Strategic Voting in Two-Party Legislative Elections

Niall Hughes

King's Business School, King's College London

niall.hughes@kcl.ac.uk

May 2022

Abstract

I study multi-district legislative elections with two parties and two binary dimensions of policy. Strategic voters focus on the most likely pivotal dimension in the legislature. Anticipating this, candidates select different policies than they would in single-district elections. The final policy is: (i) uniquely pinned down by voter preferences, (ii) preferred by a majority of districts on each dimension, (iii) a Condorcet winner if one exists. These properties are not guaranteed in single-district elections. Furthermore, I show that (iv) parliamentary systems generate superior policies to presidential systems and (v) voter polarisation affects outcomes in single-district elections but not legislative elections.

Keywords: Strategic Voting, Legislative Elections, Multi-dimensional Policy, Pivotal Voting, Plurality Rule, Large Elections.

JEL Classification Number: C72, D72, D78.

1 Introduction

Voters often find themselves between a rock and a hard place. Consider the plight of an economically-right, socially-liberal voter - a “Rockefeller Republican” - in the US 2020 Presidential Election. Should she vote for economically-left, socially-liberal Biden or economically-right, socially-conservative Trump? Many traditional working-class voters faced a similar dilemma in the UK 2019 General Election - do they plump for the economically-left, socially-liberal Labour party, or the economically-right, socially-conservative Tory party? At first, the answer seems straightforward: hold your nose and vote for the least bad of the two. If you’re a Rockefeller Republican who cares more about the economy than social issues, vote Trump, otherwise vote Biden. Traditional working-class voters in the UK who care more about social than economic issues should back the Tories, while the others should back Labour.

This intuition is correct in a single-district election such as a mayoral election or a presidential election. Whoever wins the election implements policy, so voters are essentially voting directly over two policy bundles. Given the two options presented, a voter’s best course of action is simply to vote for the bundle they dislike least. In this paper, I show that this intuition does not carry over to legislative elections. In a legislative election, voters in a district elect a representative to the legislature, and it is there that policy is determined. Their district will only matter for the final policy outcome if it changes the legislative majority on one of the two dimensions. Here, a voter needs to ask: “when will my district be pivotal nationally? Is it more likely to swing the legislative majority on the economic dimension or the social issues dimension?”. It may be that a voter cares more about the economic dimension but, because of how other districts behave, her vote is more likely to make a difference on the social issues dimension. How should this voter behave? I show that as the size of the electorate gets large, voters should focus exclusively on the dimension where they are more likely to be pivotal. By focusing on this single dimension, voters no longer face a trade-off - strategic voting allows them to unbundle policies. This is true regardless of how much or little they care about the pivotal dimension. In a nutshell, a voter’s preference intensity is irrelevant for her voting decision.

The critical difference between single-district and legislative elections is in the behaviour of conflicted voters - those who prefer one candidate on one dimension but the other candidate on the other dimension. We might wonder if this actually has any meaningful aggregate effects. After all, if candidates are strategic and choose policies to maximise their probability of winning, we might expect their policy platforms to reflect the preferences of the vast bulk of voters. In other words, if most voters are either liberal on both dimensions or conservative on both dimensions, why should Rockefeller Republicans have any impact on candidate platforms? On the contrary, I show that the behaviour of conflicted voters has substantial effects on the platforms that candidates choose and on the final implemented policy. First, outcomes are more predictable. In legislative elections, the platforms of candidates and the final policy are uniquely pinned down by voter preferences as the size of the electorate gets large. This is not the case in single-district elections. There, candidates often play non-degenerate mixed strategies when selecting their platforms, meaning the platforms voters face and the implemented final policy is random. Second, policy outcomes are more representative. In single-district elections, party loyalists always vote along party lines, so only a subset of voters are swing voters. Candidates target these swing voters, and this may lead to platforms and final policies which do not reflect the preferences of the majority of voters. Instead, in a legislative election, every type of voter is a swing voter. Candidates can no longer ignore the preferences of party loyalists as they too may cross party lines. The knock-on effect is that, in each district, one candidate always chooses a platform which has majority support on both dimensions. When we aggregate across districts, this means that the implemented policy is the one preferred by a majority of districts on each dimension. Third, there is a welfare windfall from the strategic behaviour of conflicted voters: Condorcet winner policies are guaranteed to be implemented in legislative elections but not so in single-district elections. After the main analysis, I study applications of the model to two key political debates - which system of government produces better outcomes? And what effect does voter polarisation have on political outcomes? I show that parliamentary systems typically result in higher welfare for median voters than presidential systems. Then, I find that increased polarisation of voter preferences does not affect outcomes in legislative elections (implemented policies remain optimal) but can change outcomes (for better or worse) in single-district elections. The overall message is, in a world with multiple dimensions of policy, legislative elections outperform single-district elections.

I argue that this distinction between legislative and single-district elections matters for several reasons. First, the most important elections worldwide for determining policy are legislative elections, and many of these are district-by-district first-past-the-post elections e.g. US Congress, UK House of Commons, Canadian House of Commons, Indian Lok Sabha.[1] Therefore, if there is a difference in optimal voter behaviour between single-district and legislative elections, it may have large and widespread consequences. Second, the difference in behaviour depends on voters having preferences over more than one dimension of policy. Several studies, polls and newspaper accounts have documented the evolution of the

[1]First-past-the-post is also used to elect (lower house) legislatures in Bangladesh, Botswana, Malaysia, Myanmar, Nigeria, Pakistan, Uganda, Zambia and many former British colonies.

traditional left-right voter divide into a more complex structure where voters care about multiple dimensions and these cut across party allegiances.[2] Thus, while the difference in voting incentives between single-district and legislative elections may have been purely theoretical in the past, that is no longer the case. Third, voters of all stripes want to maximise the impact of their vote, but they may be wasting their vote or even helping an adversary without the correct analysis to base their choices on. A notable example was during the UK 2019 General Election, where several tactical voting groups offered advice to voters. Some, such as Tactical Vote and Turf Out The Tories, wanted to prevent a Conservative majority while others, such as Best for Britain and Remain United, wanted to prevent Brexit. Their advice on how to vote often conflicted as their objectives were not fully aligned, leaving potential strategic voters unsure how to vote optimally. My analysis shows that there is no conflict between such groups in a legislative election - voters (and interest groups) should focus only on the dimension where the legislative majority is most likely to change. Fourth, the distinction between single-district and legislative elections matters because they result in different policy outcomes for the same distribution of voter preferences. This is important for constitutional design as in presidential systems presidents are typically chosen via single-district elections and have a large impact on policy, while in two-party parliamentary systems, legislative elections determine who forms government and which policies are implemented. Fifth, the distinction between single-district and legislative elections will matter if they respond differently to shocks. I show that this is the case for shocks to voter preferences, i.e. increased polarisation.

There are two binary dimensions of policy in the model: a left-right dimension and another dimension which I call reform/anti-reform. We could instead think of that second dimension being Brexit/anti-Brexit, globalisation/anti-globalisation, or pro-life/pro-choice, etc. Alternatively, we could think of this second dimension as a policy which is decided at the national level but has heterogeneous local effects. Two key features of the model are that (i) both of these dimensions are binary, and (ii) candidates have a fixed position on one dimension. The model shares these features with the model of Krasa and Polborn (2010) who examine single-district elections. By keeping these same assumptions, I can easily compare my results to theirs. In Section 6, I describe how my results are robust to increasing the number of dimensions or allowing non-binary policies. In reality, many issues are binary or at least perceived to be binary by voters, e.g. whether a candidate is pro- or anti- Brexit, abortion, or gun control. Furthermore, if policies are continuous, it is well known that

[2]See e.g. Buisseret and Van Weelden (2019),Gethin *et al.* (2021), https://www.economist.com/britain/2019/01/19/the-great-rescrambling-of-britains-parties, https://www.voterstudygroup.org/publication/political-divisions-in-2016-and-beyond.

equilibria may not exist.[3] I assume candidates are fixed on the left-right dimension as, in practice, party affiliation is fixed - politicians find it very difficult to switch from one party to another or from a left-wing to a right-wing position.[4] I could allow candidates to choose their positions on both dimensions, but then they would both choose the same position, and each would win with probability 0.5 - something we don't see in reality. The model differs from that of Krasa and Polborn (2010) in that I focus on legislative elections which in turn allows voters to behave strategically.

