar{y}|g)$. The last inequality rewrites the integral as a conditional expectation.

Finally, the state-dependent intensity of the limit marginal types $t(\bar{y},k)$ is linear in the total intensity, $t_\alpha = k(1-\bar{y})$; compare to (4) and (5). So, $\mathrm{E}(k^{\frac{2}{d-1}}|g,\bar{y}) = \mathrm{E}(t_\alpha^{\frac{2}{d-1}}|g,\bar{y})(1-\bar{y})^{\frac{2}{d-1}}$. Together with (87) and $\mathrm{E}(x_n(t(y,k))|g) = \frac{1}{h(g)}\int_{t\in g} x_n(t)dH(t)$, this shows (29).

G Proof of Theorem 1: Remaining cases

In the main text, we have provided the proof of the first item of Theorem 1 for the case when $W(\kappa,L,\alpha) < W(\kappa,C,\alpha)$.

Here, we finish the proof of Theorem 1. First, an auxiliary result. This auxiliary result generalizes the observation illustrated in Figure 5b.

Lemma 8 *Let $d > 1$ and $t \in [0,1]$. Take $(\mathbf{q}_n^t)_{n\in\mathbb{N}}$ as in Lemma 4. If $|q_n^t(\alpha) - \frac{1}{2}| \geq \epsilon$ for all n and some $\epsilon > 0$, then,*

$$\lim_{n\to\infty}\frac{|q(\alpha;\sigma^{\mathbf{q}_n^t}) - q_n^t(\beta)|}{s(\alpha;\mathbf{q}_n^t)} = 0. \qquad (88)$$

Proof. Suppose that the vote share of A in α, $q_n^t(\alpha)$, is bounded away from $\frac{1}{2}$ by some positive constant. Given (35), the same is true for $q_n^t(\beta)$. As a consequence, the incentives to acquire information are small. In fact, the pivotal likelihood becomes exponentially small, given (26), and so the precision of any voter type under the best response; see (17). Given so little information acquisition, the vote shares of the best response do not differ by a standard deviation, as $n \to \infty$; that is, (88) holds.[49] ■

[49] Here, recall that the standard deviation of the vote share is of the order of $\frac{1}{\sqrt{n}}$, $s(\omega;\mathbf{q}_n) =$

In the following, let $d > 3$.

First item of Theorem 1. The following case is left:

Case 2 $W(\kappa, L, \alpha) > W(\kappa, C, \alpha)$.

Take $\epsilon, \Delta > 0$ as in Lemma 4 and let the electorate be sufficiently large so that the map v_n and the vote share pairs $\mathbf{q}_n^t$ are defined for the case $(x(\alpha), x(\beta)) = (1, 0)$.

Recall from Lemma 4 that $q_n^0(\alpha) = \frac{1}{2} + \Delta_n n^{-\frac{1}{2}}$. Thus, (35) implies $\lim_{n\to\infty} \frac{\frac{1}{2} - q_n^t(\beta)}{s(\alpha;\sigma^{\mathbf{q}_n^t})} \in \mathbb{R}$. Since $\mathbf{q}_n^0$ solves (34), it holds $q_n^0(\beta) = q(\beta; \sigma^{\mathbf{q}_n^0})$. Hence,

$$\lim_{n\to\infty} \frac{q_n^0(\alpha) - q(\beta; \sigma^{\mathbf{q}_n^0})}{s(\alpha; \sigma^{\mathbf{q}_n^t})} \in \mathbb{R}. \tag{89}$$

The condition of Lemma 5 is satisfied, so that Lemma 5 yields

$$\lim_{n\to\infty} \frac{|q(\alpha; \sigma^{\mathbf{q}_n^0}) - q_n^0(\beta)|}{s(\alpha; \mathbf{q}_n^0)} = \infty. \tag{90}$$

Note that if $W(L, \kappa, \alpha) > W(C, \kappa, \alpha)$, Lemma 3 and (31) together imply that for any $\mathbf{q}_n$ and n large enough,

$$q(\alpha; \sigma^{\mathbf{q}_n}) > q(\beta; \sigma^{\mathbf{q}_n}). \tag{91}$$

Together, (89) - (91) imply that

$$\hat{v}_n(0) = q_n^0(\alpha) - q(\alpha; \sigma^{\mathbf{q}_n^0}) < 0 \tag{92}$$

for n large enough. Recall from Lemma 4 that $q_n^1(\alpha) = \frac{1}{2} + \epsilon$. Given (35), $q_n^1(\beta) \to \frac{1}{2} - \epsilon$ and, given Lemma 8, $q_n(\alpha; \sigma^{\mathbf{q}_n^1}) \to \frac{1}{2} - \epsilon$, as $n \to \infty$. Together,

$$\hat{v}_n(1) = q_n^1(\alpha) - q(\alpha; \sigma^{\mathbf{q}_n^1}) > 0 \tag{93}$$

for n large enough.

