).

Interest groups and policy uncertainty: Competition in coordination *

Carl Heese †

Abstract

We study a voting model in which voters have conflicting interests and becoming better informed about policy consequences is costly to individual voters. Outcomes are driven by *competition in coordination*: opposing interest groups compete and each interest group faces the problem to coordinate on supporting the ex-post preferred policy. Coordination is difficult since the information about policy consequences is costly and asymmetric within a group. There are three equilibria and we characterize all, based on an *index* that captures the relative severity of each group's coordination problem. The index depends on the joint distribution of the primitives (cost, intensities, prior beliefs). Correlation and dispersion matter. In particular, we find that outcomes may not align with the majoritarian principle in *any* equilibrium, unlike in related standard settings.

We study a voting model with two central features: agents have conflicting interests and becoming better informed about the consequences of a collective choice is costly to individual voters. This model applies to many situations: in most general elections, national referenda over reforms, and parliamentary votes there are opposing interest groups. Information in the digital age—though in some respects abundant— is costly to filter, pay attention to, and to process.

Outcomes are driven by *competition in coordination*: opposing interest groups compete and each interest group faces the problem to coordinate on supporting the ex-post preferred policy. Coordination is difficult since the information about payoff consequences is costly and asymmetric within a group. We characterize

---

*I am grateful for helpful discussions with Nageeb Ali, Johannes Hörner, Daniel Krähmer, Stephan Lauermann, Philipp Strack, and Thomas Tröger as well as comments from audiences at Bonn, Yale, the CRC TR 224 Conferences 2018 and 2019, CED 2019, SAET 2019, Stony Brook 2019, CMID 20, the World Congress of the Econometric Society 2020, 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, carl.heese@univie.ac.at

how the internal information asymmetries and outcomes depend on the joint distribution of prior beliefs, information cost, and preference intensities, as well as the selected equilibrium.

In particular, we find that outcomes may *not* align with the majoritarian principle in *any* equilibrium and that there are three equilibria with diverging predictions. This stands in stark contrast to the existing result of "information aggregation" in standard voting settings: if each agent holds a small piece of information, majority voting aggregates the information effectively, so that outcomes are full-information-equivalent in *all* equilibria, that is, the outcome preferred by the majority is elected, state-by-state (Feddersen and Pesendorfer, 1997; Bhattacharya, 2013).[1]

Our results highlight how differences in the informational environment between citizens shape political outcomes. Understanding these relations is of particular importance since recent social and political developments have systematic compositional effects on the citizens' information: Information sources in the digital age are individualized; social networks and the fragmentation of the news landscape shape the individuals' political opinions (prior beliefs), access to information (cost), and individual perceptions of what are the important topics (intensities). Our results suggest that these developments may have large impact on political outcomes. We discuss how our model and results speak to a variety of recent public debates in detail in Section 9.

We model a simple majority election. Policy outcomes are binary (A and B) and policy consequences depend on an uncertain state (α or β). Information about the state is costly. Prior beliefs, preference intensities, and information cost are heterogeneous. Under which conditions do election outcomes benefit a majority? When do they help special interest groups or protect a minority's interest?

Voters sharing a common interest under full information are an interest group. Central to the analysis is that each interest group faces a coordination problem when the state is unknown, that is, when policy consequences are uncertain. Take a reform of the public education system and the interest group of citizens who are in favor if it benefits children from lower income households (state α) and otherwise not (state β). Each citizen's ranking of policies depends on her private

[1]This existing work has considered settings with conflicts of interest, like this paper does. But, in such settings, it has assumed exogenous, not costly, and i.i.d. information conditional on the pay-off relevant state. (Bhattacharya, 2013). See also Feddersen and Pesendorfer (1997) for earlier work showing that elections aggregate information effectively. We discuss the existing work on information aggregation in elections in detail in Section 8.1, including Martinelli (2006).

belief about the reform. There may be conflicts of opinion. In the extreme, half of the citizens believe that the reform is beneficial, while the other half believes the opposite. Voting according to personal beliefs, the interest group mis-coordinates and their votes cancel out each other entirely. Were the citizens fully informed about policy consequences, they would coordinate perfectly on voting for their preferred policy. What matters for how well coordinated an interest group acts is the informedness of its constituents.[2] The voters' heterogeneous beliefs and the incentives to become politically informed take center stage.

Equilibrium multiplicity. We show that the *strategic interdependence* of the costly information acquisition creates an equilibrium multiplicity. It is well-known that the incentive of a voter to acquire costly information is weakened since she may "free-ride" on others. Importantly, how severe a voter's incentive to free-ride is, depends on the voter' s expectations about the closeness of the election. When citizens believe that the election will not be close, they will stay largely uninformed because they know that their individual vote is unlikely to affect the election outcome. Conversely, when voters believe that the election will be close, information about policy consequences is more valuable to the voters. Critically, when other citizens vote in an informed manner, this may change a voter's belief about the closeness of the election, and we show that this may dampen, but also *spur* the information acquisition of the given voter. These complementarities drive the existence of multiple equilibria. In particular, there are multiple *informative equilibria*, that is, equilibria in which the endogenous information acquisition shapes outcomes. We characterize the condition of their existence.[3]

Equilibrium characterization. We completely characterize all equilibria of large elections. The analysis of voting settings with costly information poses technical challenges. The classical result states the paradox that costly information

[2] An important precedent to our work is Fernandez and Rodrik (1991) with a related observation about asymmetric information and coordination: they show that policy uncertainty may lead to the rejection of efficient reforms since a benefiting majority may be divided by conflicting opinions. Relative to Fernandez and Rodrik (1991), this paper provides an analysis how interest groups *compete in coordination* when coordination of a group is difficult because of internal information asymmetries. Further, the information asymmetries in our setting are endogenous. Bouton and Castanheira (2012) point at a distinct coordination problem: elections with more than two candidates may not align with the majoritarian principle because of majorities that may be *divided* between two similar candidates.

[3] Formally, what matters for the existence is how fast cost goes to zero when a voter chooses an arbitrarily uninformative signal. The condition identified is similar to a sufficient condition identified by Martinelli (2006) in a symmetric setting with one interest group.

acquisition is inconsistent with voter rationality (Downs, 1957). More recent work, e.g., by Martinelli (2006), has considered richer environments and shown that large elections may lead to full-information equivalent outcomes even when information is costly. This more recent work imposes symmetry conditions on voter preferences and beliefs to make the analysis of equilibrium outcomes tractable; in particular, all voters share a common interest.[4] In this paper, we work in a setting without symmetry, allowing for general distributions of preference intensities, cost types, and prior beliefs. This way, we can study how internal information asymmetries arise and how they drive the coordination of competing interest groups. Doing so demands new techniques to handle the multidimensionality of the voter strategies.

Close elections as an informational phenomenon. To this end, a key observation is that vote shares are split between the two policies (arbitrarily) close to 50-50 in any informative equilibrium of an (arbitrarily) large election. This observation is important in itself since it rationalizes the frequent occurrence of close election races as an informational phenomenon.[5] It it technically relevant: in any informative equilibrium, the *marginal* types, i.e. the types that are indifferent without further information about the state, and those close-by, drive the information acquisition. We show that the closeness of the election pins down the indifferent marginal types, allowing us to use "local" techniques along the curve of indifferent types, to tackle the multidimensionality. Besides this, we provide several further technical innovations to handle the multidimensionality.[6]

The index. We provide a characterization of the endogenous average precision of the citizens of an interest group (the *index*). As argued, we may think of this index as relating to the severity of the group's coordination problem. The index is given by a measure of the marginal types of the group. It depends on the likelihood of the marginal types in the group and on the conditional distribution of cost and preference intensities. Not only first-order moments, but also dispersion matters. In a baseline in which prior beliefs and information cost are symmetric, the index interpolates between two extremes, depending on the elasticity of the information cost at zero precision, $\lim_{x\to 0}\frac{c'(x)x}{c(x)}$, which varies how "cheap" infor-

[4]See e.g. Martinelli (2006) and Proposition 5 in Oliveros (2013)

[5]The intuition is this: Given an equilibrium in which some citizens find it beneficial to get informed, the citizens must expect the election to be sufficiently close. We show that otherwise the citizens' expected benefits from becoming better informed are too small so that the information cost outweigh.

[6]For example, we represent equilibria in a compact way as fixed points in the space of vote shares (see Section 4.4). To use this representation effectively, we apply the local central limit theorem (Gnedenko, 1948) to study the limit behaviour of individual best responses as functions of the vote shares only (see Section 4.2).

mation of low precision is:[7] When information is arbitrarily cheap, the index is purely ordinal and proportional to the likelihood of the marginal types. When information cost are intermediate, the index corresponds to utilitarian welfare. Generally, if information is more expensive, preference intensities matter more.

There are two informative equilibria, if they exist. In the "most informative" equilibrium, the policy preferred by the interest group with the higher index is elected state-by-state. This result implies that even arbitrarily large interest groups may lose an election. In such equilibria, the large group remains relatively misinformed and mis-coordinates the votes more strongly than a smaller interest group with opposed interests. The result also shows that small groups with high stakes are more likely to overcome the dominance of an opposed majority if information is more expensive since then preference intensities matter more. In the other informative equilibrium, both the exogenous prior information and the endogenous information play a role. The same policy is elected in both states. Which policy this is, depends on the which interest group has the highest index and if the prior information is biased towards α or β. Again, the endogenous information matters for how well each interest group coordinates, but voting behavior is biased in the direction of the prior information.

Additionally, there is an equilibrium strategy in which almost all voters vote according to their prior beliefs. This equilibrium describes a "tyranny of the uninformed" in the following sense: Whichever policy is preferred by a majority given the prior beliefs, this policy is elected in all states. Intuitively, if one party receives the support of more than 50 % in expectation given "voting according to the prior beliefs", the election is not expected to be close to being tied. We show that, as a consequence, almost all voters shy the cost of additional information and vote according to their prior beliefs also under the best response.

Further connections to the literature. The results contribute more generally to the understanding of (a) the effects of asymmetric information in politics, (b) the political competition between opposed interest groups. Section 8.2 provides a discussion of some related existing work. In particular, we discuss other voter competition models, including models of vote-buying (Eguia and Xefteris, 2018) and costly voting (Krishna and Morgan, 2011, 2015), and the literature on "special interest politics" (as pioneered by Olson (1965) and Tullock (1983)).

In Section 1, we provide intuition for some central ideas with an example. In

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

Section 2, we present the model with general type distributions. In Section 3, we analyze a baseline setting. We present the main results for the baseline setting in Section 4 and for the general setting in Section 5. In Section 6, we study the non-informative equilibrium sequences. Section 7 contains further discussion and extensions. In Section 8, we discuss the literature.

1 Example

An example illustrates how conflicts of opinion may weaken the electoral prospects of an interest group and that the marginal types, that is the voter types that are indifferent without additional information about the state, are critical since their incentives to acquire information are the largest.

There are $2n + 1 \geq 3$ voters (or citizens). With probability $0 < \lambda < \frac{1}{2}$, a voter is *contrarian* and prefers the reform A over the status quo B in β and B over A in α. With probability $1 - \lambda$, a voter is *aligned* and prefers A in β and B in α. The aligned (contrarians) receive a utility of $k_L > 0$ ($k_C > 0$) when their preferred policy is elected, and zero otherwise.

Aligned and contrarian voters are of three types: a "reform leaning" type, a "reform skeptical" type and an "unbiased" type. These types differ in their prior beliefs. The reform-leaning aligned (contrarian) type believes that the likelihood of state α is $\frac{3}{4}$ ($\frac{1}{4}$), the reform-skeptical aligned (contrarian) type believes that the likelihood of state α is $\frac{1}{4}$ ($\frac{3}{4}$), and the unbiased types hold a symmetric prior belief. The aligned and contrarian types are distributed conditionally independently, with the reform-leaning and the reform-skeptical type being equally likely. Each voter receives a private, binary signal $s \in \{a, b\}$ about the state, drawn independently of the signals and types of the other voters.

The timing is as follows: 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 $d > 0$. 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 following sketches an argument showing that, when the electorate size is sufficiently large, there is an equilibrium in which the types with extreme prior beliefs, that is, the reform-leaning and the reform-skeptical types, choose to receive uninformative signals and the unbiased types choose a non-zero precision.[8] For such a candidate strategy, one can show that the symmetry of the signals and types implies that the vote of a single citizen affects the election outcome with the same likelihood in both states.[9] So, if a type votes A, she expects to tip the election outcome from B to A with the same probability in both states.[10] Doing so benefits her in one state and comes with a utility loss in the other. What matters for the voter‘s decision is her prior belief about the state, that is, about which event is more likely.

Conflicts of opinion. The aligned and contrarians suffer from conflicts of opinion. The skepticals within each group hold the prior belief that the state in which they prefer the status quo has a probability of $\frac{3}{4}$. So, a skeptical believes that voting for the reform A will shift the outcome from B to A more often in the state in which she does not benefit from it. Skepticals thus strictly prefer to vote for the status quo without additional information about the state. Analogously, the reform-leaning strictly prefer to vote for the reform without additional information. Any signal that could turn around these types' strict preference would have to have a sufficiently high precision $x > \bar{x}$ for some $\bar{x} > 0$.[11] 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, hence, is not worth the cost either. The reform-leaning and reform-skeptical types choose to receive an uninformative signal. Since the reform-leaning and reform-skeptical types are equally likely, their votes split $50 - 50$ between both policies in expectation, effectively canceling out each other.

Marginal types and coordination through information. An unbiased type who votes A expects to benefit from tipping the election from B to A with the same probability as she expects to lose from it, given the uniform prior. The same argument for a B-vote shows that she is indifferent between voting A or B. For

[8]For illustration: take e.g. $k_L = 1$, $k_C = 2$, $\lambda = \frac{1}{3}$, and $d = 3$. One can show that such an equilibrium exists as long as $n \geq xxx$.

[9]From an ex-ante perspective, a single citizen's vote is decisive for the election outcome with a probability smaller than 1, that is, only in the event in which the votes of the other citizens split into n votes for A and n votes for B.

[10]Similarly, voting B tips the election outcome from A to B with the same probability in both states.

[11]Precisely, the precision x has to be larger than $\frac{1}{4}$, so that $\Pr(\alpha|b; x) < 1/2 < \Pr(\alpha|a; x)$

the unbiased types, the benefit of a higher precision $x > 0$ is to tip the election into the preferred direction more often in each state. When choosing the precision x, she receives the “correct” signal with probability $\frac{1}{2} + x$ and follows it, voting for the preferred policy with probability $\frac{1}{2} + x$ in each state. So, in each state, she tips the election into the preferred direction with a probability proportional to $(\frac{1}{2} + x)$. Any time, she tips the election, she gains a utility of k_g. The unbiased types choose the precision $x_g > 0$ equating marginal benefits and marginal cost,

$$\begin{aligned} & c'(x_g) = k_g e \\ \Leftrightarrow \quad & x_g = (k_g e)^{\frac{1}{d-1}}, \end{aligned} \tag{1}$$

where $e < 1$ is a constant capturing that the type’s vote is not always decisive for the election outcome. The larger their preference intensity k_G, the more informed the unbiased aligned and contrarian types will choose to be and the better they coordinate on voting for their preferred policy in each state respectively. What matters for how the vote shares of A and B compare in each state, is the likelihood of the unbiased types within each group and the preference intensities.

Numerical illustration. One can show thatm there is a unique equilibrium in this example.[12] Figure 1 shows the equilibrium outcomes when varying the information cost and the preference intensities.[13] 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. Going down the rows, the intensity k_C of the contrarians increases, and so does the likelihood of their preferred outcome. Comparing the column for $d = 2$ and for $d = 3$, we see how 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”.[14]

[12] This has to do with the symmetry between reform-leaning and reform-skeptical. Generically, there are multiple equilibria.

[13] Varying the likelihood of the unbiased types effects similar to changes in the intensities.

[14] Here, for $k_C > 2$, the contrarian types acquire full information, which explains why outcomes are the same for $k_C = 3$ and $k_C = 4$.

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: 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.

2 Model

The model generalizes the example by allowing for general type distributions. Besides that, the voting game is as in 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 according to 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, i.e., $\mu(t, s)$ is the probability that a voter of type t with signal s votes for A. We only consider non-degenerate strategies.[15] 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 common 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 idiosyn-

[15] 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.

cratic differences captured by r), where a higher d means that information of low precision is "cheaper".[16]

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 isolates the effect of heterogeneity in preference intensities only, making results particularly comparable to existing work, and simplifies the exposition. In Section 5, we analyse the general setting. For this, we leverage that for any joint distribution of types, there is a auxiliary distribution which 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 way, we can study the effect of heterogeneity in cost and beliefs, simply by leveraging the analysis for the restricted setting.

Preference types. Slightly abusing the notation, we denote by H the distribution of (t_α, t_β), and assume in the following that it has a continuous density on its support and a connected and compact support $K_\alpha \times K_\beta \subseteq \mathbb{R}2$ with $(0, 0)$ in its interior.[17] Figure 2 shows the area of possible preference types. Voters having types t in the north-east quadrant prefer A for all beliefs and voters having types t in the south-west quadrant always prefer B (*partisans*). Voters having types t in the south-east quadrant prefer A in state α and B in β (*aligned voters*), and voters having types t in the north-west quadrant prefer B in state α and A in β (*contrarian voters*). To simplify the exposition, in the rest of the paper, we only consider strategies σ where the partisans use the (weakly) dominant strategy to vote for their preferred policy.[18] Given that the preference distribution has full support, this implies $q(\omega; \sigma) \in (\epsilon, 1 - \epsilon)$ for some $\epsilon > 0$.

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

[17]The assumption that $(0, 0)$ is in the interior of the support ensures that there are types with conflicting interests.

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

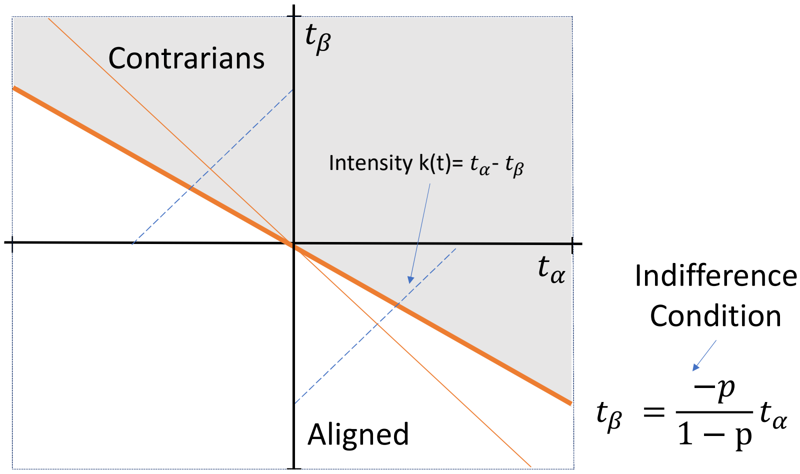

Figure 2: For any given belief $p = \Pr(\alpha) \in (0,1)$, the set of types t that are indifferent given p is given by $t_\beta = \frac{-p}{1-p} t_\alpha$. Voter types north-east of the indifference line (shaded area) prefer A given p. Contrarian and aligned types are uniquely identified by their *(total)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

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

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 Φ: 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 $\Phi(p)$. The indifference set has a slope of $\frac{-p}{1-p}$ and an increase in p corresponds to a clockwise rotation of it. Given that H has a continuous density, Φ is continuously differentiable in p.

We assume that

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

such that the median-voter preferred outcome is A in α and B in β. In particular, this excludes the (trivial) cases when there is a majority of partisans for one policy in expectation. We also assume that Φ is strictly monotone.[19] The non-monotone case is discussed in Section 7.2. Henceforth, I will call distributions H for which Φ satisfies (3) and is strictly increasing *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

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

and

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

[19]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).