The paper contributes to the various strands of the political economy literature. First, to the best of my knowledge, this is the first paper to combine legislative elections with multiple dimensions of policy. Most models have either examined legislative elections with a single dimension of policy (Austen-Smith, 1984; Callander, 2005; Eyster and Kittsteiner, 2007) or multiple dimensions of policy in a single district (Besley and Coate, 2008; Krasa and Polborn, 2010; Aragonès *et al.*, 2015; Xefteris, 2017; Bonomi *et al.*, 2021).[5] Legislative elections have garnered increasing attention from researchers of late but these have either focused on a single dimension with three parties (Hughes, 2016) or where one dimension voters care about is local and the other is national (Polborn and Snyder, 2017; Krasa and Polborn, 2018).

Second, it contributes to the literature on the drivers of voter behaviour, and more specifically strategic voting.[6] Here, the emphasis has typically been on multi-candidate elections (Myerson, 2002; Myatt, 2007; Patty *et al.*, 2009; Bouton and Castanheira, 2012)) or where information aggregation has a role to play (Feddersen and Pesendorfer, 1997; Krishna and Morgan, 2011). I show that even where there is full information and voters only face a choice between two candidates, there are still incentives to vote strategically, as votes can be pivotal in two different ways.[7] The type of strategic voting I describe also differs from the "conditional sincere voting" of Alesina and Rosenthal (1995, 2000) and Llavador (2006). In these papers, final policy is a combination of the majority and minority party policies, with weights given by seat shares - thus there is an incentive for centrist voters to engage in split-ticket voting to achieve policy moderation. In contrast, policy is solely determined by the majority in my model and, given that policies are binary, there is no scope for policy

[3]See Xefteris (2017) for a discussion of these issues.

[4]I use the left-right distinction as it is the most common and is used elsewhere Krasa and Polborn (2010); Buisseret and Van Weelden (2019); Buisseret and Van Weelden (forthcoming). If the main party cleavage was something else, e.g. religious/secular, as in Fourati *et al.* (2019), all the results still hold.

[5]See Osborne (1995) for a review of candidate competition in single-district and legislative elections with one dimension.

[6]Recent work on the determinants of voter behaviour include Bonomi *et al.* (2021), Gethin *et al.* (2021) and Cantoni and Pons (2022)

[7]There is a related literature where voters can either be pivotal for the electoral outcome or for the signal that is sent to politicians about voter preferences (Razin, 2003; Myatt, 2017).

moderation. The nature of strategic voting is thus very different here.

Third, the paper speaks to the vast literature on how political institutions affect policy choices (Persson, 2002; Morelli, 2004; Persson and Tabellini, 2005; Bordignon *et al.*, 2016). I show that in majoritarian systems, the type of election affects the congruence between voters and policies: legislative (parliamentary) elections generate superior policies for the median voter on each dimension than single-district (mayoral or presidential) elections. Finally, an application of the model in Section 5 adds to the literature on how political polarisation interacts with electoral institutions (Matakos *et al.* (2016); Gentzkow (2016); Canen *et al.* (2020)).

The paper proceeds as follows. I introduce the model and define an equilibrium in the Section 2. In Section 3, I solve the model. Then, in Section 4, I highlight the main results of the model and compare them to the literature. In Section 5, I use the model to generate additional results on (i) presidential versus parliamentary systems, and (ii) the impact of polarisation. Section 6 discusses the assumptions of the model, while Section 7 explores implications for the real-world.

2 Model

Candidates & Policy-Making A legislature is made up of D seats, one for each district, where D is an odd number. There is a left and right candidate in each district. Candidates must choose a platform of either supporting a social reform (Y) or opposing it (N). That is, each candidate is constrained to run under the relevant party platform, but free to be pro- or anti-reform. Candidates choose their platforms to maximise their probability of being elected in the district. Let $\mu_d = (\mu_{L,d}, \mu_{R,d})$ be the probabilities with which each candidate in district d chooses a pro-reform platform. The strategies of all candidates in all districts are summarised by $\boldsymbol{\mu} \equiv (\mu_1, \ldots, \mu_d, \ldots, \mu_D)$. Once candidates choose their platforms, each district faces one of four platform pairs: $a_d \in \{(a_{LN}, a_{RN}), (a_{LY}, a_{RY}), (a_{LN}, a_{RY}), (a_{LY}, a_{RN})\}$. In each district, an election is held, and whichever candidate wins a majority of votes is elected to the legislature. The post-election seat distribution in the legislature is given by $S = (s_{LN}, s_{LY}, s_{RN}, s_{RY})$, where each element denotes the number of seats won by a candidate with that platform. The final policy $z \in \{z_{LN}, z_{LY}, z_{RN}, z_{RY}\}$ is decided in the legislature by separate majority votes on each dimension. That is, I model "winner take all" legislatures, where policy is determined by a simple majority on each dimension.[8] Each

[8] Several other papers model the policy outcome of a legislative election as a combination of the minority and majority party's policies, with weights given by seat share (see e.g.Alesina and Rosenthal (1995, 2000); Llavador (2006)). While these setups capture some elements of multi-party legislatures, they seem less appropriate for 2-party legislatures where one party always has a majority.

legislator is committed to supporting the platform he was elected on.[9]

Voters A voter does not care who wins her district *per se*, all that matters is the final policy, z, decided in the legislature. A voter's type $t \in T \equiv \{t_{LN}, t_{NL}, t_{LY}, t_{YL}, t_{RN}, t_{NR}, t_{RY}, t_{YR}\}$ characterises not only her preferred position on each dimension but also their relative importance to her. I explicitly define the preference ordering of two types of voters, where the remaining six are similarly defined:

$$\begin{aligned} U(z_{LN}|t_{LN}) > U(z_{LY}|t_{LN}) > U(z_{RN}|t_{LN}) > U(z_{RY}|t_{LN}) \\ U(z_{LN}|t_{NL}) > U(z_{RN}|t_{NL}) > U(z_{LY}|t_{NL}) > U(z_{RY}|t_{NL}) \end{aligned} \tag{1}$$

Utility is bounded for all voter types. That is, for any two policies z, z' and any voter type t we have $\underline{U} < |U(z|t) - U(z'|t)| < \overline{U}$. It will often be easier to refer to sets of types t_L, t_R, t_N, t_Y where $t_L = \{t_{LN}, t_{NL}, t_{LY}, t_{YL}\}$, $t_R = \{t_{RN}, t_{NR}, t_{RY}, t_{YR}\}$, $t_N = \{t_{LN}, t_{NL}, t_{RN}, t_{NR}\}$, $t_Y = \{t_{LY}, t_{YL}, t_{RY}, t_{YR}\}$.

Following Myerson (2000, 2002), the number of voters in each district is not known but rather is a random variable k_d, which follows a Poisson distribution and has mean k.[10] Appendix A summarises several properties of the Poisson model.[11] The *actual* population of voters in d consists of k_d independent and identically distributed draws from a distribution f_d. The probability that a randomly drawn voter in district d is of type t is $f_d(t)$, with $\sum_t f_d(t) = 1$. The draws in d from f_d are independent of the draws in any other district d' from $f_{d'}$. To rule out knife-edge events, I impose that $f_d(t_i) \neq f_d(t_j)$ for any $i, j \in \{L, R, N, Y\}$ in each district.[12] A voter knows her own type, the distribution from which she was drawn, and the distribution functions of the other districts, $\mathbf{f} \equiv (f_1, \ldots, f_d, \ldots, f_D)$, but she does not know the actual distribution of voters that is drawn in any district. The left candidate is the *advantaged candidate* if $f_d(t_L) > 0.5$ while the right candidate is advantaged if the inequality is reversed. Candidates know $\mathbf{f}$, but not its realisation.

Let $V \equiv \{v_L, v_R\}$ be the set of actions a voter can take.[13] A voting strategy is $\sigma : T \times D \to \Delta(V)$ where $\sigma_{t,d}$ is the probability that a type t voter in district d casts ballot v_L, with $1 - \sigma_{t,d}$ being the probability that this type votes v_R. In a Poisson game, all voters of

[9] As shown by Kramer (1972), with separable preferences sincere and strategic voting lead to the same equilibrium, furthermore, the order of voting or allowing for reconsideration will not affect the outcome.

[10] The probability that there are exactly η voters in a district is $Pr[k_d = \eta] = \frac{e^{-k}k^{\eta}}{\eta!}$.

[11] The use of Poisson games in large election models is now commonplace as it simplifies the calculation of probabilities while still producing the same predictions as models with fixed but large populations. See e.g. Krishna and Morgan (2011), Bouton and Castanheira (2012), Bouton (2013).

[12] This ensures there is a clear majority of types on each dimension and these two majorities are not exactly the same.