Finally, using (92)- (93), an application of the intermediate value theorem shows that there is $t \in (0, 1)$ so that $\mathbf{q}_n^t$ solves (33). Recall that $\mathbf{q}_n^t$ also solves (34), by construction. Thus $\sigma^{\mathbf{q}_n^t}$ is an equilibrium. Further, it must be that $\lim_{n\to\infty} \frac{q_n^t(\alpha) - \frac{1}{2}}{s(\alpha;\mathbf{q}_n^t)} = \infty$ since otherwise (92) holds as we just argued. Hence, also $\lim_{n\to\infty} \frac{\frac{1}{2} - q_n^t(\beta)}{s(\beta;\mathbf{q}_n^t)} = \infty$, given (35). So, the distance of the vote shares to the majority threshold becomes arbitrarily large in terms of standard deviations, which implies that B gets elected in β and A in α as $n \to \infty$, given (24). Thus, the outcome preferred by the aligned is elected in all states, as claimed in the first item of Theorem 1.

Second item of Theorem 1. We present the proof for one case only. In the case presented, the outcome that is preferred by the minority given the prior beliefs is elected in all states. For the other cases, the proof is completely analogous.

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

Case 1 $W(\kappa, L, \alpha) < W(\kappa, C, \alpha)$ *and* $\Psi(\Pr(\alpha)) < \frac{1}{2}$.

Take $\epsilon, \Delta > 0$ as in Lemma 4 and let the electorate be sufficiently large so that the map $\hat{\eta}_n$ and the vote share pairs $\mathbf{q}_n^t$ are defined for the case $(x(\alpha), x(\beta)) = (1, 1)$.

Recall from Lemma 4 that $q_n^0(\alpha) = \frac{1}{2} + \Delta_n n^{-\frac{1}{2}}$. Thus, (35) implies $\lim_{n\to\infty} \frac{\frac{1}{2} - q_n^t(\beta)}{s(\alpha;\sigma^{\mathbf{q}_n^t})} \in \mathbb{R}$. Since $\mathbf{q}_n^0$ solves (34), it holds $q_n^0(\beta) = q(\beta; \sigma^{\mathbf{q}_n^0})$. Hence,

$$\lim_{n\to\infty} \frac{q_n^0(\alpha) - q(\beta; \sigma^{\mathbf{q}_n^0})}{s(\alpha; \sigma^{\mathbf{q}_n^0})} \in \mathbb{R}. \tag{94}$$

The condition of Lemma 5 is satisfied, so that Lemma 5 yields

$$\lim_{n\to\infty} \frac{|q(\alpha; \sigma^{\mathbf{q}_n^t}) - q_n^t(\beta)|}{s(\alpha; \mathbf{q}_n^t)} = \infty. \tag{95}$$

Note that if $W(L, \kappa, \alpha) < W(C, \kappa, \alpha)$, Lemma 3 and (31) together imply that for any $\mathbf{q}_n$ and n large enough,

$$q(\alpha; \sigma^{\mathbf{q}_n}) < q(\beta; \sigma^{\mathbf{q}_n}). \tag{96}$$

Together, (94) - (96) imply that

$$\hat{v}_n(0) = q_n^0(\alpha) - q(\alpha; \sigma^{\mathbf{q}_n^0}) > 0 \tag{97}$$

for n large enough.