I also refer to $q(\omega;\sigma)$ 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

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

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

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

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

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

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

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

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

3.1.2 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

$$\Pr(\text{piv}|\sigma', n)\Pr(\alpha|\text{piv}; \sigma)(\frac{1}{2} + x)t_\alpha \tag{15}$$

in state α, and

$$\Pr(\text{piv}|\sigma', n)\Pr(\beta|\text{piv}; \sigma)(\frac{1}{2} - x)t_\beta \tag{16}$$

in state β, where we used Lemma 1 and that the utility from B is normalized to zero.[20] Therefore, summing (15) and (16) and taking the derivative, the marginal benefit of a higher precision x is

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

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

$$c'(x) = MB\Big[\sigma', n\Big], \tag{18}$$

which has the unique solution

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

[20] 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.

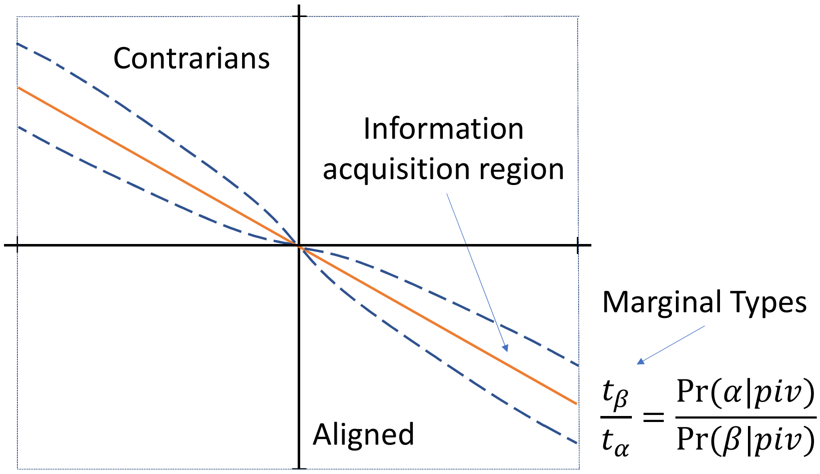

Figure 3: Types in the area between the dashed lines acquire information. Types outside that area stay uninformed.

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

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

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

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 (11) - (14), and called the *marginal types*. Lemma 2 shows that, for each total intensity $k = k(t) \in [0, \sqrt{2}]$, only types in a certain interval around the marginal types acquire information.

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 $k \in (0, 2)$ and any $g \in \{L, C\}$ there are $y_g^-(k) < \Pr(\alpha|\text{piv}; \sigma', n) < y_g^+(k)$ for such that for any best response $\sigma = (x, \mu)$ to σ' and any type $t \in g$ with $k(t) = k$,*

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

Note that for $d \leq 1$ marginal cost are bounded away from zero, $c'(0) > 0$. Thus, (18) has no solution when n is large and all types stay uninformed.

Figure 3 illustrates the functions y_g^- and y_g^+. To see the result formally, take e.g. a boundary type t with $y(t) = y_L^-(k)$. We can rewrite the indifference condition,

$$\chi(y(t)) + \frac{1}{2} = \frac{(d-1)}{d} x^*(t) \tag{23}$$

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 A. The left hand side captures the bias towards policy A without additional information: the bias is zero at the indifferent type's threshold $\bar{y} = \Pr(\alpha|\text{piv};\sigma,n)$. Intuitively, for an aligned type, the lower the threshold $y(t9$, the higher the bias towards policy A. So, the left hand side increases in the distance of $y(t) < \bar{y}$ to $\bar{y}$. Since (19) implies that the optimal informative signal $x^*(t)$ varies little with y when n is large (the derivative is proportional to the pivotal likelihood), the indifference equation has a unique solution $y_L^-(k) < \bar{y}$.

Figure 3 suggests that types with a larger total intensity are more likely to acquire information, ceteris paribus. In fact, (23) shows that the likelihood that a type acquires information increases in the total intensity in the same way that the precision $x^*(t)$ increases in it; see (19).

The following studies the equilibria of the election as the number of citizens $2n+1$ grows without bound. Considering a large number, 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 $\mathbf{q}(\sigma_n) = (q(\alpha;\sigma_n), q(\beta;\sigma_n))$. Then,

$$\delta_n(\omega) = \frac{q(\omega;\sigma_n) - \frac{n}{2n+1}}{s(\omega;\sigma_n)}. \tag{24}$$

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

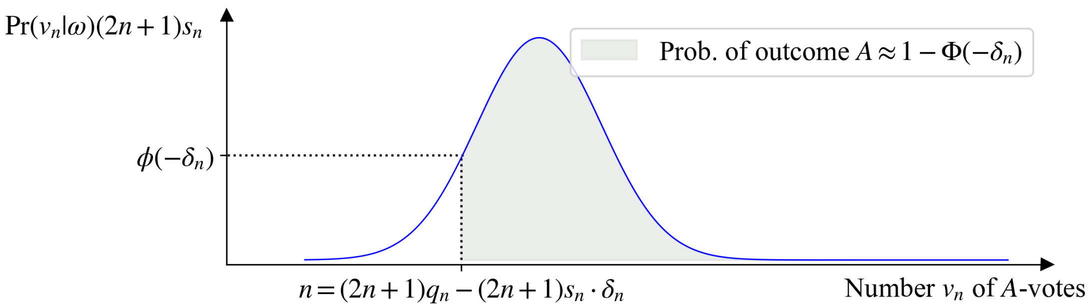

Figure 4: Illustration of the Normal approximation of the Binomial distribution of the number of A-votes v_n with mean $(2n+1)q_n$ for $q_n = q(\omega;\sigma_n)$ and with standard deviation $(2n+1)s_n = ((2n+1)(q_n(1-q_n))^{\frac{1}{2}}$ for $s_n = s(\omega;\sigma_n)$.

$\omega \in \{\alpha,\beta\}$, where $s(\omega;\sigma_n)^{-1} = \sqrt{\frac{(2n+1)}{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}}$.[21] A normal approximation of the distribution of the number of A-votes (see Figure 4) shows that, as $n \to \infty$, the probability that A gets elected in ω converges to[22]

$$\lim_{n\to\infty} \Pr(A|\omega;\sigma_n) = \lim_{n\to\infty} 1-\Phi(-\delta_n(\omega)), \tag{25}$$

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

An equilibrium sequence is *informative* if $\lim_{n\to\infty}\delta_n(\alpha)-\delta_n(\beta) \neq 0$. Informativeness captures that the aggregate effect of the voters' information acquisition on vote shares is large enough as to impact outcomes. Given (25), it is a necessary condition for the outcome distribution to be different in the two states.

[21]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 $s(\omega;\sigma_n)$.

[22]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 $\lim_{n\to\infty} 1-\Phi(\frac{(2n+1)(\frac{n}{2n+1}-q_n)}{((2n+1)q_n(1-q_n))^{\frac{1}{2}}}) = \lim_{n\to\infty} 1-\Phi(-\delta_n(\omega))$.

3.2.2 Close elections: An equilibrium Outcome

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

$$\lim_{n\to\infty} q(\omega;\sigma_n^*) = \frac{1}{2} \tag{26}$$

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

Formally, a voters' individual incentives to acquire information depend on the pivotal likelihood; recall e.g. the cost-benefit analysis for the optimal (interior) precision, (17). A Stirling approximation of the pivotal likelihood yields[23]

$$\Pr(\text{piv}|\omega;n) \approx 4^n (n\pi)^{-\frac{1}{2}} \Big[q(\omega;\sigma_n)(1-q(\omega;\sigma_n))\Big]^n, \tag{27}$$

which implies that the pivotal likelihood is exponentially small unless (26) holds.[24] If (26) does not hold in *any* state, given (17), voters acquire exponentially little information under the best response so that the difference of the vote shares in the two states, measured in standard deviations, goes to zero, i.e. $\lim_{n\to\infty}\delta_n(\alpha) - \delta_n(\beta) = 0$.

The reason why the election is close in all and not just in one state (i.e. (26)) is that the likelihood that a random citizen votes A is asymptotically the same across states. This is because, for any strategy sequence, the signal precision of a random voter is of an order weakly smaller than $n^{-\frac{1}{2(d-1)}}$, given (27), (17), and (19). So, the definition (2) together with (9) and (10) implies

$$q(\omega;\sigma_n^*) \to \Phi(\Pr(\alpha|\text{piv};\sigma_n^*,n)) \tag{28}$$

for both states ω.

3.2.3 Limit marginal types

The closeness of elections, (26) 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

[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 (8) yields $\Pr(\text{piv}|\omega;n) \approx 4^n (n\pi)^{-\frac{1}{2}} (q(1-q))^n$ for $q = q(\omega;\sigma_n)$.

[24] This is because the function $q(1-q)$ takes the maximum $\frac{1}{4}$ at $q=\frac{1}{2}$ only.

necessarily satisfies

$$\lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \sigma_n^*, n)) = \frac{1}{2}, \tag{29}$$

given (134). Since Φ is continuous and 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 $\Phi(\bar{y}) = \frac{1}{2}$.

4 Main results: Baseline setting

The literature 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.[25] 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 larger group wins.

In our setting, outcomes may not align with the majoritarian principle. Outcomes are driven by how well the interest groups coordinate internally. Internal incentive constraints, due to information being costly, drive the asymmetries in information within a group and thereby how well each group coordinates on voting for the ex-post preferred policy.[26]

Theorem 1 characterizes when informative equilibrium sequences exist. Further, it characterizes *all* informative equilibrium sequences, based on a measure (the *index*) that we will later relate to the endogenous information of each interest group (Lemma 3).

Recall that d is the elasticity of the cost functions, t_ω is the type's utility from the reform in ω and r the cost type. In the following, we denote by $\mathrm{E}(-|g)$ and $f(-|g)$ 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 $f(g)$ for the unconditional likelihood and $\mathrm{E}(-|y)$ and $f(-|y)$ when conditioning on the set

[25] See e.g. Feddersen and Pesendorfer (1997); Austen-Smith and Banks (1996).

[26] The importance of internal incentive constraints for the coordination of competing interest groups has been stressed, in particular, by the literature on special interest politics which we discuss in Section 8.

of types with threshold of doubt $y(t) = y$, et cetera.

$$W(\kappa, g, \omega) = \underbrace{f(g)f(\bar{y}|g)}_{\text{likelihood of limit marginal types}} \quad \underbrace{\mathrm{E}(||t_\omega||^\kappa | g, \bar{y}, \omega)}_{\kappa\text{-measured intensity}}, \tag{30}$$

Since all marginal types have the same *relative* intensities across the states, i.e. $\frac{t_\alpha}{t_\beta} = -\frac{1-\bar{y}}{\bar{y}}$, the index differs only by a scalar across states, $W(\kappa, g, \alpha) = -\frac{1-\bar{y}}{\bar{y}} W(\kappa, g, \beta)$.

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 Φ is strictly monotone and the richness condition* (3) *holds:*

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

Welfare. In the benchmark setting common prior and cost type, the index has a compelling interpretation in terms of welfare.[27] 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, when the informational regime is so that information of low precision is "cheaper", that is, d is lower and κ is higher,[28] then preference intensities matter more.

Full-Information Outcomes. Theorem 1 implies that, when the contrarians have a higher index, there is no equilibrium in which the full-information outcome (A in α, B in β) is chosen in both states as the electorate grows large. This stands in stark contrast to the existing literature.

[27] In Section 5.2, we define a generalized version of the group index for the setting from Section 2 with heterogeneity in prior beliefs and cost types. There, we explain how prior and cost heterogeneity drives a wedge between the index and the welfare interpretation given here.

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

Factors of political power. In the first equilibrium in which the group with the higher index wins the election, the index is a measure of political power. Here, the parameter κ captures exactly how intensities substitute with the mass of (marginal) citizens of a group in determining the political power.

Intuition for Theorem 1. Two observations from the example of Section 1 are useful to gain intuition. The first observation is that, in the example, given (1), the marginal types of an interest group receive the "correct signal" with a probability proportional to $\frac{1}{2} + x_g$ with x_g proportional to the power k_G^κ of their intensities, with $\kappa = \frac{1}{d-1}$. The second observation is that, in the example, the votes of those who acquire no information (the types that are not marginal) cancel out each other completely. Together, the two observations imply that, state-by-state, the policy preferred by a given interest group receives a larger vote share in expectation if the κ-index is higher.

We will establish similar observations for the benchmark setting. First, the average precision chosen by the voters of an interest group is proportional to the κ-index with $\kappa = \frac{2}{d-1}$, as $n \to \infty$ (Lemma 3).[29] Second, *mis-coordination of the uninformed* is a general phenomenon: as observed in Section 3.2.2, in *any* informative equilibrium sequence, the voters who stay uninformed necessarily mis-coordinate their votes in a way, so that the expected vote shares are relatively close to 50-50. Only this closeness creates high enough incentives for some voter types (close to the marginal types) to costly acquire information; compare to the discussion in Section 3.2.2. However, unlike in the example, the votes of the uninformed do not "cancel out each other" entirely. There are two equilibria and they differ in how strongly the uninformed mis-coordinate. In the first equilibrium of Theorem 1, mis-coordination is stronger and the outcomes are maximizing the κ-index when $n \to \infty$, similar to the example. In the second equilibrium, the asymmetry of the the preferences given the prior beliefs, $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$, creates a bias towards one of the policies and the same policy is elected in both states.

In the following, first, we establish Lemma 3 in Section 4.1. In Section 4.2 and Section 4.3, we discuss the economic forces driving the existence of multiple informative equilibria and explain the relevant condition from Theorem 1 ($d > 3$). In Section 4.4, we prove Theorem 1.

[29]The relevant κ is different in the example since the example the preference distribution has atoms at the marginal types. We omit a detailed explanation at this point, but refer to the proof of Lemma 3 for a detailed understanding.

4.1 The endogenous information of the interest groups

For any sequence of strategies with interior limit marginal types and for each interest group, the κ-index is proportional to the expected size of the group times the average precision acquired by a random voter of the group as $n \to \infty$.

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

$$\lim_{n\to\infty} \frac{f(g)\mathrm{E}(x_n(t)|g)}{\Pr(\text{piv}|\sigma_n, n)^{\frac{2}{d-1}} e(\hat{y}, d)} = W(g, \kappa, \alpha). \tag{31}$$

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

Sketch of the proof. Details are in Appendix C. Recall the analysis of the best response. The precision of each type t acquiring information (i.e., $x(t) > 0$) is pinned down by the total intensity $k(t)$; combining (17) and (19),

$$\frac{x(t)}{\Pr(\text{piv}|\sigma_n, n)^{\frac{1}{d-1}}} = \Big[k(t)e(y(t))\Big]^{\frac{1}{d-1}}. \tag{32}$$

This suggests that the parametrisation of the types of an interest group g through the threshold of doubt y and the total intensity k will be useful:[30] in fact, what makes the analysis tractable is that we evaluate the integral $\mathrm{E}(x_n(t)|g)$ iteratively, first along the y-dimension, then along the k-dimension.[31] We write $t(y, k)$ for the unique type of an interest group with threshold of doubt y and total intensity k.

Fix k. Recall that only types close to the indifferent marginal types have enough incentives to acquire information in a large election.[32] We use Taylor approximations around the marginal type with intensity k to show: the precision of any type $t(y, k)$ acquiring information under the best response is asymptotically

[30]Recall that there is a one-to-one relation between types t and pairs of thresholds of doubt $y(t)$ and total intensities $k(t)$: $t_\alpha = k(t)(1 - y(t))$ and $t_\beta = k(t)y(t)$, given (4)) and (5).

[31]We can integrate iteratively by an application of the law of iterated expectation.

[32]Formally, this can be seen from the indifference conditions that pin down which types acquire information: e.g. it follows from (17) and (19) that the right hand side of (63) goes to 0 as $n \to \infty$. This implies that $\chi(y(t)) \to \frac{1}{2}$ for the threshold $y(t)$ of the boundary type. However, this is equivalent to $y(t) \to \Pr(\alpha|\text{piv}; \sigma_n, n)$ where $\hat{y}_n = \Pr(\alpha|\text{piv}; \sigma_n, n)$ is the threshold of doubt of the marginal type.

equivalent to the precision of the marginal type,[33]

$$x_n(t(y,k)) \approx x_n(t(\hat{y}_n,k)), \tag{33}$$

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

Then, we show that the likelihood that a type with intensity k acquires information is asymptotically proportional to the product of likelihood and precision of the marginal type,

$$\Pr(x_n(t)>0|g,k) \approx f(t(\hat{y}_n,k)|g,k)x_n(t(\hat{y}_n,k))e_2(\hat{y},d), \tag{34}$$

where f is the conditional density of the type and e is a constant that only depends on the limit marginal type $\hat{y} = \lim_{n\to\infty}\hat{y}_n$ and d.[34] To show this, we evaluate the interval of types with intensity k that acquire information. A Taylor approximation show that the mass of types in the interval is asymptotically proportional to the likelihood of the marginal type times the interval length. In the relevant step, we leverage that the precision of each boundary type of this interval enters *linearly* into the boundary conditions that define the length of the interval (see (23)) (and then that this precision is asymptotically equivalent to the marginal type's precision given(33)). This way, we establish (34). Combining (33) and (34), we then show that

$$\mathrm{E}(x_n(t(y,k))|g,k) \approx f(t(\hat{y}_n,k)|g,k)x_n(t(\hat{y}_n,k)^2e_2(\hat{y},d). \tag{35}$$

Using (32) for $t = t(\hat{y}_n,k)$, 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,

$$\frac{\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\text{piv}|\sigma_n,n)^{\frac{2}{d-1}}} \approx \left[f(t(\hat{y}_n,k)|g,k)k^{\frac{2}{d-1}}\right]e_3(\hat{y},d). \tag{36}$$

for $e_3(\hat{y},d) = e_2(\hat{y},d)e(\hat{y})^{\frac{2}{d-1}}$. The state-dependent intensity of the marginal types $t(\hat{y}_n,k)$ is linear in the total intensity, $t_\alpha = k(1-\hat{y}_n)$, compare to (4) and (5). So, (36) implies that, fixing k, the mean precision of a type in the interest group is proportional to the likelihood of the marginal type and the power $t_\alpha^{\frac{2}{d-1}}$ of the state-dependent intensity. Finally, integrating over k, we show that the analogous

[33]Two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ are asymptotically equivalent if $\lim_{n\to\infty}\frac{x_n}{y_n} = 1$.
[34]Recall that the lemma applies when $\hat{y}_n = \Pr(\alpha|\text{piv};\sigma_n,n)$ converges in $(0,1)$.

observation holds, that is, the mean precision of a type in the interest group is proportional to the κ-index (30), meaning that (31) holds.

4.2 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 a common interest benefit from her tipping the election into her preferred direction. 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 observations, showing that, if $d < 3$, no informative equilibrium sequence exists.

We take a candidate informative equilibrium sequence in which the election is close in both states, i.e., (26) holds. Given (25), what matters for the "informativeness" of the aggregate voting behaviour is the distance of the expected vote share in the two states in terms of standard deviations as $n \to \infty$,

$$\begin{aligned}\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) &= \lim_{n\to\infty} \frac{q(\alpha;\sigma_n) - q(\beta;\sigma_n)}{s(\alpha;\sigma_n)} \\ &= \lim_{n\to\infty} \frac{2\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t)}{s(\alpha;\sigma_n)} \qquad (37)\end{aligned}$$

where we used the definition (24) and that $\lim_{n\to\infty} \frac{s(\alpha;\sigma_n)}{s(\beta;\sigma_n)} = 1$, given (26). Hence, the relevant comparison is how fast the precision acquired by a random individual voter decreases relative to the standard deviation of the vote share.

We make two observations. The first observation is that, depending on $d < 3$ or $d > 3$, the precision acquired by a random voter of the interest group is of an order smaller or larger than the pivotal likelihood as a consequence of Lemma 3.