[13] Voting is costless; thus, there will be no abstention.

the same type in the same district must follow the same strategy.[14] The expected vote share of the left candidate in the district is

$$\nu_d = \sum_{t \in T} \sigma_{t,d} f_d(t) \tag{2}$$

which can also be interpreted as the probability of a randomly selected voter playing v_L. The left candidate is the *expected winner* if $\nu_d > 0.5$ while the right candidate is if $\nu_d < 0.5$.[15] As $k \to \infty$ the probability that the expected winner wins the district goes to one.[16] Let $\boldsymbol{\sigma} \equiv (\sigma_1, \ldots, \sigma_d, \ldots, \sigma_D)$ denote the profile of voter strategies across districts and $\boldsymbol{\nu} \equiv (\nu_1, \ldots, \nu_d, \ldots, \nu_D)$ denote the profile of expected vote shares for the left candidates. Thus, we have $\boldsymbol{\nu}(\boldsymbol{\sigma}, \mathbf{f})$.

Pivotality A district is pivotal if the policy outcome z depends on which candidate that district elects. When deciding on her strategy, a voter need only consider cases in which her vote affects the policy outcome. Therefore, she will condition her vote choice on her district being pivotal.[17] There are five distinct types of pivotal event:

$$P \equiv \{piv(LR|N), piv(LR|Y), piv(NY|L), piv(NY|R), piv(LR, NY)\}$$

where $piv(LR|N)$ is when the district is pivotal between z_{LN} and z_{RN}, $piv(NY|L)$ is when the district is pivotal between z_{LN} and z_{LY}, and so on. The event $piv(LR, NY)$ is when the district is jointly pivotal on both dimensions of policy. In this case, depending on the platforms in the district, all four policy outcomes are possible. Where convenient I will simply refer to $piv(LR|N)$ and $piv(LR|Y)$ as $piv(LR)$ events and to $piv(NY|L)$ and $piv(NY|R)$ as $piv(NY)$ events. I denote the pivotal events in district d as $piv_d(LR), piv_d(NY)$ and so on.

For a given district, d, the probabilities of each of the five distinct pivotal events are functions of voter strategies $\boldsymbol{\sigma_{-d}}$, candidate strategies $\boldsymbol{\mu_{-d}}$, as well as distribution functions $\mathbf{f_{-d}}$ in the other $D-1$ districts. The *magnitude* of an event is speed at which its probability of an event goes to zero as k increases.[18] Events with greater probability have larger magni-

[14] This stems from the very nature of population uncertainty. See Myerson (1998)) pg. 377 for more detail.

[15] If $\nu_d = 0.5$, a coin toss determines the winner.

[16] This is because as the number of draws from f_d gets large, the probability that a majority of draws are from minority supporters goes to zero.

[17] I could additionally have voters condition on their vote being pivotal within their district. This would reduce the magnitude of each national pivotal event by the same constant but would leave the ranking of magnitudes unchanged. As voter best responses depend only on the pivotal event with largest magnitude, this would not alter any results.

[18] See Section 7 for a formal definition.

tudes. Let $piv_d^n(\boldsymbol{\sigma_{-d}}, \boldsymbol{\mu_{-d}}, \mathbf{f_{-d}})$ denote the n-th most likely pivotal event for district d. The pivotal event with largest magnitude is piv_d^1, the one with second largest magnitude is piv_d^2, and so on.

Assumption 1. *Candidates believe that if $piv_d^1 \neq piv(LR, NY)$, voters know the correct piv_d^1, while if $piv_d^1 = piv(LR, NY)$, either all voters in d believe $piv_d^1 = piv(LR)$ or all believe $piv_d^1 = piv(NY)$.*

That is, candidates believe that voters believe the double-pivotal event $piv(LR, NY)$ is never the most likely pivotal event. For convenience, I will assume that these candidate beliefs about voters are correct, but this is not necessary for the paper's main results.[19] Assumption 1 is important for my results. It means candidates can ignore double-pivotal events when choosing their platforms. This simplifies the candidate competition game, meaning candidates only need to consider the strategies of players within their own district, and gives us Proposition 2. This, in turn, leads to the positive results of Propositions 3-5. The assumption is not required for Lemma 1, Proposition 1, and its first corollary.

Given importance for my results, I now discuss why Assumption 1 is reasonable. First, there are very few double-pivotal events compared to single-pivotal events. For a legislature of size D there are only $\frac{D+1}{2}$ possible seat distributions which would make a given district pivotal on both dimensions while there are $\sum_{d=1}^{\frac{D-1}{2}} 4d$ single-pivotal events. The larger a legislature, the smaller the share of events that are pivotal on both dimensions. For example, in the 435-seat US House there are 218 double-pivotal events but 94,612 single-pivotal events, in the 650-seat UK House of Commons there are 325 double-pivotal events but 210,600 single-pivotal events. Hence, it is natural that candidates believe voters consider single-pivotal events to be more likely. Second, candidates know that voters get their information on likely pivotal events via imperfect election polls. As double-pivotal events are rare, polling may under-report this likelihood or only report on single-pivotal events.[20] If the latter were the case, voters would always believe a single-pivotal event was the most likely. Third, as I show in Appendix C, the chain of events necessary for a district to have $piv_d^1 = piv(LR, NY)$ is very specific. If candidates believe that, even with perfect polls, voters cannot precisely calculate and compare the magnitudes of different chains of events, then the assumption is reasonable.

There are two concerns one might have. Firstly, one might worry that voters and candidates choose their strategies assuming Assumption 1 holds but that these strategies then

[19] That is, I assume Assumption 1a in Appendix C holds. If Assumption 1 holds but voters always know the correct piv_d^1 (Assumption 1a does not hold), then Corollary 2 to Proposition 1 no longer holds: Types t_{ij} and t_{ji} vote differently if the district faces $(a_{ij'}, a_{i'j})$ and $piv_d^1 = piv(LR, NY)$.

[20] For example, if $piv_d^1 = piv(LR, NY)$, polls may report $piv_d^1 = piv(LR)$ with probability x and $piv_d^1 = piv(NY)$ with probability $1 - x$.

result in $piv(LR, NY)$ being the only feasible pivotal event. This would make players' beliefs at odds with equilibrium play. Secondly, one might worry that candidate beliefs may be inconsistent with voter behaviour when $piv_d^1 = piv(LR, NY)$. To tackle both of these concerns, I introduce Assumption 2, below. It ensures that $Pr[piv_d^1 = piv(LR, NY)] \to 0$ as $k \to \infty$. Let D_{LN} denote the number of districts with both $f_d(t_L) > 0.5$ and $f_d(t_N) > 0.5$. Let D_{LY}, D_{RN}, and D_{RY} be defined equivalently. I make the following assumption on $\mathbf{f}$.

Assumption 2. $|D_{LN} - D_{RY}| > 1$ *and* $|D_{LY} - D_{RN}| > 1$.

Lemma 2 in Appendix C shows that if Assumption 1 holds and $\mathbf{f}$ is consistent with Assumption 2, then $Pr[piv_d^1 = piv(LR, NY)] \to 0$ as $k \to \infty$. The beliefs of candidates when $piv_d^1 = piv(LR, NY)$ - whether reasonable or not - are never tested in a large election. Why is Assumption 1 needed if Assumption 2 rules out the cases it applies to? Because candidates need to know how voters would behave *if* $piv_d^1 = piv(LR, NY)$, in order to pin down their own best responses.