Recall from Lemma 4 that $q_n^1(\alpha) = \frac{1}{2} + \epsilon$. We claim that $q_n^1(\beta)$ is multiple standard deviations larger than $q_n^1(\alpha)$ when n is large, that is,

$$\lim_{n\to\infty} \frac{q_n^1(\beta) - q_n^1(\alpha)}{s(\alpha; \mathbf{q}_n^1)} > 0. \tag{98}$$

For the case $x(\beta) = 1$, by construction, $q_n^1(\beta) > \frac{1}{2}$. Then, (35) implies $\lim_{n\to\infty} q_n^1(\beta) = \frac{1}{2} + \epsilon$. Given (27), it must therefore hold that

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n^1) = \Psi^{-1}(\frac{1}{2} + \epsilon). \tag{99}$$

Given (32), $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\mathbf{q}_n^1,n)}{\Pr(\beta|\text{piv};\mathbf{q}_n^1,n)} = \frac{\Pr(\alpha)}{\Pr(\beta)} \frac{\phi(\delta_n(\alpha;\mathbf{q}_n^1))}{\phi(\delta_n(\beta;\mathbf{q}_n^1))}$.[50] Thus, (99), the assumption $\Psi(\Pr(\alpha)) < \frac{1}{2}$ and that Ψ is strictly increasing together imply that

$$\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha; \mathbf{q}_n^1))}{\phi(\delta_n(\beta; \mathbf{q}_n^1))} > 1. \tag{100}$$

This is equivalent to

$$\begin{aligned} \lim_{n\to\infty} e^{-\frac{1}{2}(\delta_n(\alpha;\mathbf{q}_n^1)^2 - \delta_n(\beta;\mathbf{q}_n^1)^2)} &> 1, \\ \Leftrightarrow \lim_{n\to\infty} \delta_n(\beta; \mathbf{q}_n^1)^2 - \delta_n(\alpha; \mathbf{q}_n^1) &> 0, \end{aligned} \tag{101}$$

[50]See also Appendix K for a comprehensive lemma on the voter's pivotal inference.

Now, note that the earlier observation $\lim_{n\to\infty} q_n^1(\beta) = \lim_{n\to\infty} q_n^1(\alpha) = \frac{1}{2}+\epsilon$ implies $\lim_{n\to\infty} \frac{s(\alpha;\mathbf{q}_n^t)}{s(\beta;\mathbf{q}_n^t)} = 1$, recalling that $s(\omega;\mathbf{q}_n^1) = (2n+1)^{-\frac{1}{2}}(q_n(\omega)(1-q_n^1(\omega)))^{\frac{1}{2}}$. Thus, (101) is equivalent to (98), given the definition of $\delta_n(\omega;\sigma_n)$ in (23).

Since the conditions of Lemma 8 are satisfied for $t = 1$, the lemma implies

$$\lim_{n\to\infty} \frac{q(\alpha;\sigma^{\mathbf{q}_n^1}) - q(\beta;\sigma^{\mathbf{q}_n^1})}{s(\alpha;\mathbf{q}_n^1)} = 0. \tag{102}$$

Then, (98), (102), and the property $q_n^1(\beta) = q(\beta;\sigma^{\mathbf{q}_n^1})$ of $q_n^1(\beta)$ together imply

$$\hat{v}_n(1) = q_n^1(\alpha) - q(\alpha;\sigma^{\mathbf{q}_n^1}) < 0 \tag{103}$$

for n large enough.

Finally, using (97) and (103), an application of the intermediate value theorem shows that there is $t \in (0,1)$ so that $\mathbf{q}_n^t$ solves (33). Recall that $\mathbf{q}_n^t$ also solves (34), by construction. Thus $\sigma^{\mathbf{q}_n^t}$ is an equilibrium. Further, it must be that $\lim_{n\to\infty} \frac{q_n^t(\alpha)-\frac{1}{2}}{s(\alpha;\mathbf{q}_n^t)} = \infty$ since otherwise (97) holds as we just argued. Hence, also $\lim_{n\to\infty} \frac{q_n^t(\beta)-\frac{1}{2}}{s(\beta;\mathbf{q}_n^t)} = \infty$, given (35). So, the distance of the vote shares to the majority threshold becomes arbitrarily large in terms of standard deviations, which implies that A gets elected in both states as $n \to \infty$, given (24). Since $\Psi(\Pr(\alpha)) < \frac{1}{2}$, this is the outcome that is preferred by a minority given the prior beliefs. Hence, outcomes are as claimed in the second item of Theorem 1.