The second observation is that the approximation (25) also holds locally (illustrated in Figure 4),[35]

$$\lim_{n\to\infty} \Pr(\text{piv}|\omega;\sigma_n)(2n+1)s(\omega;\mathbf{q}(\sigma_n)) = \lim_{n\to\infty} \phi(\delta_n(\omega)), \qquad (38)$$

where ϕ the density of the standard normal distribution. Let $s_n = s(\omega;\mathbf{q}(\sigma_n))$ and $q_n = q(\omega_n;\sigma_n)$. Given (38), the pivotal likelihood is a finite multiple of

[35] 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.

$((2n+1)s_n)^{-1} = s_n(q_n(1-q_n)^{-1}$, so, a finite multiple of the standard deviation.[36] Combining both observations, (31) and (38), we see that the amount of information acquired by a random voter vanishes relative to the standard deviation if $d < 3$. Hence, given (37), the candidate sequence cannot be informative.

4.3 Existence: Information complementarities

Given that the precision of all voter types vanishes to zero uniformly as the electorate grows large (see e.g. (19)), one may conjecture that limit equilibrium outcomes in all states are given by whichtever policy is favored by a strict majority given the prior beliefs (generically, there is one such policy). What drives the existence of the informative equilibrium sequences of Theorem 1—besides information of low precision being sufficiently cheap ($d > 3$)—is that the voters' information acquisition exhibits complementarities, as sketched below.

Fix a vote share $q(\alpha) \neq \frac{1}{2}$. We can vary the informativeness of a voter strategy σ_n with $q(\alpha;\sigma_n) = q(\alpha)$ (see Section 3.2.1) by varying $q_n(\beta) = q(\beta;\sigma_n)$. Figure 5 shows the limit vote share for the reform A under the best response σ_n' as $n \to \infty$, as a function of $q(\beta) > \frac{1}{2}$; the limit vote share is $\lim_{n\to\infty} q(\omega;\sigma_n') = \Phi(\Pr(\alpha|\text{piv};\sigma_n))$ since the precision of all types vanishes as the electorate grows large. As the vote share in β is less close to the majority threshold, voters believe the state α to be more likely conditional on the election being tied. Therefore, the support for the reform increases since preferenes are "monotone", i.e., Φ is strictly increasing. Importantly, there are vote shares in β so that the election becomes close to being tied under the best reponse, $\Phi(\Pr(\alpha|\text{piv};\sigma_n)) \approx \frac{1}{2}$. Hence, certain levels of informative voting induce a close election and thereby high incentives to acquire information.

Formally, when the expected vote share in β is given by $\frac{1}{2}$, then $\frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} \to 0$ as $n \to \infty$ since the distance $\delta_n(\alpha)$ of the expected vote share $q(\alpha)$ to $\frac{1}{2}$ will become arbitrarily large in terms of standard deviations. Hence, given (38), $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\sigma_n,n)}{\Pr(\beta|\text{piv};\sigma_n,n)} = \frac{\Pr(\alpha)}{\Pr(\beta)}\frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} \to 0$. Conversely, if $q_n(\beta)$ is farther away from $\frac{1}{2}$ than $q(\alpha)$ by arbitrarily many standard deviations, then, $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\sigma_n,n)}{\Pr(\beta|\text{piv};\sigma_n,n)} = \infty$. When n is large, varying $q(\beta)$ lets the expected vote shares under the best response vary on almost the whole interval between $\Phi(0)$ and $\Phi(1)$, where $\Phi(0) < \frac{1}{2} < \Phi(1)$.

[36] 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 (24) and thereafter.

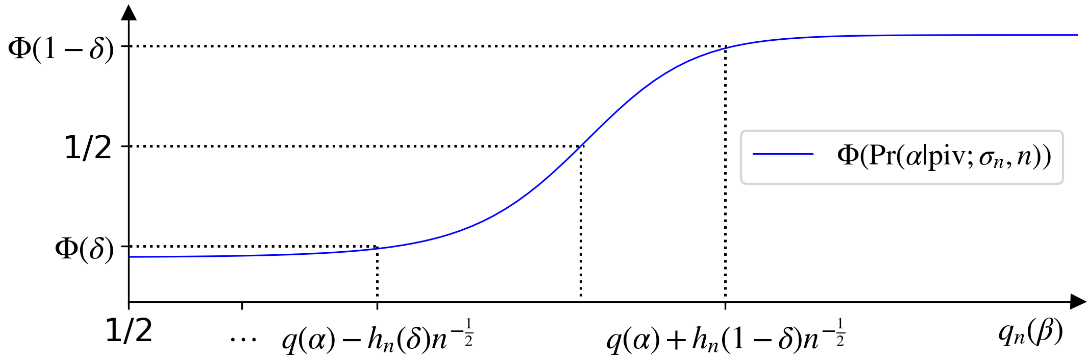

Figure 5: Fix $q(\alpha) < \frac{1}{2}$. The figure shows the limit vote share for policy A under the best response as $n \to \infty$, i.e. $\Phi(\Pr(\alpha|\text{piv}); \sigma_n)$, as a function of the expected vote share in β, for $q_n(\beta) > \frac{1}{2}$. The function $h_n(x)$ is so that, given $(q_n(\beta) - \frac{1}{2}) - (\frac{1}{2} - q(\alpha) = h_n(x)n^{-\frac{1}{2}}$.

4.4 Proof of Theorem 1

We represent informative equilibrium sequences in a compact way as sequences of fixed points of one-dimensional maps.[37] First, we show that equilibrium can be alternatively characterized in terms of the vector of the expected vote shares of the reform in state α and β, i.e.,

$$\mathbf{q}(\sigma) = (q(\alpha; \sigma), q(\beta; \sigma)). \tag{39}$$

Note that for any σ and any $\omega \in \{\alpha, \beta\}$, the vote share $q(\omega; \sigma)$ pins down the likelihood of the pivotal event conditional on ω, given (8). Given (11)-(14), (19), and (22), the vector of the pivotal likelihoods is a sufficient statistic for the best response, and therefore $\mathbf{q}(\sigma)$ as well. Given some vector of expected vote shares $\mathbf{q} = (q(\alpha), q(\beta)) \in (0, 1)$, let $\sigma^{\mathbf{q}}$ be the best response, 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.,

$$q^*(\alpha) = q\left(\alpha; \sigma^{\mathbf{q}^*}\right), \tag{40}$$
$$q^*(\beta) = q\left(\beta; \sigma^{\mathbf{q}^*}\right), \tag{41}$$

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

4.4.1 The fixed point maps

Lemma 4 builds on the insights from Section 4.3. The lemma shows that when the electorate size $2n+1$ is large, for any $q(\alpha)$ close to the majority threshold, but sufficiently many standard deviations away, there is a vote share $q_n(\beta) = q_n^+(q(\alpha)) > \frac{1}{2}$ and a vote share $q_n(\beta) = q_n^-(q(\alpha)) < \frac{1}{2}$, so that under the best response the vote share in β is again given by $q_n(\beta)$, i.e. $q_n^+(q(\alpha))$ and $q_n^-(q(\alpha))$ both solve (41).

Lemma 4 *Let $\Phi(\Pr(\alpha)) \neq \frac{1}{2}$. There are $\epsilon > 0$, $\bar{n} \in \mathbb{N}$, and $\Delta > 0$, so that for $n \geq \bar{n}$ and any $q(\alpha) \in \mathrm{B}_\epsilon(\frac{1}{2}) \setminus \mathrm{B}_{\Delta\sqrt{n}^{-1}}(\frac{1}{2})$, the equation (41) has a unique solution $q_n^+(q(\alpha))$ satisfying $\frac{1}{2} < q_n^-(q(\alpha)) < \frac{1}{2}+2\epsilon$ and a unique solution $q_n^-(q(\alpha))$ satisfying $\frac{1}{2}-2\epsilon < q_n^-(q(\alpha)) < \frac{1}{2}$. Further, $\frac{1}{2}+2\epsilon < \Phi(1)-\epsilon$ and $\Phi(0)-\epsilon < \frac{1}{2}-2\epsilon$.*

The proof of Lemma 4 is in Appendix G: we analyze $q(\beta;\sigma^{\mathbf{q}_n})$, the expected vote share in β under the best response to $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$. We show that, for $q_n(\beta) > \frac{1}{2}$ ($q_n(\beta) < \frac{1}{2}$) similarly close to $\frac{1}{2}$ as $q_n(\alpha)$, the derivative of $q(\beta;\sigma^{\mathbf{q}_n}) - q_n(\beta)$ becomes positive and arbitrarily large as $n \to \infty$. This will imply that it crosses 0 once, but only once, and establishes existence of a unique solution $q_n(\beta) > \frac{1}{2}$ ($q_n(\beta) < \frac{1}{2}$) to (41). To do so, we decompose

$$q(\beta;\sigma^{\mathbf{q}_n}) = \Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n,n)) + \epsilon_n(\mathbf{q}_n). \tag{42}$$

Here, $\Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n,n))$ is the vote share of A under the constrained best response when no information acquisition is possible and $\epsilon_n(\mathbf{q}_n)$ captures the difference to the actual best response, that is, the effect of information acquisition.

Intuitively, the effect of information acquisition is small when the electorate is large (the precision of all voters goes to zero uniformly). Conversely, the effect of the pivotal inference is large for large n. In particular, $\Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n,n))$ is increasing rapidly when the distance of $q_n(\beta)$ to $\frac{1}{2}$ increases, as Figure 5 illustrates. This basic intuition suggests that the derivative of $q(\beta;\sigma^{\mathbf{q}_n})$ is positive and large for large n. This intuition is incomplete since the incentives to acquire information, although small, increase sharply the closer $q_n(\beta)$ is to $\frac{1}{2}$, when n is large. The key step is to show that, nevertheless, the basic intuition upholds and that the effect on information acquisition, $\frac{\delta}{\delta q_n(\beta)}\epsilon_n(\mathbf{q}_n)$, vanishes relative to the effect of the pivotal inference, $\frac{\delta}{\delta q_n(\beta)}\Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n,n))$.

It follows from Lemma 4 that fixed points of the maps

$$\nu_1 : q(\alpha) \to q(\alpha;\sigma^{\mathbf{q}}) \quad \text{for} \quad \mathbf{q} = (q(\alpha), q_n^-(q(\alpha)) \tag{43}$$

and

$$\nu_2 : q(\alpha) \to q(\alpha;\sigma^{\mathbf{q}}) \quad \text{for} \quad \mathbf{q} = (q(\alpha), q_n^+(q(\alpha)) \tag{44}$$

satisfy (40)-(41). So, they correspond to equilibria of the voting game.

In the following, we show that these two maps have two fixed points when n is large enough and that these fixed points correspond to the informative equilibrium sequences of Theorem 1.[38] In Section 5.4.2 and Section 5.4.3, we consider the case when $W(\kappa, L, \alpha) < W(\kappa, C, \alpha)$. This is in order to highlight the cases that contrast the existing literature most: in this case, there can be equilibrium sequences in which the inverse of the full information outcome is elected state-by-state (Section 5.4.2) and in which the outcome preferred by a minority of the voters given the prior beliefs is elected in all states (Section 5.4.3). The proofs for the remaining cases are analogous and discussed in Appendix F.

4.4.2 Proof: Inverse of the full information outcome

Let $d > 3$. Take any type distribution so that $W(\kappa, L, \alpha) < W(\kappa, C, \alpha)$. We prove the first item of Theorem 1 and show that there is an equilibrium sequence in which the election outcome is the inverse of the full information outcome. We also show that full-information-equivalent outcomes are not possible. This highlights how the results contrast the existing literature where, in similar settings, the typical finding is that outcomes are full-information equivalent in all equilibria.[39]

To understand when full-information equivalent outcomes (A in α and B in β) are impossible, consider the effect of endogenous information acquisition. Any type choosing a precision $x > 0$ votes for her preferred policy in each state with probability $\frac{1}{2} + x$ (compare to Lemma 1), e.g. an aligned votes A with probability $\frac{1}{2}+x$ in α and with probability $\frac{1}{2}-x$ in β. The difference in the expected vote shares across the states is only due to informed voting, so that integrating over types, $q(\alpha;\sigma_n') - q(\beta;\sigma_n') = 2\int_{t\in L} x(t)dH(t) - \int_{t\in C} x(t)dH(t)$. Suppose that the κ-index

[38] It turns out that, in some cases, μ_1 has two fixed points, in some cases μ_2 has two fixed points, and, in some cases, each of the two maps has one fixed point. For details, see the proof in Section 4.4.2, Section 4.4.3, and in Appendix F.

[39] See e.g. Feddersen and Pesendorfer (1997), Martinelli (2006), or Oliveros (2013)).

of the aligned is smaller than that of the contrarians, i.e., $W(\kappa, L, \alpha) < W(\kappa, C, \alpha)$. Then, when n is sufficiently large, $\int_{t\in L} x(t)dH(t) < \int_{t\in C} x(t)dH(t)$, given Lemma 3.[40] The contrarians shift the election outcome towards their preferred policy in expectation in both states and

$$q(\alpha; \sigma_n') < q(\beta; \sigma_n'). \tag{45}$$

This means, that A is less likely to be elected in α than in β in *any* equilibrium; in particular, there is no informative equilibrium sequence in which A is elected in α and B in β, as $n \to \infty$.

To construct an equilibrium sequence with the inverse full-information outcomes, let the electorate be sufficiently large. Take a candidate equilibrium vote share $q_n(\alpha)$ from an interval close to the majority threshold as in Lemma 4, so that there are two solutions $\tilde{q}_n(q(\alpha))$ and $\hat{q}_n(q(\alpha))$ to (41). Let $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ with $q_n(\beta) \in \{q_n^-(q_n(\alpha)), q_n^+(q_n(\alpha))\}$.
The next part of the proof goes back to the analysis of the voters' incentives to free-ride. Varying the vote share in α, we can vary the closeness of the election in α. First, consider the case when the vote share of the reform is finitely many standard deviations away from the majority threshold in α. Then, the expectation of a close election in α creates relative large incentives to acquire information. Claim 1, 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 $\mathbf{q}_n$ differ by arbitrarily many standard deviations in the two states. The proof of Claim 1 is in Appendix D.

Claim 1 *If* $\lim_{n\to\infty} \frac{|q_n(\alpha)-\frac{1}{2}|}{s(\alpha;\mathbf{q}_n)} \in \mathbb{R}$*, then,*

$$\lim_{n\to\infty} \frac{|q(\alpha; \sigma^{\mathbf{q}_n}) - q_n(\beta)|}{s(\alpha; \mathbf{q}_n)} = \infty. \tag{46}$$

Second, we consider the case when the vote share of the reform A is bounded away from $\frac{1}{2}$ by some constant. We show that the same must be true for the candidate equilibrium vote share in β (see the following Claim 3). As a consequence, the incentives to acquire information under the best response are small. In fact, the pivotal likelihood becomes exponentially small, given (27), and, hence, also

[40]The condition of Lemma 3 is fulfilled for any candidate informative equilibrium sequence since they satisfy $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n, n) = \bar{y} \in (0, 1)$ where $\bar{y}$ is given by $\Phi(\bar{y}) = \frac{1}{2}$ (recall the discussion in Section 3.2.3).

the precision of any voter type under the best response, see (19). Since there is so little information acquisition, the vote shares of the best response do not differ by a standard deviation, as $n \to \infty$, as can be seen from (37).[41] We conclude:

Claim 2 *If $|q_n(\alpha) - \frac{1}{2}| \geq \epsilon$ for all n, then,*

$$\lim_{n\to\infty} \frac{|q(\alpha; \sigma^{\mathbf{q}_n}) - q_n(\beta)|}{s(\alpha; \mathbf{q}_n)} = 0. \tag{47}$$

For any candidate equilibrium vote shares $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ where $q_n(\beta)$ solves (41), the expected margin of victory must be "similarly" large in both states:

Claim 3

$$\lim_{n\to\infty} \frac{|q_n(\alpha) - \frac{1}{2}| - |q_n(\beta) - \frac{1}{2}|}{s(\alpha; \mathbf{q}_n)} \in \mathbb{R}. \tag{48}$$

Given (38), the pivotal likelihood with the local central limit theorem, the voters' inference from the election being tied depends on the distance of the vote shares to the majority threshold in terms of standard deviations. If the distance in one state is arbitrarily many standard deviations larger, i.e. if (48) does not hold, $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) \in \{-\infty, \infty\}$. Using (38), $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\mathbf{q}_n,n)}{\Pr(\beta|\text{piv};\mathbf{q}_n,n)} = \frac{\Pr(\alpha)}{\Pr(\beta)} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} \in \{0, \infty\}$, and the voters infer that the other state almost certainly holds conditional on the election being tied.[42] If the voters are almost certain of one state, the vote shares under the best response are arbitrarily close to $\Phi(0)$ or $\Phi(1)$. But, this cannot be true in the candidate equilibrium since the vote share $q_n(\beta)$ is bounded away from $\Phi(0)$ and $\Phi(1)$, see Lemma 4.

Take any sequence $(\Delta_n)_{n\in\mathbb{N}}$ for which $\lim_{n\to\infty} \Delta_n \in [-\Delta, \Delta]$ with a $\Delta > 0$ as in Lemma 4. For $q_n(\alpha) = \frac{1}{2} - \Delta_n n^{-\frac{1}{2}}$, Claim 3 implies $\lim_{n\to\infty} \frac{q_n(\beta) - q_n(\alpha)}{s(\alpha;\mathbf{q}_n)} \in \mathbb{R}$. Since $q(\alpha; \sigma^{\mathbf{q}_n}) < q(\beta; \sigma^{\mathbf{q}_n})$ for n sufficiently large, by (45), Claim 1 together with (41) implies $\lim_{n\to\infty} \frac{q_n(\beta) - q(\alpha;\sigma^{\mathbf{q}_n})}{s(\alpha;\sigma^{\mathbf{q}_n})} = \infty$. We conclude that

$$q_n(\alpha) > q(\alpha; \sigma^{\mathbf{q}_n}) \tag{49}$$

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

[42] Lemma 6 in Appendix E gives a self-contained analysis of the voter inference that does not rely on the local central limit theorem.

for n large enough. Take $q_n(\alpha) = \frac{1}{2} - \epsilon$ and let $q_n(\beta) = q_n^-(q_n(\alpha))$. Given Claim 3, $q_n(\beta) \to \frac{1}{2} + \epsilon$ and, given Claim 2, $q_n(\alpha; \sigma^{\mathbf{q}_n}) \to \frac{1}{2} + \epsilon$. Together,

$$q_n(\alpha) < q(\alpha; \sigma^{\mathbf{q}_n}) \tag{50}$$

for n large enough. Finally, using (49)- (50), an application of the intermediate value theorem shows that there is $q_n^*(\alpha) < \frac{1}{2}$ so that $q_n^* = (q_n^*(\alpha), q_n^-(q_n^*(\alpha)))$ solves (40) and (41). Further, it must be that $\lim_{n\to\infty} \frac{\frac{1}{2} - q_n^*(\alpha)}{s(\alpha; \mathbf{q}_n)} = \infty$ since otherwise (49) holds as we just argued. Hence, also $\lim_{n\to\infty} \frac{q_n(\beta) - \frac{1}{2}}{s(\beta; \mathbf{q}_n)} = \infty$, given Claim 3. 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 α and A in β as $n \to \infty$, given (25).

4.4.3 Proof: the minority-preferred outcome given the prior beliefs

Take any type distribution so that $W(\kappa, L, \alpha) < W(\kappa, C, \alpha)$ and $\Phi(\Pr(\alpha)) < \frac{1}{2}$. We prove the second item of Theorem 1 and show that there is a limit equilibrium where the election outcome is the minority-preferred policy given the prior beliefs as $n \to \infty$. This highlights that equilibrium outcomes may non-trivially depend on the prior beliefs of the voters, a finding that is novel to the literature on information aggregation in elections.