Payoffs We can now turn to voter payoffs. Let $G_{t,d}(v_L|k\boldsymbol{\nu}, \boldsymbol{a})$ denote the expected gain for a voter type t in district d of voting for the left candidate rather than the right candidate, given the vector of platforms $\boldsymbol{a}$ and expected vote shares $\boldsymbol{\nu}$. The expected gain of voting v_L when she is in a (a_{LN}, a_{RN}) or (a_{LY}, a_{RY}) district is given by

$$\begin{aligned} G_{t,d}(v_L|(a_{LN}, a_{RN}), k\boldsymbol{\nu}, \mathbf{a_{-d}}) &= \\ G_{t,d}(v_L|(a_{LY}, a_{RY}), k\boldsymbol{\nu}, \mathbf{a_{-d}}) &= Pr[piv_d(LR|N)](U(z_{LN}|t) - U(z_{RN}|t)) \\ &+ Pr[piv_d(LR|Y)](U(z_{LY}|t) - U(z_{RY}|t)) \end{aligned} \tag{3}$$

Here, the only way a vote can be pivotal is by switching the legislative majority from L to R. The set of pivotal events in a (a_{LY}, a_{RN}) district also includes those cases where the reform is passed if the left candidate wins but not if the right candidate wins.

$$\begin{aligned} G_{t,d}(v_L|(a_{LY}, a_{RN}), k\boldsymbol{\nu}, \mathbf{a_{-d}}) = Pr[piv_d(LR|N)](U(z_{LN}|t) - U(z_{RN}|t)) \\ +Pr[piv_d(LR|Y)](U(z_{LY}|t) - U(z_{RY}|t)) \\ +Pr[piv_d(NY|L)](U(z_{LY}|t) - U(z_{LN}|t)) \\ +Pr[piv_d(NY|R)](U(z_{RY}|t) - U(z_{RN}|t)) \\ +Pr[piv_d(LR, NY)](U(z_{LY}|t) - U(z_{RN}|t)) \end{aligned} \tag{4}$$

Similarly, the set of pivotal events in a (a_{LN}, a_{RY}) district includes those cases where the reform is not passed if the left candidate wins but is if the right candidate wins.

$$\begin{aligned}
G_{t,d}(v_L|(a_{LN}, a_{RY}), k\boldsymbol{\nu}, \mathbf{a_{-d}}) = & Pr[piv_d(LR|N)](U(z_{LN}|t) - U(z_{RN}|t)) \\
& + Pr[piv_d(LR|Y)](U(z_{LY}|t) - U(z_{RY}|t)) \\
& + Pr[piv_d(NY|L)](U(z_{LN}|t) - U(z_{LY}|t)) \\
& + Pr[piv_d(NY|R)](U(z_{RN}|t) - U(z_{RY}|t)) \\
& + Pr[piv_d(LR, NY)](U(z_{LN}|t) - U(z_{RY}|t))
\end{aligned} \tag{5}$$

A voter's best response is to vote v_L if the gain function is positive, v_R if negative. A voter is *unconflicted* if all the utility differential terms in her gain function have the same sign. A voter is *conflicted* is she has both positive and negative utility differential terms in her gain function. All voter types in (a_{LN}, a_{RN}) and (a_{LY}, a_{RY}) districts are unconflicted voters. In (a_{LY}, a_{RN}) districts, types $t_{LY}, t_{YL}, t_{RN}, t_{NR}$ are unconflicted while types $t_{LN}, t_{NL}, t_{RY}, t_{YR}$ are conflicted. In (a_{LN}, a_{RY}) districts, types $t_{LN}, t_{NL}, t_{RY}, t_{YR}$ are unconflicted while types $t_{LY}, t_{YL}, t_{RN}, t_{NR}$ are conflicted.

Timing The timing of the game is as follows:

1. Nature draws a population of k_d voters from each f_d
2. Candidates simultaneously choose their platforms, giving a_d in each district.
3. In each district, voters elect a single legislator by plurality rule, giving a seat distribution S.
4. Final policy z is chosen in the legislature by two separate majority votes.

Equilibrium Concept The mapping from seat shares into policy is mechanical, so I omit it from the definition of equilibrium. Equilibrium in this game consists of a candidate equilibrium in stage 2, $\boldsymbol{\mu}^*$, and a voting equilibrium in stage 3, $\boldsymbol{\sigma}^*$. In a voting equilibrium, voter types best respond to the strategies of all other voters in all districts. In a candidate equilibrium, candidates best respond to the strategies of all other candidates and voter types.[21] At both stages, I restrict attention to strictly perfect equilibria.[22] A candidate

[21]In fact, I show in Proposition 2 that, in large elections, candidates need only best respond to the strategies of voters and candidates within their own district.

[22]The original formulation of strictly perfect equilibrium is due to Okada (1981). Bouton and Gratton (2015) extend the definition of strictly perfect equilibrium to Poisson games.

strategy profile $\boldsymbol{\mu}^*$ is a strictly perfect equilibrium if, for any slightly perturbed version of the game where players must play completely mixed strategies, there exists a completely mixed strategy equilibrium which converges to $\boldsymbol{\mu}^*$ as the perturbations go to zero.[23] This rules out cases where candidates play weakly dominated strategies. Strictly perfect equilibrium is commonly used in Poisson voting games (Bouton and Gratton, 2015; Hughes, 2016; Durand *et al.*, 2021) as it rules out knife edge cases which may not be ruled out by weaker concepts such as perfectness or properness (De Sinopoli and Pimienta, 2009; De Sinopoli *et al.*, 2014). As shown by Bouton and Gratton (2015), in a Poisson game, a necessary and sufficient condition for a strategy profile $\boldsymbol{\sigma}^*$ to be a strictly perfect equilibrium is that it remains a best response to any small tremble change in the expected vote shares $\boldsymbol{\nu}$.[24] Given that $\boldsymbol{\nu}$ is a function of $\boldsymbol{\sigma}$ and $\mathbf{f}$, we can think of the equilibrium as being robust to trembles in players' strategies or to perturbations the distribution of voter preferences. As the paper's focus is large national elections, I analyse the properties of equilibria as the number of voters grows large. At the voting stage, I characterise equilibrium behaviour when the expected number of voters is greater than a threshold $\bar{k}$ while at the candidate competition stage, I characterise equilibrium behaviour for $k > \bar{k}$ as $k \to \infty$.

3 Equilibrium

Policymaking in the Legislature Given the legislative voting rule, there is a direct mapping from seat shares to policy outcomes: if $s_{LN} + s_{LY} > s_{RN} + s_{RY}$ and $s_{LN} + s_{RN} > s_{LY} + s_{RY}$, the final policy is z_{LN} while if $s_{LN} + s_{LY} > s_{RN} + s_{RY}$ and $s_{LN} + s_{RN} < s_{LY} + s_{RY}$, the final policy is z_{LY}. Reversing the first inequality in each case gives us policies of z_{RN} and z_{RY} respectively.

Voting A voter's decision will depend on the platform pair in her district but may, in principle, also depend on platforms and voter strategies in the other $D-1$ districts. The following lemma shows this is the case for conflicted voters.

[23]More formally, a perturbation of the game Γ is a vector $\alpha = (\alpha_{L,1}, \alpha_{R,1}, \ldots, \alpha_{L,2D}, \alpha_{R,2D})$ such that we have $\alpha_{i,d} = (\alpha_{i,d}(N), \alpha_{i,d}(Y))$ satisfies $\alpha_{i,d}(N), \alpha_{i,d}(Y) > 0$ and $\alpha_{i,d}(N) + \alpha_{i,d}(Y) < 1$ for $i \in \{L, R\}$ and $d \in D$. An ϵ-perturbed game of Γ, denoted $\Gamma(\epsilon)$ restricts the strategy set available to each player to $\mu_{i,d}(\epsilon) = \{\mu_{i,d} \, | \alpha_{i,d}(Y) < \mu_{i,d} < 1 - \alpha_{i,d}(N)\}$ for some $\alpha_{i,d}$ such that $0 < \alpha_{i,d}(N), \alpha_{i,d}(Y) \leq \epsilon$. A Nash equilibrium $\boldsymbol{\mu^*}$ is a strictly perfect equilibrium of Γ if for any arbitrary sequence of perturbations α^η such that $0 < \alpha^\eta_{i,d} < \epsilon^\eta$ and $\epsilon^\eta \to 0$ as $\eta \to \infty$, there exists a sequence of totally mixed strategy Nash equilibria $\boldsymbol{\mu}^\eta$ of the game $\Gamma(\epsilon^\eta)$ such that $\boldsymbol{\mu}^\eta \to \boldsymbol{\mu}^*$ as $\eta \to \infty$.

[24]More formally, $\boldsymbol{\sigma}^*$ is a strictly perfect equilibrium if there exists $\epsilon > 0$ such that if $\forall \tilde{\nu_d} \in \Delta V$: $|\tilde{\nu_d} - \nu_d(\boldsymbol{\sigma}^*, \mathbf{f})| < \epsilon$, then $\sigma^*_{t,d} \in BR_{t,d}(k\tilde{\boldsymbol{\nu}})$ for all $t \in T$. See Appendix C in Bouton and Gratton (2015) for more details on strictly perfect equilibria in Poisson games.