H Proof of Theorem 2

Existence of non-informative equilibrium sequences. 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; see (33) and (34). Let $\mathrm{Q}_{\epsilon,n}$ be the set of vote share pairs $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ satisfying

$$|q_n(\alpha) - q_n(\beta)| \leq \frac{1}{n^2}, \tag{104}$$

and

$$|q_n(\omega) - \frac{1}{2}| > \epsilon \tag{105}$$

for $\omega \in \{\alpha,\beta\}$. We claim that when $\epsilon > 0$ is small enough and $n \in \mathbb{N}$ large enough, the best response is a self-map on $\mathrm{Q}_{\epsilon,n}$,

$$\mathbf{q}_n \in \mathrm{Q}_{\epsilon,n} \Rightarrow \mathbf{q}(\sigma^{\mathbf{q}_n}) \in \mathrm{Q}_{\epsilon,n}. \tag{106}$$

Take a sequence of candidate equilibrium vote shares $\mathbf{q}_n \in \mathrm{Q}_{\epsilon,n}$. The first condition (104) implies that the voters do not learn anything about the state from conditioning on being pivotal,

$$\lim_{n\to\infty} \frac{\Pr(\mathrm{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\mathrm{piv}|\beta;\mathbf{q}_n,n)} = 1. \tag{107}$$

To see why, note that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)} = \lim_{n\to\infty} \frac{\phi(\delta_n(\alpha;\mathbf{q}_n))}{\phi(\delta_n(\beta;\mathbf{q}_n))}$, given (32). Further, $\lim_{n\to\infty} \delta_n(\alpha;\mathbf{q}_n) - \delta_n(\beta;\mathbf{q}_n) = 0$, given that $\mathbf{q}_n \in$ satisfies (104).[51] Thus, $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha;\mathbf{q}_n))}{\phi(\delta_n(\beta;\mathbf{q}_n))} = 1$ since the density ϕ of the standard normal is continuous.

The second condition (105) implies that the pivotal likelihood becomes exponentially small as $n \to \infty$, as can be seen from (26). Hence, also the precision of any voter type under the best response becomes exponentially small, given (17), and, further, the distance of the best response's vote share in α to the vote share in β, given (31). We see that the vote shares of the best response again satisfy (104) when n is large. Further, they converge to

$$\lim_{n\to\infty} q_n(\omega) = \lim_{n\to\infty} \Psi(\Pr(\text{piv}|\mathbf{q}_n, n)), \tag{108}$$

given (27). Since Ψ is continuous, (107) and (108) imply $\lim_{n\to\infty} q_n(\omega) = \Psi(\Pr(\alpha))$ for $\omega \in \{\alpha, \beta\}$. Given the assumption $\Psi(\Pr(\alpha)) \neq \frac{1}{2}$ of Theorem 2, there is $\epsilon > 0$ small enough so that (105) holds when n is large enough. We conclude that the best reponse is a self-map on the set $\mathrm{Q}_{\epsilon,n}$ of vote shares satisfying (104) and (105), when n is sufficiently large and $\epsilon > 0$ sufficiently small.

An application of Kakutani's fixed point theorem yields a sequence of equilibrium vote shares in $\mathrm{Q}_{\epsilon,n}$, and any such equilibrium sequence must satisfy (108): As we have just shown, this is a property of the best response to vote shares in $\mathrm{Q}_{\epsilon,n}$. Since any informative equilibrium sequence must, however, satisfy (25), we conclude, that the sequence of equilibrium vote shares corresponds to a non-informative equilibrium sequence.

Properties of non-informative equilibrium sequences. Suppose that an equilibrium sequence is not informative, which means that $\lim_{n\to\infty} \delta_n(\alpha;\sigma_n) - \delta_n(\beta;\sigma_n) = 0$, given the definition of informativeness in Section 3.2.1. The non-informativeness implies that the voters do not learn anything about the state from conditioning on being pivotal,

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

This is because $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv},\sigma_n,n)}{\Pr(\alpha|\text{piv},\sigma_n,n)} = \frac{\phi(\delta_n(\alpha;\sigma_n))}{\phi(\delta_n(\beta;\sigma_n))}$, given (32), and since $\lim_{n\to\infty} \delta_n(\alpha;\sigma_n) - \delta_n(\beta;\sigma_n) = 0$ implies $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha;\sigma_n))}{\phi(\delta_n(\beta;\sigma_n))} = 1$ since the density of the standard normal is continuous. Then, it follows from (27) that

$$\lim_{n\to\infty} q_n(\omega) = \Psi(\Pr(\alpha)) \tag{110}$$

for $\omega \in \{\alpha, \beta\}$. The weak law of large numbers implies that $\lim_{n\to\infty} \Pr(A|\sigma_n, n) = 1$ if $\Psi(\Pr(\alpha)) > \frac{1}{2}$ and $\lim_{n\to\infty} \Pr(B|\sigma_n, n) = 1$ if $\Psi(\Pr(\alpha)) < \frac{1}{2}$.