Take $q_n(\alpha) = \frac{1}{2} + \Delta_n n^{-\frac{1}{2}}$. We have $\lim_{n\to\infty} \frac{q_n(\beta) - q_n(\alpha)}{s(\alpha; \mathbf{q}_n)} \in \mathbb{R}$ by Claim 3. Since $W(\kappa, L, \alpha) < W(\kappa, C, \alpha)$, the vote shares of the best response are ordered as $q(\alpha; \sigma^{\mathbf{q}_n}) < q(\beta; \sigma^{\mathbf{q}_n})$; see (45). Since further $q_n(\beta) = q(\beta; \sigma^{\mathbf{q}_n})$, given (41), it follows from Claim 1 that

$$q(\alpha) > q(\alpha; \sigma^{\mathbf{q}_n}) \tag{51}$$

for n large enough.

Take $q_n(\alpha) = \frac{1}{2} + \epsilon$ and $q_n(\beta) = \hat{q}_n(q_n(\alpha))$. When $\Phi(\Pr(\alpha)) < \frac{1}{2}$, a majority of the voters prefers B given the prior beliefs. We claim that $q_n(\beta)$ is multiple standard deviations larger than $q_n(\alpha)$ when n is large, i.e.,

$$\lim_{n\to\infty} \frac{q_n(\beta) - q_n(\alpha)}{s(\alpha; \mathbf{q}_n)} > 0. \tag{52}$$

This is necessary for $\lim_{n\to\infty} \Pr(\alpha | \text{piv}; \mathbf{q}_n) \geq \Phi^{-1}(\frac{1}{2})$ since $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv};\mathbf{q}_n,n)}{\Pr(\beta|\text{piv};\mathbf{q}_n,n)} =$

$\frac{\Pr(\alpha)}{\Pr(\beta)}\frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))}$, given Lemma 6 and Claim 3. Claim 2 together with (52) implies

$$q_n(\alpha) < q(\alpha; \sigma^{\mathbf{q}_n}) \tag{53}$$

for n large enough. Finally, using (51) and (53), an application of the intermediate value theorem shows that there is $q_n^*(\alpha) > \frac{1}{2}$ so that $q_n^* = (q_n^*(\alpha), q_n^-(q_n^*(\alpha)))$ solves (40) and (41). Further, it must be that $\lim_{n\to\infty} \frac{q_n^*(\alpha)-\frac{1}{2}}{s(\alpha;\mathbf{q}_n)} = \infty$ since otherwise (49) holds as we just argued. Hence, also $\lim_{n\to\infty} \frac{q_n(\beta)-\frac{1}{2}}{s(\beta;\mathbf{q}_n)} = \infty$, given Claim 3. Hence, 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 (25).

5 General setting

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. We characterize all informative equilibrium sequences (Theorem 2).

We show that for any joint distribution of types, there is an auxiliary distribution which only admits heterogeneity in the state-dependent intensities and is *outcome-equivalent*, that is, it leads to the same set of equilibrium outcome distributions (Section 5.1). Based on this, we use the previous results for the baseline setting to prove Theorem 2 (Section 5.3). Finally, we discuss in detail the consequences of the heterogeneity of cost types and priors, and also of the correlation and dispersion of the primitives (Section 5.4).

5.1 Outcome-equivalent type distributions

We make two observations that allow us to construct outcome-equivalent distributions. Lemma 5 states the first observation.

Lemma 5 *Take any strategy σ'. The best response of a type $t = (v, r, t_\alpha, t_\beta)$ only depends on*

$$t'_\alpha = \frac{vt_\alpha}{r}, \quad \textit{and } t'_\beta = \frac{(1-v)t_\beta}{r}. \tag{54}$$

The proof of Lemma 5 is provided in Appendix H. The proof simply revisits the analysis of the best response as in Section 3.1, with minor modifications.

The second observation is that, as in the baseline setting, equilibria can be understood as a vector $\mathbf{q} = (q(\alpha), q(\beta))$ of expected vote shares for which the vote share vector of the best response is again $\mathbf{q}$, compare to (40) and (41). This is because the vote share vector is a sufficient statistic for the pivotal likelihood in each state, and therefore for the best response. The first observation implies that for each type t, the type

$$\zeta(t) = (\hat{v}, \hat{r}, \hat{t}_\alpha, \hat{t}_\beta)$$

given by $\hat{t}_\alpha = 2t'_\alpha$, $\hat{t}_\beta = 2t'_\beta$, $\hat{v} = \frac{1}{2}$, and $\hat{r} = 1$ best responds in the same way, given any $\mathbf{q}$ and n. Thus, for any distribution H, the push-forward distribution $\zeta_*(H)$ has the same equilibrium vote shares. In other words, $\zeta_*(H)$ is outcome-equivalent. Importantly, under $\zeta_*(H)$, all types share the same prior $\hat{v} = \frac{1}{2}$ and $\hat{r} = 1$.

The next section generalizes some definitions from the baseline setting by replacing the intensities t_ω with their counterpart t'_ω.

5.2 Definitions

We generalize the definition (2) of Φ. Let

$$\Psi(p) = \Pr_H(\{t : pt'_\alpha + (1-p)t'_\beta \geq 0\}). \tag{55}$$

for $p \in (0, 1)$. Figure 6 in Appendix H illustrates $\Psi(p)$. Comparing to the definition (2), the function Ψ is identical to Φ for the distributions H considered in the baseline setting for which all types in the support have $q = \frac{1}{2}$ and $r = 1$, so that $t_\omega = t'_\omega$ for all types in the support and all states ω. Similar to before, for the main analysis, we consider *monotone* type distributions H, that is, Ψ is strictly increasing and crosses $\frac{1}{2}$.[43]

We use the same arguments as for the baseline setting in Section 3.2.2 and Section 3.2.3 to observe that for any informative equilibrium sequence, (a) the outcome is close to being tied in *all* states ω, that is, $\lim_{n\to\infty} q(\omega; \sigma_n^*) = \frac{1}{2}$; (b) the closeness of elections pins down the marginal types as $n \to \infty$: the limit marginal types are those for which $\bar{p}t'_\alpha + (1-\bar{p})t'_\beta = 0$ where $\bar{p}$ is the unique belief satisfying

[43] The non-monotone case is discussed in Section 7.2.

$\Psi(\bar{p}) = \frac{1}{2}$. For details, see Appendix 3.2.3. Denote by $T_{\bar{p}}$ the set of the limit marginal types.

The *generalized κ-index* of an interest group g is

$$I(\kappa, g, \omega) = \underbrace{f(g)f(T_{\bar{p}}|g)}_{\text{likelihood of the limit marginal types}} \mathrm{E}(||\, t'_\omega \,||^\kappa \mid g, T_{\bar{p}}, \omega), \tag{56}$$

where t'_α and t'_β are the projections as in (54).[44] This definition is a generalization of the κ-index defined in Section 4: for the distributions H considered in the baseline setting for which all types have $t_\omega = t'_\omega$ for all ω, the generalized index is proportional to the index $W(\kappa, g, \omega)$ defined in (30).

5.3 Main results

As in the baseline setting, we consider type distributions H for which: (a) the distribution of (t_α, t_β) has a continuous density on its support and a connected and compact support $K_\alpha \times K_\beta \subseteq \mathbb{R}2$ with $(0,0)$ in its interior. Further, (b) the support of the cost type r is compact, connected and does not contain 0.[45] The condition (b) guarantees that the outcome-equivalent type distribution $\psi_*(H)$ also satisfies (a), which can be seen from (54).

Theorem 2 *Let $d = \lim_{x\to 0} \frac{c'(x)x}{c(x)} > 3$ and $\kappa = \frac{2}{d-1}$. Take a monotone type distribution H.*

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

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

Proof. Take the outcome-equivalent distribution $\psi_*(H)$, for which all types in the support have the prior $v = \frac{1}{2}$ and cost type $r = 1$, so that Theorem 1 applies.

[44]Recall that we use the following notation: $f(g)$ for the unconditional likelihood of the set of types $\{t : t \in g\}$, $f(-|g)$ for the conditional likelihood of the types, and $\mathrm{E}(-|g)$ for the conditional expectation, et cetera.

[45]The condition that 0 is in the interior of the support of t_ω for $\omega \in \{\alpha, \beta\}$, ensures that types may have conflicting interests when the state is known.

Since, for $\psi_*(H)$, the indices W and I are proportional and $\Phi = \Psi$, Theorem 2 simply restates Theorem 1 with the more general notation, and characterizes the informative equilibrium sequences given $\psi_*(H)$. Since H is outcome-equivalent, it has the same types of informative equilibrium sequences. Further, one checks that the index $I(\kappa, g, \omega)$ and the function Ψ are the same for two outcome-equivalent distributions. Hence, the same characterization of the informative equilibrium sequences in terms of I and Ψ applies, given H. ■

We discuss how having a different prior and cost type affects the "weight" of a marginal type: In the baseline setting, the "weight" of a marginal type is proportional to a power of the ex-post utility t_ω. This way, the index has a compelling interpretation as a welfare measure, see the discussion in Section 4. When cost types and priors are heterogeneous, this heterogeneity is reflected in the weights, creating a wedge between the index and ex-post utility: When types have higher cost r, they acquire less information. As a consequence, their weight in the index is smaller (it is divided by r, see (54) and (56)). Similarly, types with a prior belief that attaches more likelihood to α (higher q) have a relatively higher weight in the index $I(\kappa, g, \omega)$ when the state is $\omega = \alpha$, and a relatively lower weight when the state is $\omega = \beta$.

Changes in the distribution of the priors or cost types have additional non-trivial effects on equilibrium outcomes since the set of marginal types varies with the distribution.

The next section discusses the effect of varying the correlation and dispersion of the primitives for a certain class of symmetric distributions.

5.4 Discussion: Correlation and Dispersion

We discuss the role of correlation and dispersion of primitives, with a focus on the prior beliefs. First, recall the example from Section 1 with a common cost type $\gamma = 1$. In the example, all types of an interest group receive a utility of k_g if their ex-post preferred policy is elected and 0 otherwise. This implies that $t_\alpha = -t_\beta$, so that the common threshold of doubt is $\frac{-t_\beta}{t_\alpha - t_\beta} = \frac{1}{2}$. In the example, the marginal types are those with a uniform prior, $v = \frac{1}{2}$. They decide the election. The non-marginal types mis-coordinate so that their votes cancel out each other. The marginal types choose the precision x_g, proportional to the power $k_g^{\frac{1}{d-1}}$ of their total intensity (see (1)) and vote for their ex-post preferred policy with probability $\frac{1}{2} + x_g$. This way, the policy preferred by the aligned receives the larger vote share if and only if the intensity-weighted share of the aligned marginal types is larger, that is, if $k_L^{\frac{1}{d-1}} \Pr(L|q = \frac{1}{2}) > k_G^{\frac{1}{d-1}} \Pr(C|v = \frac{1}{2})$.

The example from Section 1 yields two observations. First, the dispersion of priors matters. The fewer types of an interest group have a uniform prior, ceteris paribus, the lower the vote share of the interest group's preferred policy in each state. Second, the correlation of the prior and the total intensity matters. The larger the intensities of the types with uniform prior, ceteris paribus, the higher the vote share of the interest group's preferred policy in each state. In Appendix J, we show that these two observations generalize to a class of symmetric type distributions, for which, compared to the example, we allow each interest group to have any symmetric distribution of the priors. Garbling the priors of an interest group, ceteris paribus, leads to a lower index $I(\kappa, g, \omega)$ of that group in all states. Further, if having intermediate priors (close to 1/2) is correlated more strongly with high total intensities, the index of the group is higher. For this class of symmetric type distributions, there is only the limit equilibrium in which the interest group with the higher index "wins" the election (see Theorem 2), so that these comparative statics yield clear predictions for policy outcomes.

6 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 $\Psi(\frac{1}{2}) > \frac{1}{2}$ and B if $\Psi(\frac{1}{2}) < \frac{1}{2}$.

Theorem 3 *Let* $\Psi(\frac{1}{2}) \neq \frac{1}{2}$.

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* $\lim_{n\to\infty} q(\omega; \sigma_n^*) = \Psi(\frac{1}{2})$ *for all states* $\omega \in \{\alpha, \beta\}$. *Hence,* $\lim_{n\to\infty} \Pr(A|\sigma_n^*, n) = 1$ *if* $\Psi(\frac{1}{2}) > \frac{1}{2}$ *and* $\lim_{n\to\infty} \Pr(B|\sigma_n^*, n) = 1$ *if* $\Psi(\frac{1}{2}) < \frac{1}{2}$.

Sketch of the proof. Suppose that an equilibrium sequence is not informative, which means that $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0$, 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,

$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n, n) = \Pr(\alpha)$. This is because $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv}, \mathbf{q}_n), n)}{\Pr(\alpha|\text{piv}, \mathbf{q}_n), n)} = \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))}$, given (38),[46] and because $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} = 1$ if $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0$ since the density ϕ of the standard normal is continuous. One can show that the precision of the types converges to zero uniformly, as in the baseline setting see e.g. (19). Since citizens do not learn anything from their signals or from conditional on being pivotal as $n \to \infty$, the behaviour of the voters converges to "voting according to the prior".

The genericity condition $\Psi(\frac{1}{2}) \neq \frac{1}{2}$ excludes the case in which a random voter is equally likely to prefer the reform and the status quo given the prior belief. This way, when the citizens vote according to the prior beliefs, the pivotal likelihood is exponentially small (recall (27)) and, hence, the incentives to acquire information. We show that this guarantees that the best response is again "non-informative", meaning that the vote shares in α and β do not differ by a standard deviation as $n \to \infty$.

Putting the arguments together, when the citizens follow the non-informative strategy to vote according to their prior beliefs, the sequence of best responses is again non-informative and close to voting according to the prior beliefs. In Appendix K, we provide the formal fixed point argument, proving the existence of non-informative equilibrium sequences.

7 Extensions and discussion

7.1 Ordering the equilibrium sequences along the informativeness

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

In any non-informative equilibrium sequence, by definition,

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

For the informative equilibrium sequence in which the policy preferred by the

[46] Alternatively, this follows from Lemma 6 in Appendix E, to give a self-contained argument that does not rely on the local central limit theorem.

interest group with the higher index is elected as $n \to \infty$, the distribution of the limit outcomes is degenerate and depends on the state. Thus, (25) implies that

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

Take an informative equilibrium sequence in which the limit outcome is the same in both states. In Appendix L, we show that the informativeness lies in between that of the other two types of equilibrium sequences, that is,

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

7.2 Non-monotone preferences

So far, we provided the analysis for the setting in which preferences are "monotone". When Φ is non-monotone, there may be multiple $\bar{p} \in (0,1)$ for which $\Psi(\bar{p}) = \frac{1}{2}$. This motivates the definition of a *local* κ-index, defined in the same way as $I(\kappa, g, \omega)$ (see (56)), but which depends on the selection of $\bar{p}$ satisfying $\Psi(\bar{p}) = \frac{1}{2}$.

One can show that, for any such $\bar{p}$ with $\Psi(\bar{p}) = \frac{1}{2}$ and $\Phi'(\bar{p}) \neq 0$, the statements analogous to those of Theorem 2 hold, where we simply substitute the κ-index $I(\kappa, g, \omega)$ with the local index $I(\kappa, g, \omega, \bar{p})$:[47] Precisely, if $d > 3$, $\Psi(\frac{1}{2}) \neq \frac{1}{2}$, $I(\kappa, L, \omega, \bar{p}) \neq I(\kappa, C, \omega, \bar{p})$, there are two informative equilibrium sequences for which $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} = \frac{\bar{p}}{1-\bar{p}}$.There is one informative equilibrium sequence for which the outcome preferred by the group with the larger power $I(g, \kappa, \omega, \bar{p})$ is elected state-by-state as $n \to \infty$. There is another informative equilibrium sequence for which the outcome that is preferred by the majority of the citizens given the prior beliefs is elected with probability converging to 1 (to 0) if the *local* index $I(g, \kappa, \omega, \bar{p})$ of the aligned ($g = L$) is larger (smaller) than that of the contrarians ($g = C$).

One important implication is that it may happen that the *order* of the local index of the interest groups varies with $\bar{p}$, so that there are informative equilibrium sequences for which on group wins the election with probability converging to 1 as $n \to \infty$, but also other informative equilibrium sequences in which the other interest group wins.[48]

[47] We omit the proof since it is completely analogous to the proof of Theorem 2.

[48] These results are reminiscient of known results about equilibrium multiplicity for the model with exogenous information: Take the baseline setting from Section 3. If citizens were to receive

8 Literature

8.1 Dispersed information in elections

The Condorcet Jury Theorem (CJT) states that elections effectively aggregate information that is dispersed among citizens, so that outcomes of a large election are full-information equivalent. Modern versions of the CJT have been established for a rich set of environments with strategic voters (Austen-Smith and Banks, 1996; Feddersen and Pesendorfer, 1997), costly information (Martinelli, 2006, 2007), and when voters have conflicting interests Bhattacharya (2013).[49]

The literature has also identified circumstances in which information may fail to aggregate. Bhattacharya (2013) observes that preference monotonicity is necessary for information aggregation.[50] Feddersen and Pesendorfer (1997) show that information aggregation fails when the aggregate distribution of preferences is uncertain conditional on the realized state.[51] Mandler (2012) observes a similar failure if the aggregate distribution of signals is uncertain to the voters. In these settings, the effective state is multi-dimensional. Intuitively, this creates an invertibility problem: the order statistic of the vote shares can only separate the payoff-relevant states inefficiently. In a recent paper, Barelli *et al.* (2017) provide a more general analysis of such an invertibility problem, showing that the relative richness of the signal and the state space matter for the possibility of information aggregation.

Further, it has been shown that information aggregation fails when there is policy uncertainty (Gul and Pesendorfer, 2009), and when voters have a signalling motive to affect the policy choice after the election (Razin, 2003). Several papers have considered "extended" election games: Ekmekci and Lauermann (2020) consider a conflicted organizer who can select the number of participating voters, Bond and Eraslan (2010) consider a third-party makes a strategic proposal that the voting body can collectively reject or accept, Heese and Lauermann (2017) study persuasion of voters by a third-party.[52]

---

a costless, binary, conditionally i.i.d. signal about the state with precision $0 < x < \frac{1}{2}$ and if Φ 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).

[49]See also Persico (2004) for another setting with costly information: there, the focus is on *how* to acquire information efficiently as a group, not so much, *if* information aggregates when information is costly.

[50]See also Acharya (2016), Bhattacharya (2018).

[51]See Section 6 of their paper.

[52]See also Alonso and Câmara (2016) and Bardhi and Guo (2018) for earlier work on voter

In our setting, there is no policy uncertainty, there are no signalling motives, and the game is not extended by an interested party. Conditional on the state, there is no residual aggregate uncertainty about the distribution of signals or preferences, and preferences are monotone.[53] Relative to the findings in the "standard" settings by Feddersen and Pesendorfer (1997) and Martinelli (2006), there are two main contributions. First, in our setting information aggregation may fail in *all* equilibria because outcomes are driven by *competition and coordination*: opposing interest groups compete and each interest group faces the problem to coordinate on supporting the ex-post preferred policy. Second, predictions depend on the selected equilibrium. We show that the strategic interdependence of the costly information acquisition drive the existence of three limit equilibria. When preferences are non-monotone, even more equilibria with yet diverging predictions may exist, by a logic similar to that in Bhattacharya (2013), as discussed in Section 7.2.

8.2 Political competition and asymmetric information in distributive politics

Studying the competition between opposing interest groups for political influence is a central issue in several streams of the literature in political science and economics. The literature has studied a variety of models in which opposing groups compete in costly actions to advance their interests. We discuss some related work in Section 8.2.1 and Section 8.2.2.