Lemma 1. *Conflicted voters are strategic voters - their best response depends on the strategies of voters and candidates in the other $D-1$ districts. Unconflicted voters have a dominant strategy for any given $\boldsymbol{a}$.*

Proof. See Appendix D. □

Whether a voter's gain function is positive or negative will depend on (i) the utility difference between each pair of policies and on (ii) the probability of her district being pivotal between each pair of policies. Her type determines the former while the latter depends on the strategies of voters and candidates in the other $D-1$ districts. An unconflicted voter can ignore (ii) because each utility differential in her gain function has the same sign. A conflicted voter has both positive and negative utility differentials, so the sign of her gain function may depend on the strategies of voters and candidates in the other districts. One can easily construct strategy profiles such that all the probability falls on a given pivotal event. As such, the gain function of conflicted voters will be positive for some strategy profiles and negative for others. Therefore, conflicted voters are *strategic voters* - their best response depends on the strategies of other players.

I have already established that a conflicted voter will either vote v_L or v_R depending on the various elements of her gain function. I now show that as the number of voters gets large and the probability of each pivotal event goes to zero, she need only consider the pivotal event with largest magnitude; that is, the most likely pivotal event. A property of the Poisson model is that, as the population size grows large, almost all the probability of being pivotal is contained in the most likely pivotal event piv_d^1. This greatly simplifies the best response of a conflicted voter.

Proposition 1. *There exists a $\bar{k}$ such that for $k > \bar{k}$, the best response for a voter is to vote solely based on the unique most likely pivotal event, piv_d^1. These best responses are represented in Table 1.*

Proof. See Appendix D. □

The proposition follows from a straightforward application of Lemma 1 in Hughes (2016), reproduced in Appendix A. If the expected seat share is such that a voter's district is pivotal in expectation, she votes based on that pivotal event. If her district is not pivotal in expectation, then, in theory, she would need to calculate the probability of all the various pivotal events. These various pivot probabilities go to zero as the number of voters increases, but they do so at different rates. That is, they have different magnitudes. If one event A has a larger magnitude than another event A' then $\lim_{k\to\infty} \frac{Pr[A']}{Pr[A]} = 0$. Therefore, conditional on being pivotal, the probability of being in the pivotal event with largest magnitude goes to

District	Vote v_L	Vote v_R
(a_{LN}, a_{RN}) (a_{LY}, a_{RY}) (a_{LN}, a_{RY}) with $piv_d^1 = piv(LR)$ (a_{LY}, a_{RN}) with $piv_d^1 = piv(LR)$	$t_{NL}, t_{LN}, t_{YL}, t_{LY}$	$t_{NR}, t_{RN}, t_{YR}, t_{RY}$
(a_{LN}, a_{RY}) with $piv_d^1 = piv(NY)$	$t_{NL}, t_{LN}, t_{NR}, t_{RN}$	$t_{YL}, t_{LY}, t_{YR}, t_{RY}$
(a_{LY}, a_{RN}) with $piv_d^1 = piv(NY)$	$t_{YL}, t_{LY}, t_{YR}, t_{RY}$	$t_{NL}, t_{LN}, t_{NR}, t_{RN}$

Table 1: Best responses of each voter type

one as the size of the electorate grows. Because of this, once the population is large enough ($k > \bar{k}$), voters can safely ignore all but the most likely pivotal event when making their voting decision as it has the largest magnitude. The result holds as long as there does not exist more than one pivotal event with the largest magnitude. I show in the Appendix in Lemma 3 that in any strictly perfect equilibrium, the pivotal event with largest magnitude is unique. While the proposition applies to all voter types, its power is in pinning down the best response of conflicted voters. A number of additional results follow directly from Proposition 1. The following corollary shows that every voter type is a potential swing voter.

Corollary 1 to Proposition 1. *Depending on the realisation of platforms,* $\mathbf{a}$*, and the strategies of other voters,* $\boldsymbol{\sigma}$*, each voter type may either vote* v_L *or* v_R*.*

This result can be seen by looking at Table 1. The final two rows show that voting v_L is a best response for each voter type in one case but not in the other. If a district faces (a_{LN}, a_{RY}) and the largest pivotal event is $piv(NY)$, then all t_Y types vote v_R while all t_N types vote v_L. If, instead, a district faces (a_{LY}, a_{RN}) and $piv_d^1 = piv(NY)$, then all t_Y types vote v_L while all t_N types vote v_R. The reason is that if the left-right dimension is not the relevant one, voters choose their preferred candidate on the other dimension - regardless of that candidate's position on the left-right dimension.

Corollary 2 to Proposition 1. *For* $k > \bar{k}$*, voter preference intensity is irrelevant - types* t_{ij} *and* t_{ji} *always vote the same way.*

This result can be seen directly in Table 1. Preference intensity is irrelevant because, for any $k > \bar{k}$, voters only consider one dimension when voting - the dimension most likely to change the legislative majority. With only one dimension to focus on, the interests of t_{ij} and t_{ji} types are aligned.

We might worry that by ignoring all but one dimension of policy, voters may select sub-optimal policies, or that candidates may have incentives to choose winning policies on one dimension but arbitrary ones on other dimensions. We will see that this is not the case in equilibrium. In fact, it is precisely by focusing on a single dimension that voters can guarantee Condorcet winning policies.

Candidate Platform Choice Candidates combine information on voter best responses from Table 1 with the distribution of voter types, $\mathbf{f}$, to work out the expected vote share in their district for any of the six possible scenarios. These expected vote shares pin down the probability of each candidate winning the district. From now on, I abuse notation slightly by denoting $(\mu_{L,d}, \mu_{R,d})$ simply as (μ_L, μ_R) where it is clear I am analysing district d.

Invoking Assumption 1, let λ be the probability that candidates put on voters believing $piv_d^1 = piv(NY)$ when in fact $piv_d^1 = piv(LR, NY)$.[25] Furthermore, let $\tilde{Pr}[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] = Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] + \lambda Pr[piv_d^1 = piv(LR, NY)|\boldsymbol{\mu_{-d}}]$ and $\tilde{Pr}[piv_d^1 = piv(LR)|\boldsymbol{\mu_{-d}}] = Pr[piv_d^1 = piv(LR)|\boldsymbol{\mu_{-d}}] + (1-\lambda)Pr[piv_d^1 = piv(LR, NY)|\boldsymbol{\mu_{-d}}]$. The probability of the left candidate winning in district d is given by

$$\begin{aligned} Pr[Lwin] = Pr[Lwin|piv_d^1 = piv(NY), \mu_L, \mu_R] * \tilde{Pr}[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] \\ + Pr[Lwin|piv_d^1 = piv(LR)] * \tilde{Pr}[piv_d^1 = piv(LR)|\boldsymbol{\mu_{-d}}] \end{aligned} \tag{6}$$

where

$$\begin{aligned} Pr[Lwin|piv_d^1 = piv(NY), \mu_L, \mu_R] &= (\mu_L\mu_R + (1-\mu_L)(1-\mu_R))Pr[Lwin|(a_{LN}, a_{RN})] \\ &+ \mu_L(1-\mu_R)Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] \\ &+ (1-\mu_L)(\mu_R)(1 - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]) \end{aligned} \tag{7}$$

From Equation 6 we see that μ_d only affects the winner of district d through Equation 7. This is because (i) candidates in d cannot affect the ranking of most likely pivotal events for their district - this only depends on the other $D-1$ districts, and (ii) voters focus on the left-right dimension when $piv_d^1 = piv(LR)$, so the probability of L winning is independent of μ_d. Taken together, this means that maximising or minimising Equation 6 is equivalent to doing so for Equation 7. Notice that Equation 7 does not include the probabilities of the different pivotal events. As such, I can fix the behaviour of candidates in the other $D-1$

[25] This λ can be arbitrary and can vary candidate to candidate.

districts and simply analyse the best responses of those in d. The following proposition characterises all of the equilibria in this candidate competition stage.

Proposition 2. *For any given* $\mathbf{f}$*, the candidate equilibrium is unique for any* $k > \bar{k}$*. In each district, both candidates choose anti-reform policies if* $f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$ *but pro-reform policies if* $f_d(t_N) < f_d(t_L), 1 - f_d(t_L)$*. For all other local preference distributions, equilibria* (μ_L^*, μ_R^*) *are in mixed strategies. As* $k \to \infty$ *in every mixed strategy equilibrium the advantaged candidate chooses the majority-preferred reform position with probability* $\to 1$ *while the disadvantaged candidate chooses the other reform position with probability* $\to 1$*. These equilibria are represented in Table 2.*

Proof. See Appendix D. □

District Preferences	μ_L^*	μ_R^*	a_d^*
$f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$	0	0	(a_{LN}, a_{RN})
$f_d(t_L), 1 - f_d(t_L) > f_d(t_N)$	1	1	(a_{LY}, a_{RY})
$f_d(t_L) > f_d(t_N) > 0.5$ or $0.5 > f_d(t_N) > f_d(t_L)$	$\tilde{\mu}_L \to 0$	$\tilde{\mu}_R \to 1$	$\to (a_{LN}, a_{RY})$
$f_d(t_L) > 1 - f_d(t_N) > 0.5$ or $0.5 > 1 - f_d(t_N) > f_d(t_L)$	$\tilde{\mu}_L \to 1$	$\tilde{\mu}_R \to 0$	$\to (a_{LY}, a_{RN})$

Table 2: Equilibrium platforms in the candidate competition game. The final two rows are mixed strategy equilibria that converge to degenerate probabilities as $k \to \infty$.