I Non-monotone type distributions

In the main text, we have provided the analysis for the setting in which preferences are "monotone". When Ψ is non-monotone, there may be multiple $\bar{p} \in (0,1)$ for which $\Psi(\bar{p}) = \frac{1}{2}$. This

[51]Here, recall that $\delta_n(\omega;\sigma_n)$ is the distance of the vote share to $\frac{n}{2n+1}$ in terms of standard deviations $s(\omega;\sigma_n) = \frac{\sqrt{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}}{\sqrt{2n+1}}$; see (23).

motivates the definition of a *local* κ-index, defined in the same way as $W(\kappa, g, \omega)$ in (28), but which depends on the selection of $\bar{p}$ satisfying $\Psi(\bar{p}) = \frac{1}{2}$. For any $\bar{p} \in (0, 1)$ satisfying $\Psi(\bar{p}) = \frac{1}{2}$, and any $\kappa > 0$, the local κ-index of an interest group $g \in \{L, C\}$ at $\bar{p}$ in $\omega \in \{\alpha, \beta\}$ is

$$W(\kappa, g, \omega, \bar{p}) = h(g)h(\bar{p}|g) \underbrace{\mathrm{E}(||t_\omega||^\kappa | g, \bar{p}, \omega)}_{\kappa\text{-measured intensity}}, \tag{111}$$

where $h(\bar{p}|g)$ is the conditional density of the threshold of doubt $y(t) = \bar{p}$ and $\mathrm{E}(-|g, \bar{p})$ the conditional expectation operator that conditions on the types of an interest group g with threshold of doubt $y(t) = \bar{p}$.

For any $\bar{p}$ with $\Psi(\bar{p}) = \frac{1}{2}$ and $\Psi'(\bar{p}) \neq 0$, statements similar to those of Theorem 1 hold, where the local index $W(\kappa, g, \omega, \bar{p})$ takes the role of the κ-index $W(\kappa, g, \omega)$. The formal results are stated in Theorem 3. We omit the proof since it is completely analogous to the proof of Theorem 1.

Theorem 3 *Let $d = \lim_{x \to 0} \frac{c'(x)x}{c(x)} > 3$ and $\kappa = \frac{2}{d-1}$. Take any preference distribution H such that Ψ satisfies the richness condition* (3)*. Take any $\bar{p} \in (0, 1)$ for which $\Psi(\bar{p}) = \frac{1}{2}$ and $\Psi'(\bar{p}) \neq 0$.*

1. *There is an equilibrium sequence in which the policy preferred by the interest group (aligned or contrarians) with the higher local κ-index at $\bar{p}$ is elected with probability converging to* 1 *as $n \to \infty$.*

2. *If $\Psi(\Pr(\alpha)) \neq \frac{1}{2}$, there is an equilibrium sequence in which the policy A is elected with probability converging to* 1 *as $n \to \infty$ if $\Pr(\alpha) > \bar{p}$ and $W(C, g, \alpha, \bar{p}) < W(L, g, \alpha, \bar{p})$ or if $\Pr(\alpha) < \bar{p}$ and $W(C, g, \alpha, \bar{p}) > W(L, g, \alpha, \bar{p})$. There is an equilibrium sequence in which the policy B is elected with probability converging to* 1 *if $\Pr(\alpha) > \bar{p}$ and $W(C, g, \alpha, \bar{p}) > W(L, g, \alpha, \bar{p})$ or if $\Pr(\alpha) < \bar{p}$ and $W(C, g, \alpha, \bar{p}) < W(L, g, \alpha, \bar{p})$.*

One important implication of this generalization is that it may happen that the *order* of the local index of the interest groups varies with $\bar{p}$. Then, there are informative equilibrium sequences for which one interest group wins the election with probability converging to 1 as $n \to \infty$, but also other informative equilibrium sequences in which the other group wins.[52]

J Literature: Attention in electoral competition models

Several papers have analyzed how the electoral competition between politicians is affected by the voters' limited attention to politics (see, for example, Matějka and Tabellini, 2021; Yuksel, 2021). A central finding in Matějka and Tabellini (2021) is that politicians cater more