Relative to the existing work, we provide a novel model of competition: actions (voting choices) are not costly and it is uncertain which action advances the own interest. The model highlights that *internal information asymmetries* shape the internal coordination of competing groups and this way play a central role for political outcomes. We analyze which information asymmetries arise endogenously when there are incentive constraints due to information being costly. In particular, we provide a detailed, yet simple description of how primitives map into equilibrium information and outcomes based on an index formula (see Theorem 1 and Theorem 2).

Prior work has studied other effects of asymmetric information in distributive politics: this work finds, for example, that asymmetric information opens up the

persuasion.

[53] We discuss non-monotone preferences in Section 7.2.

room for "signal jamming" (Esteban and Ray, 2006), and may create an adverse selection problem when pay-off types are negatively correlated conditional on the state (Ali *et al.* (2018)). In the literature on electoral competition between political platforms, Matějka and Tabellini (2017) find that when voters are rationally inattentive, this affects the policies announced by the platforms: platforms cater more to the more attentive citizens. A main difference—besides a focus on a different set of questions— is that in this prior work the information acquisition is not strategically interdependent, which is a central driving force for the results in our analysis.

8.2.1 Special interest politics

A large literature on "special interest politics" (as pioneered by Olson (1965) and Tullock (1983)) has brought forward the idea that the coordination of interest groups depends on *internal incentive constraints*. Most prominently, Olson (1965) argues that members of an interest group may shirk from contributing to a joint effort to take influence on policy-making, and that, as a consequence, a large group may be dominated by a smaller, more effective group. This is because the smaller group suffers relative less from free-riding. In our setting, the relative size of an interest group does not affect the relative free-riding incentives: the pivotal likelihood is the same for each agent, independent of the group that he belongs to. This way, the Olsonian logic of a relative group size effect is not at play.[54]

8.2.2 Vote-buying and costly voting

Theories of costly voting and vote-buying constitute another stream of literature in which opposed interest groups compete in a political contest.

Krishna and Morgan (2011, 2015) show that simple majority voting with voting costs leads to outcomes that asymptotically maximize utilitarian aggregate welfare when the electorate is large ($n \to \infty$).[55] In Casella *et al.* (2012), a competitive market for votes results in dictatorship by the most concerned individual (or group). Eguia and Xefteris (2018) study vote-buying mechanisms for collective decisions over binary policies. These vote-buying mechanisms can be viewed as re-

[54] This has to do with the fact that the number of voters is odd and abstention is not possible. In settings in which voting is costly, this relative group size effect arises: see e.g. the "underdog effect" in (Krishna and Morgan, 2015).

[55] These results depend on further assumptions, e.g., that voting cost are uncorrelated with the preference over the policies and that the lower bound of the voting cost is zero.

duced form descriptions of certain classes of political contests: under a vote-buying mechanism, each agent can take any positive a positive action $a > 0$ to support policy A or a negative action to support policy B, while incurring a cost $c(|a|)$; outcomes depend on the net contribution of all agents, that is, $\sum_{i=1}^{n} a_i$. They show that in large societies (n large), given further relatively weak assumptions on the environment, vote-buying mechanisms implement a one-parameter class of welfare functions: each agent's willingness to pay to change the policy outcome is taken to the power κ. This defines the individual's decision weight. The policy that is supported by the group of agents with the larger sum of weights is chosen as outcome.

In the baseline setting in this paper, outcomes are given by a similar welfare rule, however, only in one limit equilibrium and the rule only counts the weights of the marginal types (compare to the discussion following Theorem 1). These similarities and differences suggest that there are deeper connections between voting models with costly actions and models with costly information. For example, one may conjecture, that introducing information asymmetries (and costly information) into vote buying models may (a) affect which group is successful and (b) create a similar equilibrium multiplicity as in the setting in this paper.

9 Conclusion

We have studied a large class of voting settings in which payoff consequences are uncertain and distributive, and in which becoming better informed about policy consequences is costly. The existing result from related "standard" settings, that large elections aggregate the dispersed information of citizens in a way so that outcomes are full-information equivalent, does not uphold (see e.g. Feddersen and Pesendorfer, 1997; Bhattacharya, 2013; Martinelli, 2006).

Our first contribution is to show that in a large class of settings, outcomes may not be full-information equivalent in *any* equilibrium. Outcomes are driven by how well competing interest groups coordinate internally. Internal incentive constraints, due to information being costly, drive the asymmetries in information within a group and thereby how well each group coordinates on voting for the ex-post preferred policy. We have provided a detailed, yet simple description of how primitives map into equilibrium information and outcomes based on an index formula. More broadly, our analysis highlights the importance of further work on the relation between internal information-related incentive constraints on the effectiveness of groups.

The second main contribution is to observe that the strategic interdependence of costly information acquisition in elections creates a global coordination problem: there are three types of equilibria that are ordered according to their "informativeness".

The third main contribution is to provide a novel model of political competition between voters. Previous work has considered a variety of models in which opposing groups compete in costly actions to advance their interests (see, e.g., the models of vote-buying, costly voting or the lobbying models in the literature on special interest politics.) In our model, competition is of a different kind: consequences of actions are uncertain and interest groups compete in how well they coordinate on voting for the ex-post preferred policy.

Our model and results speak to a variety of public and academic debates that center around the informedness of citizens:

- *Populism and voter ignorance.* Voter ignorance has been brought forward as an explanation for the rise in populism in numerous Western democracies and electoral choices like the Brexit.[56] Some even argue that voter ignorance is *the* major problem for democracy, see e.g. Brennan (2016). In our model, the non-informative equilibrium—which exists for all (non-zero) levels of information cost—is consistent with this pessimistic view on voter informedness.[57] However, we also show that the electorate may coordinate on other, more informative equilibria (Theorem 2). In our theory, the closeness of the election correlates with the informativeness of the equilibrium. The frequent occurence of close elections in practice is, in this respect, an indicator of intense, but functioning electoral processes. More work is needed to study which circumstances lead to coordination on the informative equilibria.

- *Voter information campaigns and websites.* Civil society groups advocate using voter information campaigns to improve democratic representation and the fairness of elections.[58] Similarly, websites that try to help voters compare their own political positions with those of the candidates of an

[56] See, e.g., `https://www.washingtonpost.com/news/volokh-conspiracy/wp/2016/06/26/brexit-regrexit-and-the-impact-of-political-ignorance/` and `https://www.washingtonpost.com/news/volokh-conspiracy/wp/2016/06/14/brexit-and-political-ignorance/`

[57] In Martinelli (2006) and earlier work, this equilibrium only exists for relatively low cost levels. This result relies on certain symmetry assumptions; generically, it does not hold.

[58] See, e.g., `https://www.vote411.org/`.

upcoming election have gained popularity in the recent decade.[59] In the context of our model, such websites and campaigns may be understood as lowering the information cost. Our results suggest that their effects are ambiguous: when information cost are lower overall, this may hinder the representation of minority interests; however, if particularly the information cost of demographic groups with otherwise comparably high cost are lowered, this may "level the playing field" and lead to a better representation of these groups' interests by election outcomes.

- *Social Media.* There is considerable concern about the role that social media, such as Facebook and Twitter, play in promoting misinformation. Recent empirical work attempts to clarify the effects of social media on the political knowledge and misperceptions of citizens, for example in the context of the Covid-19 pandemic; see, e.g., Bridgman *et al.* (2020)) and Allcott *et al.* (2020). In the context of our model, misinformation may be viewed as altering the cost of obtaining correct information since its presence requires filtering and more cross-checking across sources., This way, social media participation may reduce voter informedness; further, if participation varies across demographic groups, the interests of particularly active groups may be less well reflected by election outcomes.

[59] See, e.g. `https://www.isidewith.com/elections/2020-presidential-quiz`, `https://2020election.procon.org/2020-election-quiz.php`, or `https://www.bpb.de/politik/wahlen/wahl-o-mat/`.

Appendix

Recall our notation: we denote by $\mathrm{E}(-|g)$ and $f(-|g)$ 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 $f(g)$ for the unconditional likelihood and $\mathrm{E}(-|y)$ and $f(-|y)$ when conditioning on the set of types with threshold of doubt $y(t) = y$, et cetera. In the following, we also use notation like $\Pr(\alpha|\text{piv}; \mathbf{q}_n, n)$ and similar for pairs of expected vote shares $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$. This notation is a shortcut for $\Pr(\alpha|\text{piv}; \sigma_n, n)$ where σ_n is a strategy with expected vote shares $\mathbf{q}_n$.

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

$$\Pr\left(\alpha|b, \text{piv}; \sigma', n\right) < \Pr\left(\alpha|a, \text{piv}; \sigma', n\right). \tag{57}$$

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

$$\Pr(\alpha|b, \text{piv}, \sigma', n) < y(t) < \Pr(\alpha|b, \text{piv}, \sigma', n). \tag{58}$$

Otherwise, given (11)-(14), there is a policy $z \in \{A, B\}$ that the voter weakly prefers, independent of her private signal $s \in \{a, b\}$. But then, she would be strictly better off by not paying for the information $x(t) > 0$ and simply voting the same after both signals. Finally, (11)-(14), and (58) together imply (20) and (21)

Proof of (57)**.** 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

$$\frac{\Pr\left(\alpha|s, \text{piv}; \sigma', n\right)}{\Pr\left(\beta|s, \text{piv}; \sigma', n\right)} = \frac{\Pr\left(\alpha\right)}{\Pr\left(\beta\right)} \frac{\Pr\left(\text{piv}|\alpha; \sigma', n\right)}{\Pr\left(\text{piv}|\beta; \sigma', n\right)} \frac{\Pr(s|\alpha; \sigma)}{\Pr(s|\beta; \sigma)}, \tag{59}$$

if $\Pr\left(\text{piv}|\beta; \sigma', n\right) > 0$, 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 (57) follows from $\Pr(a|\alpha; \sigma) = \frac{1}{2} + x$ and $\Pr(a|\beta; \sigma) = \frac{1}{2} - x$, and the analogous formula for $s = b$.

B Proof of Lemma 2

We show that the indifference condition of the type $y_L^-(k)$ can be rewritten to obtain (23). The argument in the main text then shows that there is a unique solution $y_L^-(k) < \Pr(\alpha|\text{piv}; \sigma', n)$ when n is large enough. The proof for the other boundary types $y_L^+(k), y_C^-(k), y_L^+(k)$ is analogous. In the following, for the ease of presentation, we drop the dependence on n and σ' in the notation.

The boundary type $y_L^-(k)$ is indifferent between voting A without further information and choosing the precision $x = x^*(t; \sigma', n)$. When choosing $x = x^*(t; \sigma', n)$ the expected utility from the policy elected in the pivotal event is given by (15) in α and by (16) in β. Hence, the indifference condition is

$$\begin{aligned}
& \Pr(\text{piv})\Big[\Pr(\alpha|\text{piv})(\frac{1}{2} + x)t_\alpha + \Pr(\beta)(\frac{1}{2} - x)t_\beta\Big] - c(x) \\
= \;& \Pr(\text{piv})\Big[\,\Pr(\alpha|\text{piv})t_\alpha + \Pr(\beta|\text{piv})t_\beta\Big]. \qquad (60)
\end{aligned}$$

Rearranging,

$$\begin{aligned}
& \Pr(\text{piv})\Big[(\frac{1}{2} + x)\Big[\Pr(\alpha|\text{piv})t_\alpha - \Pr(\beta|\text{piv})t_\beta\Big] + \Pr(\beta|\text{piv})t_\beta\Big] - c(x) \\
= \;& \Pr(\text{piv})\Big[\Pr(\alpha|\text{piv})t_\alpha - \Pr(\beta|\text{piv})t_\beta + 2\Pr(\beta|\text{piv})t_\beta\Big] \qquad (61)
\end{aligned}$$

Plugging (17) and (18) into (61),

$$\begin{aligned}
& (\frac{1}{2} + x)c'(x) - c(x) + \Pr(\text{piv})\Pr(\beta|\text{piv})t_\beta \\
= \;& c'(x) + 2\Pr(\text{piv})\Pr(\beta|\text{piv};)t_\beta. \qquad (62)
\end{aligned}$$

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

$$(\frac{1}{2} + x) - \frac{c(x)}{c'(x)} = 1 + \frac{\Pr(\beta|\text{piv})t_\beta}{\Pr(\alpha|\text{piv})t_\alpha + \Pr(\beta|\text{piv})(-t_\beta)}. \qquad (63)$$

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

$$(\frac{1}{2} + x) - \frac{c(x)}{c'(x)} = 1 + \frac{-\Pr(\beta|\text{piv})y(t)}{\Pr(\alpha|\text{piv})(1 - y(t)) + \Pr(\beta|\text{piv})y(t)}. \qquad (64)$$

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 (64) and rearranging gives (23), i.e.,

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

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

C Proof of Lemma 3

Take the interest group of the aligned types. The proof for the group of 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 that has $y(t) = y$ and $k(t) = k$, $H(y, k)$ for the joint distribution of k and y, and $H(k)$ and $H(y)$ for the marginal distributions. We evaluate the mean precision

$$\text{E}(x_n(t)|g) = \text{E}(\text{E}(x_n(t)|g, k)|g) \tag{65}$$

iteratively. We start by analysing $\text{E}(x_n(t)|g, k)$ 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 $\bar{y}_n = \Pr(\alpha|\text{piv}; \sigma, n)$ 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 (63) that pins down the boundary types $y_g^-(k)$; the proof for the other boundary type is analogous. It follows from (17) and (19) that the right hand side of (63) goes to 0 as $n \to \infty$. This implies that $\chi(y(t)) \to \frac{1}{2}$ 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 $y(t) \to \bar{y}_n$. ■

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 the first-order condition (19). Using a Taylor approximation of $x^*(t(y',k);\sigma_n,n)$,

$$x_n(t(y',k)) - x_n(t(\bar{y}_n,k)) = (\bar{y}_n - y')\frac{d}{dy}_{|y=\hat{y}_n(y')} x^*(t(y,k);\sigma_n,n) \tag{66}$$

for some $\hat{y}_n(y') \in [y',\bar{y}_n]$. Given (17) and (19),

$$\frac{d}{dy}_{|y=\hat{y}_n(y')} x^*(t(y,k);\sigma_n,n) = x_n(t(\bar{y}_n,k))M_n(y') \tag{67}$$

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

$$M_n(y') = (\frac{e(\hat{y}_n(y'))}{e(\bar{y}_n)})^{\frac{1}{d-1}}\frac{(1-2\bar{y}_n)}{e(\hat{y}_n(y'))}. \tag{68}$$

It follows from Step 1 that $\hat{y}_n(y') \to \bar{y}_n$ as $n \to \infty$, so that $\lim_{n\to\infty}(\frac{e(\hat{y}_n(y'))}{e(\bar{y}_n)})^{\frac{1}{d-1}} = 1$ for all y'. Thus,

$$\lim_{n\to\infty}\max_{y':x(t(y',k))>0}|M_n(y')| = \lim_{n\to\infty}|\frac{(1-2\bar{y}_n)}{e(\bar{y}_n)}| = \lim_{n\to\infty}|\frac{(1-2\bar{y}_n)}{2\bar{y}_n(1-\bar{y}_n)}|. \tag{69}$$

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 (66) and (67),

$$\begin{aligned} x(t(y',k)) &= x(t(\bar{y}_n,k)) + x(t(\bar{y}_n,k))M_n(y')(\bar{y}_n - y), \\ \Leftrightarrow \frac{x(t(y',k))}{x(t(\bar{y}_n,k))} &= (1+M_n(y'))(\bar{y}_n - y'). \end{aligned} \tag{70}$$

Finally, (70), the observation that $\lim_{n\to\infty}M_n(y') \in \mathbb{R}$, and Step 1 together imply 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 $f(t|g,k)$ the density of the types conditional on $t \in g$ and $k(t) = k$.

Step 3

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

for $e_2(y,d) = \frac{8(d-1)}{d}(1-y)^2$.

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

$$\Pr(\{t : x_n(t) > 0\}|g,k) \quad \approx \quad f(t(\bar{y}_n,k)|g.k)(y_g^+(k) - y_g^-(k)), \tag{71}$$

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 (19). Taylor approximations of the function χ from the indifference conditions (see e.g. (23)) 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

$$\chi'(\bar{y}_n)\Big[y_g^-(k) - \bar{y}_n\Big] \approx \frac{(d-1)}{d}x^*(t(y_g^-(k),k);\sigma_n,n), \tag{72}$$

$$\chi'(\bar{y}_n)\Big[\bar{y}_n - y_g^+(k)\Big] \approx \frac{(d-1)}{d}x^*(t(y_g^+(k),k);\sigma_n,n). \tag{73}$$

Algebra shows that $\chi'(\bar{y}_n) = -\frac{1}{4(1-\bar{y}_n)^2}$, so that (71)-(73) together imply Step 3.[60]
■

Combining Step 2 and Step 3, we prove the next step.

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

Proof. We rewrite the conditional expectation in integral form,

$$\mathrm{E}(x_n(t(y,k))|g,k) = \int_{t:x_n(t)>0} x_n(t)dH(t|g,k). \tag{74}$$

Given Step 2, we have $x_n(t) = (1+\epsilon_n(t))x_n(t(\bar{y}_n,k)$ for some sequence $\epsilon_n(t)$ that

[60]Consider $\chi(y) = \frac{-\bar{y}_n y}{\bar{y}_n(1-y)+(1-\bar{y}_n)y}$. Then, $\chi'(y) = \frac{-\bar{y}_n}{\bar{y}_n(1-y)+(1-\bar{y}_n)y} + \frac{\bar{y}_n y}{(\bar{y}_n(1-y)-(1-\bar{y}_n)y)^2}(1-2\bar{y}_n)$. Plugging in $y=\bar{y}_n$ gives $\chi'(\bar{y}_n) = \frac{-\bar{y}_n}{2\bar{y}_n(1-\bar{y}_n)} + \frac{\bar{y}_n(1-2\bar{y}_n)}{2\bar{y}_n(1-\bar{y}_n))^2} = \frac{-1}{2(1-\bar{y}_n)} + \frac{(1-2\bar{y}_n)}{4(1-\bar{y}_n)^2} = \frac{-1}{4(1-\bar{y}_n)}$.

converges to zero as $n \to \infty$. Hence,

$$\begin{aligned} & \mathrm{E}(x_n(t(y,k))|g,k) \\ = \;& x_n(t(\bar{y}_n,k)\Pr(\{t : x_n(t) > 0\}|g,k) + x_n(t(\bar{y}_n,k)) \int_{t:x_n(t)>0} \epsilon_n(t) dH(t|g,k) \end{aligned} \tag{75}$$

Further,

$$\begin{aligned} & |\int_{t:x_n(t)>0} \epsilon_n(t) dH(t|g,k)| \\ \leq \;& \int_{t:x_n(t)>0} |\epsilon_n(t)| dH(t|g,k) \\ \leq \;& \Pr(\{t : x(t) > 0\}|g,k) M_n(y_g^+(k) - y_g^-(k)), \end{aligned} \tag{76}$$

for $M_n = \max_{y' \in [\psi_g(k), \phi_g(k)]} |M_n(y')|$. The first inequality follows from the triangle inequality. For the second inequality, we use that $\epsilon_n(t) = M_n(y')(\bar{y}_n - y')$ given (70). For the third inequality, 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 $\lim_{n\to\infty} y^+(k) - y^-(k) \to 0$. Since $\lim_{n\to\infty} M_n \in \mathbb{R}$ (recall (69) 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 (76). Combining this with (75),

$$\mathrm{E}(x_n(t(y,k))|g,k) \approx x_n(t(\bar{y}_n,k)) \Pr(\{t : x(t) > 0\}|g,k). \tag{77}$$

Using Step 3,

$$\mathrm{E}(x_n(t(y,k))|g,k) \approx x_n^2(t(\bar{y}_n,k)) f(t(\bar{y}_n,k)|g,k) e_2(\bar{y}_n,d). \tag{78}$$

Step 4 follows since e is continuous so that $\lim_{n\to\infty} e_2(\bar{y}_n,d) = e_2(\bar{y},d)$. ∎

Recall (32) 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 (32) and Step 4,

$$\frac{\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}} \approx \left[f(t(\bar{y}_n,k)|g,k)k^{\frac{2}{d-1}}\right] e_3(\bar{y}_n,d). \tag{79}$$

for $e_3(\bar{y}_n,d) = e_2(\bar{y},d)e(\bar{y}_n)^{\frac{2}{d-1}}$. 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:

$$\begin{aligned}
\lim_{n\to\infty}\frac{\mathrm{E}(x_n(t(y,k))|g)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}} &= \lim_{n\to\infty}\frac{\mathrm{E}(\mathrm{E}(x_n(t(y,k))|g,k)|g)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}} \\
&= \lim_{n\to\infty}\int_k \frac{f(k|g)\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}}dk \\
&= \int_k \lim_{n\to\infty}\frac{f(k|g)\mathrm{E}(x_n(t(y,k))|g,k)}{\Pr(\mathrm{piv}|\sigma_n,n)^{\frac{2}{d-1}}}dk \\
&= \int_k \lim_{n\to\infty} f(k|g)f(t(\bar{y}_n,k)|g,k)k^{\frac{2}{d-1}}e_3(\bar{y}_n,d)dk \\
&= e_3(\bar{y},d)\int_k f(k|g)f(t(\bar{y},k)|g,k)k^{\frac{2}{d-1}}dk \\
&= e_3(\bar{y},d)\int_k f(\bar{y}|g)f(t(\bar{y},k)|g,\bar{y})k^{\frac{2}{d-1}}dk \\
&= e_3(\bar{y},d)f(\bar{y}|g)\mathrm{E}(k^{\frac{2}{d-1}}|g,\bar{y}) \qquad (80)
\end{aligned}$$

for $e(\bar{y}_n,d)=e(d)e(\bar{y}_n)^{\frac{2}{d-1}}$. 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 (79). The fifth equality follows from $\bar{y}_n \to \bar{y}$ as $n\to\infty$ and since e and $f(-|g,k)$ are continuous. The sixth equality follows since Bayes law implies $f(k|g)f(t(\bar{y},k)|g,k) = f(t(\bar{y},k)|g,\bar{y})f(\bar{y}|g)$. The last inequality rewrites the integral as a conditional expectation. Finally, the state-dependent intensity of the limit marginal types $t(\bar{y}_n,k)$ is linear in the total intensity, $t_\alpha = k(1-\bar{y}_n)$, 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 (80), this shows (31).