Table 2 shows that the candidate equilibrium is completely pinned down by voter preferences and falls into one of four cases. If the anti-reform majority is larger than the left or right majority, both candidates run on an anti-reform platform. Similarly, if the pro-reform majority is larger than either of the ideological majorities, both candidates choose pro-reform platforms. In each case, as both candidates choose the same reform platform, voters base their decisions on the left-right dimension. If one candidate were to deviate to the minority position on the reform dimension, their expected vote share would decrease.

In all other cases - where the left-right majority is larger than the social reform majority - the equilibrium is always in mixed strategies. The game here is similar to a matching pennies game. The advantaged candidate wants to match the policy of the disadvantaged candidate so that the left-right dimension is the one voters consider. The disadvantaged candidate wants to avoid choosing the same policy - he would ideally choose the majority policy on the social reform dimension while the advantaged candidate chooses the minority position. Clearly, these pure strategies cannot be an equilibrium, as the advantaged candidate would always seek to match the policy of the disadvantaged candidate.

Given that the equilibria are in mixed strategies, we might be concerned that many different realisations of platforms may occur for a given $\mathbf{f}$. If this were the case, it would hinder our understanding of how voter preferences map into platform choices. It turns out, however, that as the number of voters increases to infinity, the mixed strategy equilibria converge to pure strategy equilibria in which the advantaged candidate chooses the majority reform position and the disadvantaged candidate selects the minority reform position. Why does this happen? If the advantaged candidate chooses the majority reform position, his expected vote share is greater than a half for any strategy of his opponent. If he selects the minority position, he will only have a higher vote share if the opponent also chooses the same policy. The second option may give a higher overall probability of winning, but with a small risk of losing. In contrast, the safer first option has a tiny risk of losing. As the number of voters increases, the trade-off goes in favour of the less risky strategy. But why does the disadvantaged candidate choose the minority-preferred policy with probability going to one? He does so because if he chose the same platform as the advantaged candidate, voters would vote based on the left-right dimension, and he would lose with even higher probability.

The uniqueness of equilibrium was not obvious ex-ante. There could, potentially, have been multiple equilibria for the same preference distributions, e.g. where the candidates adopt a certain strategy when voters condition on one dimension but a different strategy when voter focus on another dimension. The fact that candidate strategies are completely determined by local voter preferences means that candidates can ignore the preferences and strategies of voters in other districts. This dramatically reduces the complexity of finding equilibria of the larger game.

Both Proposition 1 and Proposition 2 hold for "large" electorates. But exactly how large is large? The population threshold, $\bar{k}$, depends on (i) the bounds on voter utility differentials and (ii) the difference in the magnitudes of the pivotal events on each dimension. It becomes larger when the difference in magnitudes is closer to zero and when $\frac{\overline{U}}{\underline{U}}$ is larger. For example, if we have $\overline{U} = 1000\underline{U}$ and $-|mag[piv_d(LR)] - mag[piv_d(NY)]| = -0.1$, we get $\bar{k} = 80$. Instead, $\overline{U} = 10000\underline{U}$ gives $\bar{k} = 103$; while reducing the magnitude differential to -0.01 gives $\bar{k} = 800$.[26] In Proposition 2, the mixed strategies of candidates μ_L, μ_R depend on the size of the population and competitiveness in the district. The greater the difference in competitiveness between the two dimensions in a district, the further the equilibrium mixed strategies will be from 0 and 1. Take an extreme case of $f_d(t_L) = 0.99$ and $f_d(t_N) = 0.49$, so that the race is a foregone conclusion on the left-right dimension but very close on the

[26]These calculations make use of Equation 17 and Equation 23. Recall that $-|mag[piv_d(LR)] - mag[piv_d(NY)]|$ can range from $-D$ to 0 depending on $\boldsymbol{\nu}$. The small values of -0.1 and -0.01 in the examples correspond to cases where a district is *almost* equally likely to be pivotal on each dimension.

reform dimension. If we have $k = 100$, this gives $(\mu_L^*, \mu_R^*) \approx (0.997, 0.003)$. Instead, with $k = 1000$, we have $(\mu_L^*, \mu_R^*) \approx (0.99968, 0.00032)$.[27] Given that the average district size is above 7,000 in the UK and above 70,000 in the US, we should not be too concerned about whether $\bar{k}$ is large enough for Proposition 1 or whether the mixed strategies in Proposition 2 are well approximated by pure strategies.

District Winners & Policy Outcomes Because the mixed strategies that candidates use converge to pure strategies as $k \to \infty$, the realised vector of district platforms, $\mathbf{a}$, is unique and pinned down by voter preferences as $k \to \infty$. This does not mean that the voting equilibrium is unique, but, as I will now show, it does tell us that in any equilibrium both the winner in each district and the implemented policy are unique. The following result follows directly from Proposition 2.

Proposition 3. *For $k > \bar{k}$, the advantaged candidate in a district wins with probability $\to 1$ as $k \to \infty$.*

Proof. See Appendix D. □

In the top two cases of Table 2, the dis-advantaged candidate can never be the expected winner. Both candidates choose the majority-preferred policy on the reform dimension, meaning voters vote based on the left-right dimension and the advantaged candidate is the expected winner. In the bottom two cases of Table 2, the disadvantaged candidate is the expected winner only if he chooses the majority-preferred reform position while the advantaged candidate chooses the minority-preferred position. From Proposition 2, we know the probability that this occurs goes to zero as $k \to \infty$. At the same time, the probability that the expected winner wins the district goes to one as $k \to \infty$. Therefore, for large electorates, the winner in district d will almost certainly be the advantaged candidate with a platform which is majority-preferred on both dimensions.

I now build on this result to show that the policy implemented in the legislature is unique, and this policy is preferred by a majority of voters in a majority of districts to the alternative on each dimension. To do this, a small detour is first required. In the standard one-dimensional Downsian model, the median voter plays a key role - candidates propose policies to attract the median voter and by doing so, ensure the implemented policy has the support of a majority of voters. In a multi-dimensional setting there is no single median voter. Instead, it is useful to think of dimension-by-dimension median voters. Let $\tilde{t}_d^{LR} \in \{t_L, t_R\}$ denote the expected median voter in district d on the left-right dimension. This voter prefers left to right if $f_d(t_L) > 0.5$. Similarly, let $\tilde{t}_d^{NY} \in \{t_N, t_Y\}$ denote the expected median voter

[27] These calculations make use of Equation 13 and Equation 26.

in district d on the reform dimension. This voter is anti-reform if $f_d(t_N) > 0.5$. Order the median voters across districts so that $\tilde{t}_d^{LR}$ with lower d prefer left to right, and $\tilde{t}_d^{NY}$ with lower d oppose the reform. Then, $\tilde{t}_{\frac{D+1}{2}}^{LR}$ is the expected median voter in the median district on the left-right dimension. Similarly, $\tilde{t}_{\frac{D+1}{2}}^{NY}$ is the expected median voter in the median district on the reform dimension. Note, the median district on one dimension need not be the median district on the other dimension. By the law of large numbers, as $k \to \infty$, the realised median voters in each district (t_d^{LR}, t_d^{NY}) converge to the expected median voters $(\tilde{t}_d^{LR}, \tilde{t}_d^{NY})$, and therefore $(t_{\frac{D+1}{2}}^{LR}, t_{\frac{D+1}{2}}^{NY})$ converges to $(\tilde{t}_{\frac{D+1}{2}}^{LR}, \tilde{t}_{\frac{D+1}{2}}^{NY})$.