[52]These results are reminiscent of known results about equilibrium multiplicity for the model with exogenous information: Take the baseline setting from Section 3. If citizens were to receive costless, binary, conditionally i.i.d. signals about the state with an exogenous precision of $0 < x < \frac{1}{2}$ and if Ψ is non-monotone and not constant on any open interval, it is known that there is a multiplicity of equilibrium sequences, some of which do not aggregate information (Bhattacharya, 2013).

to the voters with more extreme ideal policies. The intuition is that these voters will pay more attention and be more responsive to marginal changes in the equilibrium policy since these changes will affect them more strongly (for example, because of utilities with quadratic loss). The politicians anticipate the voter response and consequently put more weight on voters with extreme ideal policies in their decisions. The analysis of the model in this paper brings forward a rival intuition: Voters with extreme prior beliefs, or with preferences that are extremely biased towards one policy, have low incentives to acquire costly information since the information is unlikely to change their opinion about which policy or candidate to vote for. However, election outcomes in the informative equilibria are driven by the costly informational efforts of the citizens (compare to Theorem 1). In this sense, extreme voter types matter little for outcomes in our model. The difference in these observations is due to policies being endogenous, continuous choices of politicians on the one hand, and coarse and exogenous primitives on the other hand. The latter modeling choice follows the tradition of the literature on social choice and on information aggregation in elections.

K Voter inference

Lemma 9 *Consider any sequence of strategies $(\sigma_n)_{n\in\mathbb{N}}$.*

1. *If $\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right| < \lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$, then, $\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}=\infty$.*
2. *If $\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right| > \lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$, then, $\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}=0$.*
3. *If $\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right| = \lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$ and $\delta_n(\alpha;\sigma_n)-\delta_n(\beta;\sigma_n)$ converges in the extended reals $\bar{\mathbb{R}}=\mathbb{R}\cup\{\infty,-\infty\}$, then, $\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}=\lim_{n\to\infty}\frac{\phi(\delta_n(\alpha;\sigma_n))}{\phi(\delta_n(\beta;\sigma_n))}\in\bar{\mathbb{R}}$, where ϕ is the density of the standard normal distribution.*

Proof. Let

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

From (6), $\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}=(k_n)^n$. The function $q(1-q)$ has an inverse u-shape on $[0,1]$ and is symmetric around its peak at $q=\frac{1}{2}$. Therefore, $\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right| < \lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$ implies that $\lim_{n\to\infty}k_n>1$. So, $\lim_{n\to\infty}(k_n)^n=\infty$. Similarly, $\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right| > \lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$ implies that $\lim_{n\to\infty}k_n<1$. So, $\lim_{n\to\infty}(k_n)^n=0$.

It remains to prove the third item. For this, recall the definitions of $s(\omega;\mathbf{q}(\sigma_n))$ and define

$$\begin{aligned}\hat{\delta}_n(\omega;\sigma_n) &= \frac{(q(\omega;\sigma_n)-\frac{1}{2})}{s(\omega;\mathbf{q}(\sigma_n))}\\ &= (2n+1)^{\frac{1}{2}}\frac{q(\omega;\sigma_n)-\frac{1}{2}}{(q(\omega;\sigma_n)(1-q(\omega;\sigma_n)))^{\frac{1}{2}}}.\end{aligned}\tag{112}$$

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

$$
\begin{aligned}
& \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} \\
= \quad & \Big[\frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\Big]^n \\
= \quad & \Big[1-\frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))-q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\Big]^n \\
= \quad & \Big[1-\frac{(\frac{1}{2}+(q(\beta;\sigma_n)-\frac{1}{2}))(\frac{1}{2}-(q(\beta;\sigma_n)-\frac{1}{2}))-(\frac{1}{2}+(q(\alpha;\sigma_n)-\frac{1}{2}))(\frac{1}{2}-(q(\alpha;\sigma_n)-\frac{1}{2}))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\Big]^n \\
= \quad & \Big[1-\frac{(q(\alpha;\sigma_n)-\frac{1}{2})^2-(q(\beta;\sigma_n)-\frac{1}{2})^2}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\Big]^n \\
= \quad & \Big[1-\frac{1}{2n+1}(\frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\delta;\hat{\sigma}_{n_n}(\alpha)^2-\hat{\delta}_n(\beta;\sigma_n)^2\Big]^n. \qquad (113)
\end{aligned}
$$

For the first equality, we used (6). For the fourth equality, we used the third Binomial formula. For the last equality, we used (112).