D Proof of Claim 1

The proof follows arguments similar to those in Section 4.2. There, we discussed why the condition $d>3$ is the critical condition for the severity of the free-rider problem in a large electorate. Most of the proof restates the observations from Section 4.2.

We note that the condition of Lemma 3 is satisfied for any sequence of vote share pairs $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ with $q_n(\beta) = q_n^-(q_n(\alpha))$ for all n or with $q_n(\beta) = q^+(q_n(\alpha))$ for all n. This is because for $q_n(\beta) \in \{q_n^-(q_n(\alpha)), q_n^+(q_n(\alpha))\}$, by construction, the implied vote share under the best response, $q(\beta;\sigma_n^{\mathbf{q}_n})$, lies in

$[\frac{1}{2} - 2\epsilon, \frac{1}{2} + 2\epsilon]$, see Lemma 4. Since only a vanishing fraction of types acquires information, the definition (2) together with (9) and (10) implies that this vote shares converges to $\Phi(\Pr(\alpha|\text{piv}; \sigma_n^{\mathbf{q}_n}, n))$. The continuity of Φ implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) \in [\Phi^{-1}(\frac{1}{2}-2\epsilon), \Phi^{-1}(\frac{1}{2}+2\epsilon)]$. In particular, $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) \in (0, 1)$.

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 as a consequence of Lemma 3,

$$\lim_{n\to\infty} \frac{\mathrm{E}(x(t)|g)}{\Pr(\text{piv}|\mathbf{q}_n, n)} = \infty. \tag{81}$$

for $g \in \{L, C\}$. The second observation is that the approximation (25) also holds locally (as illustrated in Figure 4),[61]

$$\lim_{n\to\infty} \Pr(\text{piv}|\omega; \mathbf{q}_n)(2n+1)s(\alpha; \mathbf{q}_n) \quad = \quad \lim_{n\to\infty} \phi(\delta_n(\omega)), \tag{82}$$

where ϕ the density of the standard normal distribution and $\omega \in \{\alpha, \beta\}$. Let $s_n = s(\omega; \mathbf{q}(\sigma_n))$ and $q_n = q(\omega_n; \sigma_n)$. Given (82), the pivotal likelihood is a finite multiple of $((2n+1)s_n)^{-1} = s_n(q_n(1-q_n)^{-1}$, so, a finite multiple of the standard deviation.[62] Given the assumption of the lemma, that is, $\lim_{n\to\infty} \delta_n(\alpha) = \lim_{n\to\infty} \frac{|q_n(\alpha)-\frac{1}{2}|}{s(\alpha;\mathbf{q}_n} \in \mathbb{R}$, we have $\lim_{n\to\infty} \phi(\delta_n(\omega)) \in \mathbb{R}$. Combining this with (81) and (82),

$$\lim_{n\to\infty} \frac{\mathrm{E}(x(t)|g)}{s(\alpha; \mathbf{q}_n)} = \infty. \tag{83}$$

Recall (37),

$$\begin{aligned} &\lim_{n\to\infty} \frac{q(\alpha; \sigma^{\mathbf{q}_n}) - q(\beta; \sigma^{\mathbf{q}_n})}{s(\alpha; \mathbf{q}_n)} \\ = &\lim_{n\to\infty} \frac{2(\Pr(L)\mathrm{E}(x_n(t)|L) - \Pr(C)\mathrm{E}(x_n(t)|C))}{s(\alpha; \mathbf{q}_n)}. \end{aligned} \tag{84}$$

Lemma 3 together with the genericity condition $W(L, \kappa, \alpha) \neq W(C, \kappa, \alpha)$ implies

[61] 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.

[62] 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 (24) and thereafter.

that either $\Pr(L)\mathrm{E}(x_n(t)|L) > \Pr(C)\mathrm{E}(x_n(t)|C)$ for all n sufficiently large, or $\Pr(L)\mathrm{E}(x_n(t)|L) < \Pr(C)\mathrm{E}(x_n(t)|C)$ for all n sufficiently large. Together with (83) and (84),

$$\lim_{n\to\infty} \frac{q(\alpha;\sigma^{\mathbf{q}_n}) - q(\beta;\sigma^{\mathbf{q}_n})}{s(\alpha;\sigma^{\mathbf{q}_n})} \in \{\infty, -\infty\}, \tag{85}$$

which is equivalent to (46). This finishes the proof of Claim 1 since $q(\beta;\sigma^{\mathbf{q}_n}) = q_n(\beta)$ for $q_n(\beta) \in \{q_n^-(q_n(\alpha)), q_n^+(q_n(\alpha))\}$.

E Voter inference

Lemma 6 *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, $\lim_{n\to\infty} \frac{\Pr(\mathrm{piv}|\alpha;\sigma_n,n)}{\Pr(\mathrm{piv}|\beta;\sigma_n,n)} = \infty$.*

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, $\lim_{n\to\infty} \frac{\Pr(\mathrm{piv}|\alpha;\sigma_n,n)}{\Pr(\mathrm{piv}|\beta;\sigma_n,n)} = 0$.*

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 $\delta_n(\alpha) - \delta_n(\beta)$ converges in the extended reals $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$, then, $\lim_{n\to\infty} \frac{\Pr(\mathrm{piv}|\alpha;\sigma_n,n)}{\Pr(\mathrm{piv}|\beta;\sigma_n,n)} = \lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} \in \bar{\mathbb{R}}$, 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 (8), $\frac{\Pr(\mathrm{piv}|\alpha;\sigma_n,n)}{\Pr(\mathrm{piv}|\beta;\sigma_n,n)} = (k_n)^n$. 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

$$\begin{aligned}
\hat{\delta}_n(\omega) &= \frac{2n+1}{s(\omega;\mathbf{q}(\sigma_n))}(q(\omega;\sigma_n) - \frac{1}{2}) \\
&= (2n+1)^{\frac{1}{2}} \frac{q(\omega;\sigma_n) - \frac{1}{2}}{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}.
\end{aligned} \tag{86}$$

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

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

where we used (86) for the equality on the last line.

Case 1 $\lim_{n\to\infty}\delta_n(\alpha)-\delta_n(\beta)\in\mathbb{R}$.

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

$$\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)}=\left[(1-\frac{1}{2n+1}x_n)^n-e^{-\frac{1}{2}x_n}\right]+e^{-\frac{1}{2}x_n}. \tag{87}$$

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

$$|(1-\frac{x_n}{(2n+1)})^n-e^{-x_n}|\le\frac{x_n^2}{(2n+1)^3}. \tag{88}$$

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,

$$\begin{aligned}
\lim_{n\to\infty}\frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} &= \lim_{n\to\infty}e^{-\frac{1}{2}x_n} \\
&= e^{\lim-\frac{1}{2}x_n} \\
&= \lim_{n\to\infty}\frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))}.
\end{aligned} \tag{89}$$

where we used that $\lim_{n\to\infty}\lim_{n\to\infty}\left|q(\alpha;\sigma_n)-\frac{1}{2}\right|=\lim_{n\to\infty}\left|q(\beta;\sigma_n)-\frac{1}{2}\right|$ 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 2 $\lim_{n\to\infty}\delta_n(\alpha)-\delta_n(\beta)=\infty$.

Then, given (87),

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\sigma_n,n)}{\Pr(\text{piv}|\beta;\sigma_n,n)} \geq \lim_{n\to\infty}(1-\frac{x}{n})^n = e^{-x} \tag{90}$$

for all $x \in \mathbb{R}$. 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 $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} = e^{-\frac{1}{2}\delta_n(\alpha)^2-\delta_n(\beta)^2} = 0$, given $\lim_{n\to\infty} \delta_n(\alpha)-\delta_n(\beta) = \infty$.

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

Then, given (87),

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

for all $x \in \mathbb{R}$. 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 $\lim_{n\to\infty} \delta_n(\alpha)-\delta_n(\beta) = -\infty$. ■

F 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 $W(\kappa, L, \alpha) \neq W(\kappa, C, \alpha)$ and the proof of the second item for the case when $W(\kappa, L, \alpha) \neq W(\kappa, C, \alpha)$ and $\Phi(\Pr(\alpha)) < \frac{1}{2}$.

The proof given in the main text has proceeded as follows: First, we established Claim 1 - Claim 3. Second, we provide fixed point arguments to construct informative equilibrium sequences. For the remaining cases, the approach is the same: based on Claim 1 - Claim 3, one uses fixed point arguments to construct informative equilibrium sequences. The fixed point arguments are analogous to those given in the main text. For the convenience, we provide the fixed point argument for a further case.

Let $d > 3$. Take a distribution of preferences types so that $W(\kappa, L, \alpha) > W(\kappa, C, \alpha)$. We show that there is an equilibrium sequence in which the full information outcome is elected with probability converging to 1 as $n \to \infty$, thereby finishing the proof of the first item of Theorem 1. Let the electorate be sufficiently large, so that for a candidate equilibrium vote share $q_n(\alpha)$ from an interval close to the majority threshold as in Lemma 4, there are two solutions $q_n^-(q(\alpha))$ and $q_n^+(q(\alpha))$ to (41). Denote $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ with $q_n(\beta) \in \{q_n^-(q(\alpha)), q_n^+(q(\alpha))\}$.

Take any sequence $(\Delta_n)_{n\in\mathbb{N}}$ for which $\lim_{n\to\infty} \Delta_n \in [-\Delta, \Delta]$ with a $\Delta > 0$ as in Lemma 4. For $q_n(\alpha) = \frac{1}{2} + \Delta_n n^{-\frac{1}{2}}$, Claim 3 implies $\lim_{n\to\infty} \frac{q_n(\beta)-q_n(\alpha)}{s(\alpha;\sigma^{\mathbf{q}_n})} \in \mathbb{R}$. Since $q(\alpha;\sigma^{\mathbf{q}_n}) > q(\beta;\sigma^{\mathbf{q}_n})$ by (45) and since $q_n(\beta) = q(\beta;\sigma^{\mathbf{q}_n})$ by (41), it follows from Claim 1 that

$$q_n(\alpha) < q(\alpha;\sigma^{\mathbf{q}_n}) \tag{92}$$

for n large enough. Take $q_n(\alpha) = \frac{1}{2}+\epsilon$ and let $q_n(\beta) = q_n^+(q_n(\alpha))$; hence, $q_n(\beta) < \frac{1}{2}$. Given Claim 3, $q_n(\beta) \to \frac{1}{2} - \epsilon$ and, given Claim 2, $q_n(\alpha;\sigma^{\mathbf{q}_n}) \to \frac{1}{2} - \epsilon$. Together,

$$q_n(\alpha) > q(\alpha;\sigma^{\mathbf{q}_n}) \tag{93}$$

for n large enough. Finally, using (92)- (93), an application of the intermediate value theorem shows that there is $q_n^*(\alpha) < \frac{1}{2}$ so that $q_n^* = (q_n^*(\alpha), q_n^+(q_n^*(\alpha)))$ solves (40) and (41). Further, it must be that $\lim_{n\to\infty} \frac{q_n^*(\alpha)-\frac{1}{2}}{s(\alpha;\mathbf{q}_n)} = \infty$ since otherwise (92) holds as we just argued. Hence, also $\lim_{n\to\infty} \frac{\frac{1}{2}-q_n(\beta)}{s(\beta;\mathbf{q}_n)} = \infty$, given Claim 3. Hence, 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 (25).

References

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

Ahn, D. S. and Oliveros, S. (2012). Combinatorial voting. *Econometrica*, **80** (1), 89–141.

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

Allcott, H., Braghieri, L., Eichmeyer, S. and Gentzkow, M. (2020). The welfare effects of social media. *American Economic Review*, **110** (3), 629–76.

Alonso, R. and Câmara, O. (2016). Persuading voters. *American Economic Review*, **106** (11), 3590–3605.

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

Bardhi, A. and Guo, Y. (2018). Modes of persuasion toward unanimous consent. *Theoretical Economics*, **13** (3), 1111–1149.

Barelli, P., Bhattacharya, S. and Siga, L. (2017). On the possibility of information aggregation in large elections. In *Working Paper*.

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. *Available at SSRN 3245949*.

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.

Brennan, J. (2016). *Against democracy*. Princeton University Press.

Bridgman, A., Merkley, E., Loewen, P. J., Owen, T., Ruths, D., Teichmann, L. and Zhilin, O. (2020). The causes and consequences of covid-19 misperceptions: Understanding the role of news and social media. *Harvard Kennedy School Misinformation Review*, **1** (3).

Casella, A., Llorente-Saguer, A. and Palfrey, T. R. (2012). Competitive equilibrium in markets for votes. *Journal of Political Economy*, **120** (4), 593–658.

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

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

Eguia, J. X. and Xefteris, D. (2018). Implementation by vote-buying mechanisms.

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

Esteban, J. and Ray, D. (2006). Inequality, lobbying, and resource allocation. *American Economic Review*, **96** (1), 257–279.

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.

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

HEESE, C. and LAUERMANN, S. (2017). *Persuasion and Information Aggregation in Elections*. Tech. rep., Working Paper.

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.

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. (2017). Electoral competition with rationally inattentive voters. *Available at SSRN 3070204*.

OLIVEROS, S. (2013). Aggregation of endogenous information in large elections.

OLSON, M. (1965). *The Logic Of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

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

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

TULLOCK, G. (1983). *Economics of income redistribution*. Boston, MA: Kluwer-Nijhoff.

Online appendix

G Proof of Lemma 4

G.1 Existence

Lemma 7 *There is $\Delta > 0$ and $\epsilon > 0$, so that for any $q_n(\alpha) \in \mathrm{B}_\epsilon \setminus \mathrm{B}_{\frac{\Delta}{\sqrt{n}}}(\frac{1}{2})$, the following holds.*

1. *There is $\bar{n} \in \mathbb{N}$, so that for $n \geq \bar{n}$ there is $q_n(\beta) \in [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$ solving* (41).
2. *For any sequence of solutions $q_n(\beta)$ of* (41) *for which the pivotal likelihood ratio converges in the extended reals and with $q_n(\beta) \in [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$,*

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) \in [\Phi^{-1}(\frac{1}{2}), \Phi^{-1}(\frac{1}{2} + 2\epsilon)]. \tag{94}$$

Proof. 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 a standard deviation $\frac{\sqrt{2n+1}}{q_n(\alpha)(1-q_n(\alpha))}$ of the (empirical) vote share distribution, it follows from Lemma 6 that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)} \leq \frac{\phi(\Delta')}{\phi(0)}$. Hence, for any prior $\Pr(\alpha) \in (0, 1)$, there is $\Delta > 0$ large enough, so that

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) < \Phi^{-1}(\frac{1}{2}) \tag{95}$$

if $q_n(\beta) = \frac{1}{2}$ and for any $q_n(\alpha) \in \mathrm{B}_\epsilon \setminus \mathrm{B}_{\frac{\Delta}{\sqrt{n}}}(\frac{1}{2})$. Take $q_n(\beta) = \frac{1}{2} + \frac{\epsilon}{2}$. The election is more close to being tied in α, and, by Lemma 6, voters become convinced that the state is α, i.e.,

$$\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \mathbf{q}_n, n) = 1. \tag{96}$$

Recall (134), which states that $\lim_{n\to\infty} q(\beta_n; \sigma_n) = \lim_{n\to\infty} \Phi(\Pr(\alpha|\text{piv}; \mathbf{q}_n, n))$ and holds since the precision acquired by all types converges to zero uniformly. This immediately implies that any solution $q_n(\beta)$ of (41) for which $q_n(\beta) \in [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$ must satisfy (94), proving the second item of Claim 7.