Proposition 4. *For $k > \bar{k}$, the policy preferred by $t_{\frac{D+1}{2}}^{LR}$ on the left-right dimension and by $t_{\frac{D+1}{2}}^{NY}$ on the reform dimension is implemented with probability $\to 1$ as $k \to \infty$.*

Proof. See Appendix D. □

The logic of the proposition is quite straightforward: we know that the implemented policy, z, requires a majority of seats on each dimension. We also know from Proposition 3 that as $k \to \infty$ the winning policy in each district is majority-preferred on both dimensions with probability $\to 1$. Combining these two gives us the result. The proposition confirms that the disciplining effect of strategic voters on candidates extends beyond local outcomes. As the number of voters gets large, candidates' mixed strategies converge to pure strategies and we get a unique implemented policy. Moreover, a majority of voters in a majority of districts prefer this unique policy to the alternative on each dimension. In other words, the result is *as if* districts held separate elections on each dimension of policy.

I can go further in exploring properties of equilibrium, but, must again define some additional terms. A Condorcet winner is a policy that would win a pairwise majority vote in a district against any other policy. This definition is not useful in a legislative setting as voters care about national policy outcomes, not local ones. The following definition extends the pairwise comparison element of a Condorcet winner to a legislative setting.

Definition. *A **Legislative Condorcet winner** is a policy z such that if it faced any other policy z' in every district, z would win a majority of seats.*

Though the notions of district Condorcet winner and Legislative Condorcet winner are related, it is worth pointing out how they differ. In a given district, z_{LN} is a Condorcet winner if $f_d(t_L) > 0.5$, $f_d(t_N) > 0.5$, and $f_d(\hat{t}_{LY}) > 0.5$ all hold, where $\hat{t}_{LY} \equiv t_{LY} \cup t_{LN} \cup t_{NL} \cup t_{NR}$. A Legislative Condorcet winner requires the three conditions to hold in a majority of districts, though not necessarily simultaneously. That is, to be a Legislative Condorcet winner z_{LN} does not need to be a Condorcet winner in a majority of districts. I can now state the following result:

Proposition 5. *For $k > \bar{k}$, if a Condorcet winner policy exists in a district, it wins with probability $\to 1$ as $k \to \infty$. If a Legislative Condorcet winner exists, it is the implemented policy z with probability $\to 1$ as $k \to \infty$.*

Proof. See Appendix D. □

The first part of the proposition states that if a Condorcet winner policy exists in a district, it will be chosen almost surely in a large electorate. This may seem like the bare minimum one would require from an electoral rule, but - as I show in the next section - it is generally not easy to guarantee. The second part of the proposition states that if a Legislative Condorcet winner policy exists, it will be implemented. This result follows the same logic as Proposition 4: the implemented policy must be preferred by a majority of voters in a majority of districts to the alternative on each dimension. By definition, a Legislative Condorcet winner policy is preferred by a majority of voters in a majority of districts. Therefore, if such a policy exists, it is sure to be implemented.

4 Analysis of Main Results

In this section, I highlight how the main results in Section 3 differ from those in the literature. I focus on three areas where the model's results are markedly different from those in other settings.

Strategic Voting The model disproves the conventional wisdom that strategic and sincere voting coincide when there are only two voting choices and there is no room for policy compromise.[28] Indeed, in single-district elections with multiple dimensions of policy, voting cannot be strategic (Krasa and Polborn, 2010; Besley and Coate, 2008). Moreover, voting is also non-strategic in legislative elections with two parties and one dimension of policy (Hinich and Ordeshook, 1974). One might reasonably conjecture that this would remain true when combining legislative elections and multiple dimensions of policy. Instead, Lemma 1 shows this is not the case: simply voting for whichever of the two candidates' platforms they prefer is not a best response for conflicted voters.

Suppose policy is not simply determined by the winner but is a combination of the majority and minority party platforms. In that case, there can be incentives for voters to behave strategically, either by split-ticket voting in presidential and legislative elections (Alesina

[28]There is a literature on strategic voting with two candidates (Feddersen and Pesendorfer, 1997) when voters have incomplete information over their own preferences. Here, voters are fully informed of their preferences over policies yet still vote strategically.

and Rosenthal, 1995, 2000) or by voting against their preferred party in order to moderate the final policy (Llavador, 2006). The "conditionally sincere" voting of those papers differs from strategic voting in this paper in a several ways. First, it relies on elections not being "winner-take-all" affairs, so that policy is continuous in vote share. This assumption seems at odds with the reality of majoritarian elections. Second, "conditionally sincere" voters vote for their preferred candidate (conditional on others' vote choices), while voters in my model are strategic in the more traditional sense: they weigh up the most likely event in which their vote will be pivotal and vote based on that dimension. Finally, preference intensity is a driver of strategic behaviour for "conditionally sincere" voting - centrist voters want to ensure moderate, compromised policies; in my model, preference intensity is irrelevant - all voters are strategic.

Strategic voting in my setting also has several features which distinguish it from strategic voting in multi-candidate elections (Myerson and Weber, 1993; Patty *et al.*, 2009; Bouton and Castanheira, 2012; Fisher and Myatt, 2017) and in information aggregation settings (Feddersen and Pesendorfer, 1997; Bhattacharya, 2013; McMurray, 2013). First, the probability of a voter's own ballot being pivotal is irrelevant in my setting. In addition, the strategies of other voters in her district do not matter at all for how she votes. Instead, what matters is the relative probability that her *district* is pivotal in the changing legislative majority on one dimension rather than the other. Proposition 1 shows this is not overly complex to calculate. She must simply focus on the most likely case in which her district is pivotal and base her decision on that dimension of policy alone. In that sense, the information she needs to vote strategically is relatively low and much less than in other models of strategic voting. For example, in multi-candidate elections, a strategic voter needs to know the strategies and distribution of voters in her district. Her vote choice will depend on which equilibrium is being played, i.e. who the two front-runners are. In information aggregation settings, a voter needs to make inferences about the likely state of the world conditional on her vote being pivotal. The cognitive burden of these types of strategic voting seems much larger than in the legislative setting.

In multi-candidate models, there is a unique equilibrium with naive voting but multiple equilibria under strategic voting. Which equilibrium is played depends purely on voter beliefs. As such, there is no guarantee that strategic voting generates "good" equilibria. On the contrary, anything goes![29] In my setting, we get the reverse: strategic voting leads to a unique policy outcome, while this is not guaranteed with naive voters.

[29]See Bol and Verthé (2019) for an overview of strategic versus sincere voting, and Eggers and Vivyan (2020) for an analysis of how strategic voting varies by demographic characteristics.

District Behaviour in Legislative v Single-District Elections In Appendix B, I solve the single-district version of the model, originally due to Krasa and Polborn (2010). Fixing the distribution of preferences in a district, f_d, I now compare a single-district election to the case where the district votes in a legislative election. Several differences immediately stand out. First, voting is strategic in legislative elections but deterministic in single-district elections. Second, all voters are potential swing voters in a legislative election. This is in sharp contrast with the case of a single-district election studied by Krasa and Polborn (2010). There, a type who cares more about the left-right dimension always votes for the same candidate, while only a type who cares more about the reform dimension may switch allegiance. The reason for the difference is that voters can never be pivotal on the reform policy alone in a single-district election. Third, from Corollary 2 to Proposition 1 we know that preference intensity is irrelevant - types t_{ij} and t_{ji} always vote the same way as voters do not need to weigh up their relative preference across multiple dimensions. Once again, this result is in sharp contrast with the analysis of single-district elections. There, conflicted types t_{ij} and t_{ji} never vote the same way - each votes for their preferred candidate on the dimension they care more intensely about.

The difference in voter behaviour between single-district and legislative elections also generates differences in candidate behaviour. Comparing Table 2 to Table 4, we see that there are many more types of equilibria in the single-district case. Furthermore, the conditions on voter preferences which pin down each equilibrium differ between the two types of election. This means that a district with the same f_d can face one platform pair in a single-district election and another pair in a legislative election. In a legislative election, candidate behaviour in the district is completely pinned down by voter preferences, but this is not the case in a single-district - in some equilibria candidates play non-degenerate mixed strategies. As a result, platforms should exhibit more variance in single-district elections (keeping f_d fixed) than the same district would in a legislative election. Moreover, this non-degenerate mixing in single-district elections means the probability of the advantaged candidate winning does not converge to one as the size of the electorate grows. In a single-district election we can have that (i) the disadvantaged candidate wins with positive probability;(ii) the winning platform is not majority-preferred on each dimension; and (iii) a Condorcet winner policy is not chosen by either candidate.[30] None of these hold true for the same district in a legislative election.

[30]This last failure occurs in two different scenarios. In the first scenario, both candidates choose a non-Condorcet winner policy as it is more attractive to swing voters than the Condorcet winner policy. In the second scenario, both candidates mix with probability 0.5, so if a Condorcet winner policy exists, it will only be on the ballot 50% of the time.