Case 2 $\lim_{n\to\infty}\delta_n(\alpha;\sigma_n)-\delta_n(\beta;\sigma_n)\in\mathbb{R}$.

Let

$$x_n=\frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}\hat{\delta}_n(\alpha;\sigma_n)^2-\hat{\delta}(\beta;\sigma_n)^2. \tag{114}$$

Then,

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}=\Big[(1-\frac{1}{2n+1}x_n)^n-e^{-\frac{1}{2}x_n}\Big]+e^{-\frac{1}{2}x_n}. \tag{115}$$

Using the Lemmas 4.3 and 4.4 of Durrett (1991), for all $n\in\mathbb{N}$,

$$|(1-\frac{x_n}{(2n+1)})^n-e^{-\frac{1}{2}x_n}|\leq\frac{(\frac{1}{2}x_n)^2}{n^3}. \tag{116}$$

Note that the limit behaviour of $\delta_n(\alpha)-\delta_n(\beta)$ is the same as that of $\hat{\delta}_n(\alpha)-\hat{\delta}_n(\beta)$, that is, $\lim_{n\to\infty}\delta_n(\alpha)-\delta_n(\beta)\in\mathbb{R}$ is equivalent to $\lim_{n\to\infty}\hat{\delta}_n(\alpha)-\hat{\delta}_n(\beta)\in\mathbb{R}$. Since we assumed $\lim_{n\to\infty}\delta_n(\alpha)-\delta_n(\beta)\in\mathbb{R}$, we see that $\lim_{n\to\infty}x_n\in\mathbb{R}$, so that $\frac{x_n^2}{(2n+1)^3}\to 0$ as $n\to\infty$. Consequently,

$$
\begin{aligned}
\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} &= \lim_{n\to\infty}e^{-\frac{1}{2}x_n} \\
&= e^{\lim-\frac{1}{2}x_n} \\
&= \lim_{n\to\infty}\frac{\phi(\delta_n(\alpha;\sigma_n))}{\phi(\delta_n(\beta;\sigma_n))}. \qquad (117)
\end{aligned}
$$

For the equality on the last line we used the definitions of $\delta_n(\omega;\sigma_n)$ and $\hat{\delta}_n(\omega;\sigma_n)$ and that the assumption $\lim_{n\to\infty}\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right|=\lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$ of the third item of Lemma 9 is equivalent to $\lim_{n\to\infty}\frac{q(\alpha;\sigma_n)(1-q(\alpha;\sigma_n))}{q(\beta;\sigma_n)(1-q(\beta;\sigma_n))}=1$ for the equality on the last line; this is true because the function $h(q)=q(1-q)$ is symmetric around $\frac{1}{2}$.

Case 3 $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = \infty$.

Then, for any $x > 0$, there is $\bar{n}(x) \in \mathbb{N}$ so that for all $n \geq \bar{n}(x)$, it holds that $\frac{x_n}{2n+1} \geq \frac{x}{n}$ with x_n given by (114). Thus, given (113),

$$\begin{aligned} \lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} &\leq \lim_{n\to\infty}(1-\frac{x}{n})^n \\ &= e^{-x} \end{aligned} \tag{118}$$

for all $x > 0$. We conclude that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} = 0$. The claim follows since $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} = \lim_{n\to\infty} e^{-\frac{1}{2}(\delta_n(\alpha)^2-\delta_n(\beta)^2)} = 0$, given $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = \infty$.

Case 4 $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = -\infty$.

Then, for any $x < 0$, there is $\bar{n}(x) \in \mathbb{N}$ so that for all $n \geq \bar{n}(x)$, it holds that $\frac{x_n}{2n+1} \leq \frac{x}{n}$. Thus, given (113),

$$\begin{aligned} \lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} &\geq \lim_{n\to\infty}(1-\frac{x}{n})^n \\ &= e^{-x} \end{aligned} \tag{119}$$

for all $x < 0$. We conclude that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} = \infty$. The claim follows since $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} = e^{-\frac{1}{2}(\delta_n(\alpha)^2-\delta_n(\beta)^2)} = \infty$, given $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = -\infty$. ■