Further, we see that (95) implies that for $q_n(\beta) = \frac{1}{2}$, it holds $q_n(\beta) > q(\beta_n; \sigma^{\mathbf{q}_n})$ when n is large enough; since $\Phi(1) > \frac{1}{2} + \epsilon$, (96) implies that for $q_n(\beta) = \frac{1}{2} + \frac{\epsilon}{2}$, it holds $q_n(\beta) < q(\beta_n; \sigma^{\mathbf{q}_n})$ when n is large enough. Existence of a vote share $q_n(\beta)$

that solves (41) follows from the continuity of $q(\beta;\sigma^{\mathbf{q}_n}) - q_n(\beta)$ in $q_n(\beta)$ and an application of the intermediate value theorem. ■

G.2 Uniqueness

Lemma 8 *There are $\epsilon > 0$ and $\Delta > 0$, so that the following holds. Take a sequence $(q_n(\alpha))_{n\in\mathbb{N}}$ of vote shares with*

$$q_n(\alpha) \in \mathrm{B}_\epsilon \setminus \mathrm{B}_\Delta \sqrt{n}. \tag{97}$$

Then, for any sequence $(q_n(\beta))_{n\in\mathbb{N}}$ of solutions to (41) given $q_n(\alpha)$ with $q_n(\beta) \in [\frac{1}{2}, \frac{1}{2}+\epsilon]$, there is $\bar{n} \in \mathbb{N}$, so that for all $n \geq \bar{n}$,

$$q_n(\beta) \in \mathrm{B}_{\frac{\Delta}{2\sqrt{n}}}(q_n(\alpha)) \cup \mathrm{B}_{\frac{\Delta}{2\sqrt{n}}}(1-q_n(\alpha)). \tag{98}$$

Proof. Take $\epsilon > 0$. For any $\Delta > 0$, we denote the set of vote share pairs $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ that satisfy (97) and (98) by $Q_n(\Delta)$. Fix a sequence $(q_n(\alpha))_{n\in\mathbb{N}}$ of vote shares in α. For any sequence $(q_n(\beta))_{n\in\mathbb{N}}$ of vote shares in β, so that the pivotal likelihood ratio converges (in the extended reals), it follows from the third item of Lemma 6 that $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)} = \lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))}$ where ϕ is the density of the standard normal. Since $\frac{\phi(x)}{\phi(y)} = e^{-\frac{1}{2}(x^2-y^2)}$,

$$\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)} = \lim_{n\to\infty} e^{-2(\delta_n(\alpha)^2-\delta_n(\beta)^2)}. \tag{99}$$

If $q_n(\beta)$ are solutions to (41), then, (94) holds. Together with (99), this implies that $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) \in \mathbb{R}$. Plugging in the definition of $\delta_n(\omega)$ (see (24)),

$$\lim_{n\to\infty} \frac{|q_n(\alpha)-\frac{1}{2}|}{s(\alpha;\mathbf{q}_n)} - \frac{|q_n(\beta)-\frac{1}{2}|}{s(\beta;\mathbf{q}_n)} \in \mathbb{R}, \tag{100}$$

for $s(\omega;\mathbf{q}_n) = \frac{q_n(\omega)(1-q_n(\omega))}{(2n+1)^{\frac{1}{2}}}$. This shows that $q_n(\beta)$ is at most finitely many multiples of $n^{-\frac{1}{2}}$ further away from $\frac{1}{2}$ than $q_n(\alpha)$ for any n sufficiently large. Hence, (98) holds for some $\Delta > 0$ large enough. ■

To show uniqueness, we analyze the derivative of $q(\beta;\sigma^{\mathbf{q}_n}) - q_n(\beta)$. First, we decompose $q(\beta;\sigma^{\mathbf{q_n}})$,

$$q(\beta;\sigma^{\mathbf{q}_n}) = \Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n,n)) + \epsilon_n(\mathbf{q}_n), \tag{101}$$

where $\Phi(\Pr(\alpha|\text{piv};\mathbf{q}_n,n))$ is the likelihood of a random citizen voting A under the constrained best response where no information acquisition is possible; hence, $\epsilon_n(\mathbf{q}_n)$ captures the difference to the likelihood under the actual best response. In the following, we evaluate the derivatives of each of the two summands of (101), one after the other.

Lemma 9 *There is a constant $M > 0$ and $\bar{n} \in \mathbb{N}$, so that for $\mathbf{q}_n \in Q_n(\Delta)$, $n \geq \bar{n}$, and $q = q_n(\beta)$,*

$$\frac{d}{dq}\Pr(\alpha|\text{piv};\mathbf{q}_n,n) \geq n(2q-1)M \tag{102}$$

Proof. First, consider the likelihood ratio $\frac{\Pr(\beta|\text{piv};\mathbf{q}_n,n)}{\Pr(\alpha|\text{piv};\mathbf{q}_n,n)} = \frac{\Pr(\beta)}{\Pr(\alpha)}\frac{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)}{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)} = \frac{\Pr(\beta)}{\Pr(\alpha)}\left[\frac{q_n(\alpha)(1-q_n(\alpha))}{q_n(\beta)(1-q_n(\beta))}\right]^n$. Taking the derivative at $q = q_n(\beta)$,

$$\begin{aligned} & n\frac{\Pr(\beta)}{\Pr(\alpha)}\left[\frac{q(1-q)}{q_n(\alpha)(1-q_n(\alpha))}\right]^{n-1}(1-2q) \\ = \; & n\frac{\Pr(\beta|\text{piv};\mathbf{q}_n,n)}{\Pr(\alpha|\text{piv};\mathbf{q}_n,n)}(1-2q)\frac{q_n(\alpha)(1-q_n(\alpha))}{q(1-q)}. \end{aligned} \tag{103}$$

Second, consider the map that sends a likelihood ratio $\ell = \frac{1-p}{p}$ to the belief $p = \frac{1}{l+1}$. The derivative of this map is $-p^2$. Together, for $p = \Pr(\alpha|\text{piv};\mathbf{q}_n,n)$,

$$\frac{d}{dq}\Phi(p) = n(2q-1)\frac{q_n(\alpha)(1-q_n(\alpha))}{q(1-q)}p.$$

Since $\mathbf{q}_n \in Q_n(\Delta)$, the ratio $\frac{q_n(\alpha)(1-q_n(\alpha))}{q(1-q)}$ and $p = \Pr(\alpha|\text{piv};\mathbf{q}_n,n)$ are bounded below by a positive constant when n is large, so that the claim (102) follows. ∎

Lemma 10 *Take any sequence $(\mathbf{q}_n)_{n\in\mathbb{N}}$ with $\mathbf{q}_n \in Q_n(\Delta)$ and for which the pivotal likelihood $\Pr(\alpha|\text{piv};\mathbf{q}_n,n)$ converges in $(0,1)$. For $q = q_n(\beta)$,*

$$\lim_{n\to\infty}\frac{|\frac{d}{dq}\epsilon_n(\mathbf{q}_n))|}{\frac{d}{dq}\Pr(\alpha|\text{piv};\mathbf{q}_n,n))} = 0. \tag{104}$$

In the following, $x_n(t)$ is the precision of a type under the best response $\sigma^{\mathbf{q}_n}$ to $\mathbf{q}_n$. Similarly, $y_g^-(k), y_g^+(k)$ denote the threshold of doubt types that are indifferent between no information acquisition and some information acquisition under the best response, et cetera.

Proof.

Step 1

$$|\frac{d}{dq}\epsilon_n(\mathbf{q}_n)| \leq \sum_{g\in\{L,C\}} |\frac{d}{dq}\Pr(\{t : t \in g, x_n(t) > 0\})| + |\int_{t:x_n(t)>0} \frac{d}{dq} x^*(t; \mathbf{q}_n, n) dH(t)| \quad (105)$$

Recall that ϵ_n captures the effect of the signals acquired under the best response on voting behaviour: for example, some aligned types with threshold of doubt below the indifferent types ware swayed to vote B when receiving a b-signal, et cetera. Formally,

$$\begin{aligned} \epsilon_n(q) &= -\int_{t\in L, y(t)\in[y_L^-(k),p]} \frac{1}{2} + x(t)dH(t) + \int_{t\in L, y(t)\in[p,y_L^+(k)]} \frac{1}{2} - x(t)dH(t) \\ &\quad + \int_{t\in C, y(t)\in[y_C^-(k),p]} \frac{1}{2} + x(t)dH(t) - \int_{t\in C, y(t)\in[p,y_C^+(k)]} \frac{1}{2} - x(t)dH(t). \end{aligned}$$

for $p = \Pr(\alpha; \mathrm{piv}; \mathbf{q}_n, n)$. Using the triangle inequality,

$$|\frac{d}{dq}\epsilon_n| \leq \frac{1}{2}\sum_{g\in\{L,C\}} |\frac{d}{dq}\Pr(\{t : t \in g, x(t) > 0\})| + |\frac{d}{dq}\int_{t:x_n(t)>0} x_n(t)dH(t)|,$$

and recalling that $x_n(t) = x^*(t; \sigma^{\mathbf{q}_n}, n)$ if $x(t) > 0$, we conclude that (105) holds.

Step 2 $\frac{d}{dq}\Pr(\mathrm{piv}|\mathbf{q}_n, n) = n\frac{q(1-q)}{(1-2q)}\Pr(\beta)\Pr(\mathrm{piv}|\beta; \mathbf{q}_n, n)$.

Since only the pivotal likelihood in β depends on the vote share $q = q_n(\beta)$,

$$\frac{d}{dq}\Pr(\mathrm{piv}|\mathbf{q}_n, n) = \Pr(\beta)\frac{d}{dq}\Pr(\mathrm{piv}|\beta; \mathbf{q}_n, n) \quad (106)$$

Given (8), $\Pr(\mathrm{piv}|\beta; \mathbf{q}_n, n) = \binom{2n}{n}(q(1-q))^n$. Taking the derivative,

$$\frac{d}{dq}\Pr(\mathrm{piv}|\beta; \mathbf{q}_n, n) = n\frac{(1-2q)}{q(1-q)}\Pr(\mathrm{piv}|\beta; \mathbf{q}_n, n), \quad (107)$$

and (106) together with (107) show the claim of Step 2.

Step 3 *Take any sequence $(\mathbf{q}_n)_{n\in\mathbb{N}}$ with $\mathbf{q}_n \in Q_n(\Delta)$ and for which the pivotal* *likelihood* $\Pr(\alpha|\text{piv};\mathbf{q}_n,n)$ *converges in* $(0,1)$*. For* $q=q_n(\beta)$,

$$\lim_{n\to\infty}\frac{\max_t|\frac{d}{dq}x^*(t;\sigma^{\mathbf{q}_n},n)|}{\frac{d}{dq}\Pr(\alpha|\text{piv};\mathbf{q}_n,n))}=0, \tag{108}$$

$$\lim_{n\to\infty}\frac{|\frac{d}{dq}\Pr(\{t:t\in g,x(t)>0\})|}{\frac{d}{dq}\Pr(\alpha|\text{piv};\mathbf{q}_n,n))}=0. \tag{109}$$

Given (17) and (19), and since $k(t)$ is uniformly bounded, there is a constant $M>0$, so that for all $t\in g$ and $g\in\{L,C\}$,

$$\frac{d}{dq}x^*(t;\sigma^{\mathbf{q}_n},n)\leq\frac{d}{dq}\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}}M. \tag{110}$$

By the chain rule,

$$\frac{d}{dq}\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}}=\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}-1}\frac{d}{dq}\Pr(\text{piv}|\beta;\mathbf{q}_n,n),$$

so that, given Step 2,

$$\frac{d}{dq}\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}}=\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}}n\frac{(1-2q)}{q(1-q)}\Pr(\beta).$$

Since $\lim_{n\to\infty}\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}}=0$,

$$\lim_{n\to\infty}\frac{|\frac{d}{dq}\Pr(\text{piv}|\beta;\mathbf{q}_n,n)^{\frac{1}{d-1}}|}{n\frac{(2q-1)}{q(1-q)}}=0. \tag{111}$$

This observation together with (110) implies $\lim_{n\to\infty}\frac{\max_t|\frac{d}{dq}x^*(t;\sigma^{\mathbf{q}_n},n)|}{n\frac{(2q-1)}{q(1-q)}}=0$. Then, (108) follows from Lemma 9.

Fix a fixed total intensity $k(t)=k$. Denote by $G(-|g,k)$ the conditional distribution of the threshold of doubt $y(t)$.[63] Then,

$$\Pr(\{t:x(t)>0\}|T_g,T_k)=G(\phi_g^+(k)|g,k)-G(\phi_g^-(k)|g,k). \tag{112}$$

To simplify notation, let in the following $y\in\{y_g^+(k),y_g^-(k)\}$ and $p=\Pr(\alpha|\text{piv};\mathbf{q}_n,n)$.

[63] Recall that we use the shortcut notation $\mathrm{E}(-|g,k)$ and similar for the expectation conditional on the set of types $\{t:t\in g\}$ and the set of types $\{t:k(t)=k\}$.

We totally differentiate and use the triangle inequality,

$$\begin{aligned}
& |\frac{d}{dq}(G(y|g,k) - G(p|g,k))| \\
= \ & |(\frac{d}{dq}y)G'(y|g,k) - (\frac{d}{dq}p)G'(p|g,k)| \\
\leq \ & |\frac{d}{dq}(y-p)G'(y|g,k)| + |(\frac{d}{dq}p)(G'(y|T_g,T_k) - G'(p|g,k))|. \qquad (113)
\end{aligned}$$

Now, suppose that

$$\lim_{n\to\infty} \frac{|\frac{d}{dq}(y-p)|}{\frac{d}{dq}p} = 0, \qquad (114)$$

Note that $G'(y|g,k) - G'(p|g,k) \to 0$ as $n \to \infty$ since $y \to p$ and since the conditional type distribution has a continuous density.[64] Then, (112)-(114) imply

$$\lim_{n\to\infty} \frac{|\frac{d}{dq}\Pr(\{t : x_n(t) > 0\}|g,k)|}{\frac{d}{dq}p} = 0 \qquad (115)$$

for all k and g. Take the conditional distribution $H(-|g)$ denote of $k(t)$.

$$\begin{aligned}
& \lim_{n\to\infty} \frac{|\frac{d}{dq}\Pr(\{t : t \in g, x_n(t) > 0\})|}{\frac{d}{dq}p)} \\
= \ & \lim_{n\to\infty} |\int_k \frac{\frac{d}{dq}\Pr(\{t : x_n(t) > 0\}|g,k) dH(k|g)}{\frac{d}{dq}p)}| \\
\leq \ & \lim_{n\to\infty} \int_k |\frac{\frac{d}{dq}\Pr(\{t : x_n(t) > 0\}|g,k)}{\frac{d}{dq}p)}| dH(k|g) \\
= \ & 0, \qquad (116)
\end{aligned}$$

where the equality on the second line follows by an application of Leibnitz integral rule, the inequality on the third line by the triangle inequality, the equality on the last line by the dominated convergence theorem and (115). Note that (116) restates (109), which is what we aim to show.

To finish the proof of (109), we establish (114) in the following. W.l.o.g., take $y = y_g^+(k)$ and let $x = x^*(t(y,k); \mathbf{q}_n, n)$. We recall the indifference equation (23)

[64] This follows from inspection of the indifference condition (23) and since $x^*(t) \to 0$ uniformly as $n \to \infty$.

that implicitly defines y,

$$-\chi(y) = \frac{1}{2} + \frac{d-1}{d}x, \tag{117}$$

and rewrite it: given the definition of χ, $\frac{\chi(y)}{1-\chi(y)} = \frac{p}{1-p}\frac{y}{1-y}$, so that (117) yields

$$\Leftrightarrow F(y,p,x) := \frac{y}{1-y} - \frac{\frac{1}{2}+\frac{d-1}{d}x}{\frac{1}{2}-\frac{d-1}{d}x}\frac{p}{1-p} = 0. \tag{118}$$

Implicit differentiation gives $\frac{d}{dq}yF_y + \frac{d}{dq}pF_p + \frac{d}{dq}xF_x = 0$. Rearranging and applying of the triangle inequality,

$$|\frac{d}{dq}(y-p)F_y| \leq |\frac{d}{dq}p(F_p+F_y)| + |(\frac{d}{dq}x)F_x|. \tag{119}$$

Note that $F_y = \frac{1}{(1-y)^2}$ and that $F_p = -\frac{1}{(1-p)^2}$. This implies that $F_y \to -F_p$ since $y \to p$ as $n \to \infty$ (see for example Step 1 in the proof of Lemma 3). Therefore,

$$\lim_{n\to\infty} \frac{|\frac{d}{dq}(y-p)F_y|}{\frac{d}{dq}p} \leq \lim_{n\to\infty} \frac{|(\frac{d}{dq}x)F_x|}{\frac{d}{dq}p}. \tag{120}$$

In the following, we analyse $\frac{(\frac{d}{dq}x)F_x}{\frac{d}{dq}p}$. Recall (17) and (19): for $x = x^*(t(y,k);\mathbf{q}_n,n)$ and $y = y_g^+(k)$,

$$x = \Pr(\text{piv}|\mathbf{q}_n,n)^{\frac{1}{[}d-1}(ke(y))^{\frac{1}{d-1}}. \tag{121}$$

Using the product rule of differentiation,

$$\frac{d}{dq}x = (\frac{d}{dq}\Pr(\text{piv}|\mathbf{q}_n,n)^{\frac{1}{d-1}})(ke(y))^{\frac{1}{d-1}} + (\frac{d}{dq}y)e'(y)(\Pr(\text{piv}|\mathbf{q}_n,n)k)^{\frac{1}{d-1}} \tag{122}$$

It follows from Step 1 and Lemma 9 that

$$\lim_{n\to\infty} \frac{\frac{d}{dq}\Pr(\text{piv}|\mathbf{q}_n,n)^{\frac{1}{d-1}}}{\frac{d}{dq}p} = 0, \tag{123}$$

so that

$$\lim_{n\to\infty} \frac{|\frac{d}{dq}x|}{\frac{d}{dq}p} \leq \lim_{n\to\infty} \frac{(\frac{d}{dq}y)\Pr(\text{piv}|\mathbf{q}_n,n)^{\frac{1}{d-1}}\bar{M}}{\frac{d}{dq}p} \tag{124}$$

for some constant $\bar{M} > 0$. Using the triangle inequality,

$$|\frac{d}{dq}y| \leq |\frac{d}{dq}(y-p)| + |\frac{d}{dq}p|, \tag{125}$$

Since $\lim_{n\to\infty} \Pr(\mathrm{piv}|\mathbf{q}_n, n) = 0$ (see e.g. (27)),

$$\lim_{n\to\infty} \frac{|\frac{d}{dq}x|}{\frac{d}{dq}p} \leq \lim_{n\to\infty} \frac{(\frac{d}{dq}y - p)\Pr(\mathrm{piv}|\mathbf{q}_n, n)^{\frac{1}{d-1}}\bar{M}}{\frac{d}{dq}p} \tag{126}$$

Note that $F_x = -\frac{4}{(1-2x)^2} \to -4$ since $x \to 0$ as $n \to \infty$ and recall that we made the initial assumption that $\lim_{n\to\infty} y \in (0,1)$. Combining (120) and (126),

$$\lim_{n\to\infty} \frac{|\frac{d}{dq}(y-p)|}{\frac{d}{dq}p} \leq \lim_{n\to\infty} \frac{|\frac{d}{dq}(y-p)|}{\frac{d}{dq}p}\bar{M}_2 \Pr(\mathrm{piv}|\mathbf{q}_n, n)^{\frac{1}{d-1}} \tag{127}$$

for $\bar{M}_2 = \bar{M}4\lim_{n\to\infty} F_y^{-1} > 0$. Since $\Pr(\mathrm{piv}|\mathbf{q}_n, n)$, $\lim_{n\to\infty} \frac{|\frac{d}{dq}(y-p)|}{\frac{d}{dq}p} = 0$, that is, (114) holds.

Taken together, Step 1 and Step 3 imply (104), which finishes the proof of Lemma 10. ■

Proof of the uniqueness. Suppose that there are two sequences of solutions to (41), $(\tilde{q}_n(\beta))_{n\in\mathbb{N}}$ and $\hat{q}_n(\beta))_{n\in\mathbb{N}}$, with $\tilde{q}_n(\beta), \hat{q}_n(\beta) \in [\frac{1}{2}, \frac{1}{2}+2\epsilon]$. Suppose that, for any $\bar{n}$, there is $n \geq \bar{n}$, so that $q_n(\beta) \neq \hat{q}_n(\beta)$. By Lemma 8, there is n' so that $\mathbf{q}_n = (q_n(\alpha), \tilde{q}_n(\beta) \in Q_n(\Delta)$ and $\hat{\mathbf{q}}_n = (q_n(\alpha), \hat{q}_n(\beta) \in Q_n(\Delta)$ for $n \geq n'$.

Take $\mathbf{q}_n \in Q_n(\Delta)$. By the definition of $Q_n(\Delta)$, $(2q-1) > \frac{\Delta}{2}\sqrt{n}$ for $q = q_n(\beta)$. Denote $p = \Pr(\alpha|\mathrm{piv}; \mathbf{q}_n, n)$. Lemma 9 implies that $\lim_{n\to\infty} \frac{d}{dq}p = \infty$ uniformly, and by the chain rule of differentiation, $\frac{\frac{d}{dq}\Phi(p)}{\frac{d}{dq}p} = \Phi'(p)$. Note that for any $\delta > 0$, there is $\delta' > 0$ so that $\Phi'(p) > \delta'$ for all $p \in (\delta, 1-\delta)$. This means that for n sufficiently large, $q(\beta; \sigma^{\mathbf{q}_n}) - q_n(\beta)$ is strictly increasing for all $\mathbf{q}_n \in Q_n(\Delta)$ for which $p \in (\delta, 1-\delta)$. Given (94), there is $\delta > 0$ and $\hat{n}$ so that for all $n \geq \hat{n}$ and any solution to (41) must satisfy $p \in (\delta, 1-\delta)$.

Consider two differing solutions for some $n \geq \max(n', \hat{n})$. Since for both solutions, it holds $p \in (\delta, 1-\delta)$ and since $q(\beta; \sigma^{\mathbf{q}_n}) - q_n(\beta)$ is strictly increasing in $q_n(\beta)$ as long as $p \in (\delta, 1-\delta)$, it cannot be that both vote shares are zeros of the function $q(\beta; \sigma^{\mathbf{q}_n}) - q_n(\beta)$, that is, they cannot both be solutions to (41). This yields a contradiction to the initial assumption and proves that for sufficiently

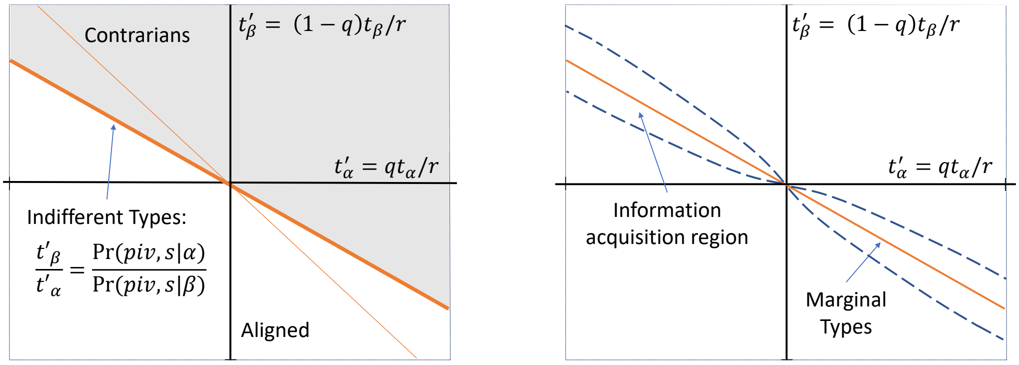

Figure 6: Voter types north-east of the indifference line (shaded area) prefer A given σ' (left panel). Types projected into the area between the dashed lines acquire information. Types projected to the outside area stay uninformed (right panel).

large n, there is a unique solution to to (41) with $\tilde{q}_n(\beta), \hat{q}_n(\beta) \in [\frac{1}{2}, \frac{1}{2} + 2\epsilon]$. The proof that any solution $q_n(\beta) \in [\frac{1}{2} - 2\epsilon, \frac{1}{2}]$ of (41) is unique, is analogous.

H Proof of Lemma 5

Figure 6 illustrates Lemma 5. In both panels (left and right), types are projected onto the two dimensions t'_α and t'_β. The left panel illustrates the voting behaviour. A voter of type t supports A if the expected value of A conditional on being pivotal and conditional on the private signal s is strictly positive. That is, if $\Pr(\text{piv}, s|\alpha; \sigma', n)t'_\alpha + \Pr(\text{piv}, s|\beta; \sigma', n)t'_\beta > 0$. These are the types with projections in the shaded area. The other types support B. The right panel illustrates the information choices. Only types with projections close to that of the indifferent marginal types (i.e., in the area between the dashed lines) choose a precision $x > 0$. Their choice is given by equating the marginal cost $c'(x) = rx^d$ with the marginal benefits MB of a higher precision. In the proof, we show that

$$\frac{\text{MB}}{r} = ||t'_\alpha \Pr(\alpha) \Pr(\text{piv}|\alpha; \sigma', n) - t'_\beta \Pr(\beta) \Pr(\text{piv}|\beta; \sigma', n)||, \tag{128}$$

compare to (17).

H.1 Proof

We show the claim for all aligned types. The argument for the contrarian types is analogous, and all partisan types vote their preference in any equilibrium.[65] In the following, we rewrite the relevant inequalities and the first-order condition from the analysis of best response in the the baseline setting in Section 3.

A voter of type t supports A if the expected value of A conditional on being pivotal and s is strictly positive and otherwise supports B. The expected value exceeds zero if

$$\begin{aligned} & \Pr(\text{piv}, s|\alpha; \sigma', n) \Pr(\alpha) t_\alpha + \Pr(\text{piv}, s|\beta; \sigma', n) \Pr(\beta) t_\beta > 0, \\ \Leftrightarrow \quad & \Pr(\text{piv}, s|\alpha; \sigma', n) \frac{q t_\alpha}{r} + \Pr(\text{piv}, s|\beta; \sigma', n) \frac{(1-q) t_\beta}{r} > 0; \end{aligned} \tag{129}$$

compare to (9) and (10). It follows from (129) that the voting behaviour under the best response is pinned down by $\frac{qt_\alpha}{r}$ and $\frac{(1-q)t_\beta}{r}$).

A voter type chooses an information precision $x = 0$ if the expected utility when choosing $x = 0$ is larger than when choosing the signal with precision $x^* = x^*(t; \sigma', n)$, where $x^*(t; \sigma', n)$ is the interior optimum characterized by equating marginal benefits of a higher precision with marginal cost,

$$c'(x^*) = \text{MB}(\sigma', n); \tag{130}$$

compare to (18) and recall (17) for the marginal benefit formula. Otherwise, he chooses the precision $x = x^*$. Rewriting (130),

$$\frac{r}{d}(x^*)^d = \Pr(\text{piv}|\alpha; \sigma', n) \Pr(\alpha) t_\alpha + \Pr(\text{piv}|\beta; \sigma', n) \Pr(\beta) t_\beta, \tag{131}$$

which shows that $\frac{qt_\alpha}{r}$ and $\frac{(1-q)t_\beta}{r}$ pin down x^*. The expected utility when choosing $x = 0$ exceeds that when choosing $x = x^*$ if

$$\begin{aligned} & \Pr(\text{piv}|\alpha; \sigma', n) \Pr(\alpha) t_\alpha + \Pr(\text{piv}|\beta; \sigma', n) \Pr(\beta) t_\beta \\ > \quad & \Pr(\text{piv}|\alpha; \sigma', n) \Pr(\alpha) t_\alpha (\frac{1}{2} + x^*) \\ & + \Pr(\text{piv}|\beta; \sigma', n) \Pr(\beta) t_\beta (\frac{1}{2} - x^*) - c(x^*), \end{aligned} \tag{132}$$

[65] Recall that we consider equilibria in undominated strategies. There are trivial equilibria where all types stay uninformed and vote for the same policy.

compare to (64). Hence, whether (132) holds, only depends on $\frac{qt_\alpha}{r}$ and $\frac{(1-q)t_\beta}{r}$ since we have shown that this pair also pins down x^*.

I Limit marginal types in the general setting

Take an informative equilibrium sequence in which the limit outcome is the same in both states. We show that the informativeness lies in between that of the other two types of equilibrium sequences, that is,

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

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

$$\lim_{n\to\infty} q(\omega; \sigma_n^*) = \frac{1}{2} \tag{133}$$

The argument is analogous to that in Section 3.2.2.

Intuitively, the election must be close in at least *some* state since otherwise the incentives to acquire costly information are too small. Formally, a voters' individual incentives to acquire information depend on the pivotal likelihood; recall e.g. the cost-benefit analysis for the optimal (interior) precision, (131). Given (27) and since the function $q(1-q)$ takes the maximum $\frac{1}{4}$ at $q = \frac{1}{2}$ only, the pivotal likelihood is exponentially small unless (133) holds. If (133) does not hold in *any* state, given (131), voters acquire exponentially little information under the best response so that the difference of the vote shares in the two states, measured in standard deviations, goes to zero, i.e. $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0$.

The reason why the election is close in all and not just in one state (i.e. (133)) is that the likelihood that a random citizen votes A is asymptotically the same across states. This is because, for any strategy sequence, the signal precision of a random voter is of an order weakly smaller than $n^{-\frac{1}{2(d-1)}}$, given (27), (131), and (27). We claim that this implies

$$q(\omega; \sigma_n^*) \to \Psi(p_n) \tag{134}$$

for both states ω and where $p_n = \frac{\Pr(\text{piv}|\alpha;\sigma_n^*,n)}{\Pr(\text{piv}|\sigma_n^*,n)}$ is the inference from the pivotal event. To see this, recall the inequality (129), divide it by $\Pr(\text{piv}|\sigma_n^*, n)$, and let s

be the uninformative signal,

$$p_n \frac{q t_\alpha}{r} + (1 - p_n) \frac{(1-q) t_\beta}{r} > 0; \tag{135}$$

The inequality describes the types that vote A under the best response given s. Since in equilibrium the precision of all types is asymptotically zero, as $n \to \infty$, the inequality describes the types that vote A in equilibrium as $n \to \infty$. Comparing the inequality with the definition (55) of Ψ shows (134).

The closeness of elections, (133) pins down the marginal types as $n \to \infty$. Since Ψ is continuous and strictly increasing, (133) together with (134), implies that $\lim_{n\to\infty} p_n = \bar{p}$ where $\bar{p}$ is so that $\Psi(\bar{p}) = \frac{1}{2}$. The marginal types are those that are indifferent between voting A and B without further information about the state. Given (135) and since $p_n \to \bar{p}$, the "limit marginal types" are the types for which $\bar{p}\frac{q t_\alpha}{r} + (1-\bar{p})\frac{(1-q)t_\beta}{r} = 0$.

J Correlation and dispersion of priors matter

We analyze the effect of dispersion in priors beliefs and the effect of correlation between priors and total intensities.

For this, we consider a setting with a symmetry assumption for the type distribution. This way, we can isolate the effect of varying either the dispersion or the correlation. All types are such that the preference intensity is constant across states, that is, $\frac{-t_\beta}{t_\alpha - t_\beta} = \frac{1}{2}$, and there is a common cost type $r = 1$. Hence, a type distribution is fully described by the following: the likelihood of a type belonging to a given interest group, $\Pr(\{t : t \in L\})$ and $\Pr(\{t : t \in C\})$ and the conditional joint distribution of total intensities $k(t)$ and prior beliefs, per interest group. Let the conditional distributions of the prior beliefs be symmetric to $\frac{1}{2}$. As in the main text, we consider situations so that the type distribution is "monotone", that is, Ψ is strictly increasing and crosses $\frac{1}{2}$.

We claim that in this symmetric setting, the κ-index $I(\kappa, g, \omega)$ (see (56)) is proportional to

$$\underbrace{f(g) f(T_{\bar{p}}|g)}_{\text{likelihood of the limit marginal types}} \qquad \mathrm{E}(|| \, k(t) \, ||^\kappa \mid g, T_{\bar{p}}, \omega), \tag{136}$$

and that $\bar{p} = \frac{1}{2}$.

First, we show that the unique belief $\bar{p} = \frac{1}{2}$ at which $\Psi(\bar{p}) = \frac{1}{2}$ is given by $\bar{p} = \frac{1}{2}$. Recalling the definition of Ψ, this means that the likelihood that a random type prefers A given the prior beliefs is $\frac{1}{2}$. The types that prefer A given the prior belief are those for which $t'_\alpha + t'_\beta \geq 0$. This is equivalent to $(2v-1)t_\alpha \geq 0$, given (54) and $r = 1$. Since all aligned types have $t_\alpha > 0$, only the aligned types with $q \geq \frac{1}{2}$ prefer A given the prior. Analogously, all contrarian types with $q \leq \frac{1}{2}$ prefer A given the prior. Since the conditional distribution of priors is symmetric for each interest group, it holds $\Psi(\frac{1}{2}) = \frac{1}{2}$.

Second, we show that $\mathrm{E}(||\, t'_\omega \,||^\kappa \mid g, T_{\bar{p}}, \omega)$ is proportional to $\mathrm{E}(||\, k(t) \,||^\kappa \mid g, T_{\bar{p}}, \omega)$. This holds since we assumed that $t_\alpha = -t_\beta$ for all types, which implies $k(t) = 2||t_\alpha||$, see (5). Together, both steps imply (136).

Dispersion and correlation of prior beliefs. Take an interest group $g \in \{L, C\}$. Fix the type distribution of the other interest group. Fix the distribution of the total intensity conditional on any prior belief v and g. Take two distributions of the prior beliefs with c.d.f. F_g and $\hat{F}_g$ respectively that are symmetric to $\frac{1}{2}$. The distribution of F_g is a mean preserving spread of that of $\hat{F}_g$ if and only if $\int_{\frac{1}{2}-\epsilon}^{\frac{1}{2}+\epsilon} F_g(v)dv \leq \int_{\frac{1}{2}-\epsilon}^{\frac{1}{2}+\epsilon} \hat{F}_g(v)dv$ for all $\epsilon > 0$.

The set $T_{\bar{p}}$ of the limit marginal types is given by the types for which $\bar{p}t'_\alpha + (1-\bar{p})t'_\beta = 0$, which is equivalent to $v = \frac{1}{2}$, given $\bar{p} = \frac{1}{2}$, $r = 1$, and (54). Hence, if F_g is a mean-preserving spread of $\hat{F}_g$, the likelihood of the limit marginal types with uniform prior is smaller given F_g. Since we fixed the distribution of the total intensity conditional on $q = \frac{1}{2}$ and g and hence the conditional expectation $\mathrm{E}(||\, k(t) \,||^\kappa \mid g, T_{\bar{p}}, \omega)$ and using (136), we see that the index of the group is smaller. Similarly, if intermediate priors are stronger correlated with high total intensities so that $\mathrm{E}(||\, k(t) \,||^\kappa \mid g, T_{\bar{p}}, \omega)$ is larger, this increases the index, ceteris paribus.

For this symmetric setting, one can show that if $d > 3$, there is only one limit equilibrium and in this limit equilibrium, the policy preferred by the interest group with the higher index is elected with probability 1. Thus, the above comparative statics yield clean predictions for policy outcomes. The reason why there is unique equilibrium, is because in this symmetric setting $\Psi(\frac{1}{2}) = \frac{1}{2}$ holds, whereas in the previous theorems we have only shown existence of three equilibria if $\Psi(\frac{1}{2}) \neq \frac{1}{2}$, see Theorem 2 and 3.

K Proof of Theorem 3

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, (40) and (41). Let $\mathrm{Q}_{\epsilon,n}$ be the set of vote share pairs $\mathbf{q}_n = (q_n(\alpha), q_n(\beta))$ satisfying

$$|q_n(\alpha) - q_n(\beta)| \leq \frac{1}{n^2}, \tag{137}$$

and

$$|q_n(\omega) - \frac{1}{2}| > \epsilon \tag{138}$$

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

$$\mathbf{q} \in \mathrm{Q}_{\epsilon,n} \Rightarrow \mathbf{q}(\sigma^{\mathbf{q}}) \in \mathrm{Q}_{\epsilon,n}. \tag{139}$$

Take a sequence of candidate equilibrium vote shares $\mathbf{q}_n \in \mathrm{Q}_{\epsilon,n}$. The first condition (137) implies that the voters do not learn anything about the state from conditioning on being pivotal, $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\mathbf{q}_n,n)} = \frac{1}{2}$. This is, because $\lim_{n\to\infty} \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\beta;\mathbf{q}_n,n)} = \lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))}$ by (38) (or alternatively by Lemma 6 in Appendix E) and because $\lim_{n\to\infty} \delta_n(\alpha) - \delta_n(\beta) = 0$ by the choice of $\mathbf{q}_n$, so that $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} = 1$.[66] The second condition (138) implies that the pivotal likelihood becomes exponentially small as $n \to \infty$, as can be seen from (27). Hence, also the precision of any voter type under the best response becomes exponentially small, given (19) and, further, the distance of the best response's vote share in α to the vote share in β, given (37). We see that the vote shares of the best response again satisfy (137) when n is large. Further, they converge to

$$\lim_{n\to\infty} q_n(\omega) = \lim_{n\to\infty} \Psi(p_n) \tag{140}$$

for $p_n = \frac{\Pr(\text{piv}|\alpha;\mathbf{q}_n,n)}{\Pr(\text{piv}|\mathbf{q}_n,n)}$, given (134). Since we have shown that $\lim_{n\to\infty} p_n = \frac{1}{2}$, since $\Psi(\frac{1}{2}) \neq \frac{1}{2}$, and since Ψ is continuous, the vote shares under the best response also satisfy (138) when n is large. We conclude that the best reponse is a self-map on the set $\mathrm{Q}_{\epsilon,n}$ of vote shares satisfying (137) and (138), when n is large.

[66]Here, recall that $\delta_n(\omega)$ is the distance of the vote share to $\frac{n}{2n+1}$ in terms of standard deviations $s(\omega;\sigma_n) = \frac{q(\omega;\sigma_n)(1-q(\omega;\sigma_n))}{\sqrt{2n+1}}$, see (24).

An application of Kakutani's fixed point theorem yields a sequence of equilibrium vote shares, and any such equilibrium sequence must satisfy (140): As we have just shown, this is a property of the best response to vote shares satisfying (137) and (138). Since any informative equilibrium sequence must, however, satisfy (133), we conclude, that the sequence of equilibrium vote shares corresponds to a non-informative equilibrium sequence.

Properties of non-informative equilibrium sequences. In the main text, we already argued that the limit behaviour in all non-informative equilibrium sequences is given by voting according to the prior beliefs (see Section 6). Hence, the vote share for policy A is $\Psi(\frac{1}{2})$ in expectation in both states. The weak law of large numbers implies that $\lim_{n\to\infty} \Pr(A|\sigma_n^*, n) = 1$ if $\Psi(\frac{1}{2}) > \frac{1}{2}$ and $\lim_{n\to\infty} \Pr(B|\sigma_n^*, n) = 1$ if $\Psi(\frac{1}{2}) < \frac{1}{2}$.

L Ordering the equilibrium sequences along the informativeness

First, since the equilibrium sequence is informative, $\lim_{n\to\infty} |\delta_n(\alpha) - \delta_n(\beta)| > 0$. Second, suppose that $\lim_{n\to\infty} |\delta_n(\alpha) - \delta_n(\beta)| = \infty$. Since the limit outcome is the same in both states, given (25), this implies that either $\delta_n(\omega) > 0$ for both states and all n sufficiently large or $\delta_n(\omega) > 0$ for both states and all n sufficiently large. Note that $\lim_{n\to\infty} \frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))} = e^{-\frac{1}{2}(\delta_n(\alpha)^2 - \delta_n(\beta)^2)} \in \{0, \infty\}$ since ϕ is the density of the standard normal. Since for any type with prior belief $q \in (0,1)$ $\lim_{n\to\infty} \frac{\Pr(\alpha|\text{piv},\sigma_n^*,n)}{\Pr(\alpha|\text{piv},\sigma_n^*,n)} = \lim_{n\to\infty} \frac{q}{1-q}\frac{\phi(\delta_n(\alpha))}{\phi(\delta_n(\beta))}$, given (38), $\lim_{n\to\infty} |\delta_n(\alpha) - \delta_n(\beta)| = \infty$ implies $\lim_{n\to\infty} \Pr(\alpha|\text{piv}; \sigma_n, n) \in \{0, 1\}$. As a consequence, all types behave as if they know the state to be α or β. The mass of types that prefers A when the state is known to be α is given by $\Psi(1)$ and the mass mass of types that prefers A when the state is known to be α is given by $\Psi(0)$, see (55). For monotone type distributions, it holds $\Psi(0) < \frac{1}{2} < \Psi(\frac{1}{2})$, so that $\lim_{n\to\infty} q(\omega; \sigma_n^*, n) \neq \frac{1}{2}$. This contradicts the observation from Section 5.2 that, for any informative equilibrium sequence, $\lim_{n\to\infty} q(\omega; \sigma_n^*, n) = \frac{1}{2}$, that is, the election is "close" as $n \to \infty$ (compare also to Section 3.2.3).