Policy Outcomes in Legislative v Single-District Elections In terms of policy, the key differences between these two types of elections are that policies are, first, more predictable and, second, more representative in legislative elections. In terms of predictability, we know from Proposition 4, that as the number of voters gets large, we are guaranteed a unique (expected) policy. This is not the case in a single-district election. There, even as $k \to \infty$ candidates may use non-degenerate mixed strategies. This means that candidates' chosen platforms and final policies are random. There is not a direct mapping from all voter preferences into policy outcomes, while there always is for legislative elections.

How representative are implemented policies in these two types of election? Proposition 4 says that the final policy in a legislative election is preferred by the median voter in the median district on each dimension. Furthermore, Proposition 5 says that if a Legislative Condorcet winner exists, it will be implemented. These results stand in stark contrast to the single-district case. There, the final policy may not be the one preferred by the median voter on each dimension. In addition, there is no guarantee that a Condorcet winner policy will be implemented. Why such a difference? In single-district elections, voters must choose between bundles of policies. As a result, preference intensity does matter - t_{LN} and t_{LY} types always vote v_L while t_{RN} and t_{RY} types always vote v_R. Only types t_{NL},t_{YL},t_{NR},t_{YR} are the swing voters. This gives enormous influence to these swing voters and pulls policies away from those preferred by the median voter on each dimension. This undue influence of minority voters with intense preferences on the secondary dimension is documented in Krasa and Polborn (2010) and Besley and Coate (2008) and finds empirical support in Bouton *et al.* (2020). Besley and Coate (2008) show that citizens initiatives can potentially unbundle policies and bring about outcomes aligned with majority preferences. Proposition 4 shows that such initiatives are unnecessary in a legislative election as voters can unbundle policies via their voting behaviour.

Two well-known limitations of majority rule are worth mentioning when discussing which election type is more representative of voter preferences. First, majority rule implements the will of the majority without regard for the intensity of voter preferences. As such, a policy preferred by a majority of voters might not be that which maximises the sum of voter utility. Second, if a Condorcet winner does not exist, we have Condorcet cycles where each policy can be defeated by at least one other policy. With such cycles it is meaningless to talk of the will of the majority. Given these two properties of majority rule, it is possible to construct preference distributions such that the sum of voter utility is greater under single-district elections than legislative elections. This can occur if (a) a single-district election implements a policy not preferred by a majority on each dimension but which happens to have intense minority support, or (b) there is no Condorcet winner but the policy implemented in a single-district gives a greater sum of voter utility than the policy implemented in a legislative election, perhaps because its majority is larger.

5 Applications

5.1 Parliamentary v Presidential Systems

One of the key choices any new democracy faces is whether to become a presidential or parliamentary system. A vast literature in economics and political science has studied the effect of having either system on size and composition of government spending, growth, responses to crises, tax rates, corruption, and electoral campaign spending.[31] Any analysis must first define how exactly these two systems differ. Persson and Tabellini (2005) consider the difference to be along two dimensions: (i) presidential systems have greater separation of power between the executive and legislature, and (ii) presidents do not require the confidence of the legislature to remain in office. In reality, however, there is large variance in the degree of separation of power both within and across these two systems. In particular, in many strong presidential systems such as Brazil, Chile and Philippines, legislative power is highly concentrated in the hands of the president. Equally, in a strong parliamentary system, power rests with legislators rather than the executive (prime-minister).[32] Here, I focus on a stark comparison between a strong presidential system and a strong parliamentary system. In the former, a president is directly elected by a national (single-district) election and he/she alone determines policy. In the latter, voters elect the legislature and a legislative majority on each dimension determines policy. These represent two extremes.[33] One can think of regimes where policy outcomes are determined jointly by the legislature and executive to varying degrees as existing between these extremes.[34] We have already seen that the incentives of voters are different under these two types of elections and that this results in different candidate platforms and implemented policies. But which system is best for voters?

If there was only one dimension of policy, I could compare under which system the national median voter, $\tilde{t}_{med}$, fares better. However, with multiple dimensions of policy there is no single median voter. The natural equivalent is to examine whether the median voter on each dimension, $\tilde{t}^{LR}_{med} \in \{t_L, t_R\}$ and $\tilde{t}^{NY}_{med} \in \{t_N, t_Y\}$, fares better under a presidential

[31] For the most part, these papers consider the effects of the system of government in models on political agency rather than one of voter preference aggregation. For an overview see Persson and Tabellini (2005).

[32] See Shugart *et al.* (1992) for a detailed discussion on defining presidential and parliamentary systems.

[33] They closely resemble the pure presidential model in Persson *et al.* (1997) and the simple legislature in Persson *et al.* (2000).

[34] The US system, where the president has weak legislative power would nonetheless be closer to the strong presidential system than a weak parliamentary system where the prime-minister exerts some power.

or parliamentary system. A further complication in comparing welfare between these two systems is that the respective national median voters $\tilde{t}^{LR}_{med} \in \{t_L, t_R\}$ and $\tilde{t}^{NY}_{med} \in \{t_N, t_Y\}$ need not coincide with the median voters in the median district on each dimension $\tilde{t}^{LR}_{\frac{D+1}{2}} \in \{t_L, t_R\}$ and $\tilde{t}^{NY}_{\frac{D+1}{2}} \in \{t_N, t_Y\}$. [35] Whether they coincide or not will depend on how voters are distributed across districts. For example, in a US context it could be that the national median voter on the left-right dimension is a Democrat but that the median voter in the median district (Illinois's 13th District) is a Republican.[36] Similarly, the national median on the social issues dimension may be pro-choice but the median in the median district on that dimension may be pro-life. Such cases occur if the distribution of voter preferences across districts is highly asymmetric. In many cases, however, the identities of the national median and the median voter in the median district will coincide. If this is so, the next proposition says that the median voter on each dimension is better off under a parliamentary system than a presidential one.

Proposition 6. *For $k > \hat{k}$, any distribution of voter preferences and any distribution of voters across districts such that $\tilde{t}^h_{med} = \tilde{t}^h_{\frac{D+1}{2}} \forall h \in \{LR, NY\}$, the utility of $\tilde{t}^h_{med}$ is weakly greater under a parliamentary system than a presidential system with probability $\to 1$ as $k \to \infty$.*

Proof. See Appendix D. □

That is, for some distributions of voter preferences, the median voter on each dimension gets the same utility under both systems, while under all other distributions they get strictly greater utility under a parliamentary system. From Proposition 4, we know that the implemented policy in a legislative election is that preferred by the median voter in the median district on each dimension. If these types coincide with national median voters, then a parliamentary election always implements the policies favoured by the national medians on each dimension. Instead, presidential systems suffer from the problems of single-district elections laid out in Appendix B - candidates choose policies to attract swing voters, resulting in final policies which may not be the preferred choice of the median voter on each dimension. The result is somewhat counter-intuitive: presidential systems have a direct election to choose their policy-maker yet it is parliamentary systems with their indirect election of policy-makers (the median legislator on each dimension) that is more representative of voters preferences. This difference occurs because there is more than one dimension of

[35] $\tilde{t}^{LR}_{med} \in \{t_L, t_R\}$ and $\tilde{t}^{NY}_{med} \in \{t_N, t_Y\}$ are respectively at the 50th percentile of the national voter population on that dimension, but $\tilde{t}^{LR}_{\frac{D+1}{2}} \in \{t_L, t_R\}$ and $\tilde{t}^{NY}_{\frac{D+1}{2}} \in \{t_N, t_Y\}$ can be anywhere between the 25th and 75th percentile.

[36] Illinois's 13th District is the median district on the left-right dimension according to Cook's Partisan Voting Index.

policy. In a single-district (presidential) election, voters must choose between two bundles of policies - it is this bundling of issues that can lead to sub-optimal policies. In a legislative (parliamentary) election voters can act strategically to unbundle the issues and vote on a dimension by dimension basis. It is this strategic play of voters that brings about better polices under parliamentary systems.

5.2 Polarisation

There has been much debate about causes and consequences of political polarisation in the US and around the world.[37] In this section, I ask whether increased polarisation can affect candidate platforms and implemented policies; and whether the impact differs between single-district and legislative elections.

Definition. *Let f_d^1 and f_d^2 be two distributions of voters in district d such that $f_d^1(t_{ij}) + f_d^1(t_{ji}) = f_d^2(t_{ij}) + f_d^2(t_{ji}) \forall i \in \{L, R\}, j \in \{N, Y\}$. Moving from f_d^1 to f_d^2 increases **polarisation** if $f_d^2(t_{ij}) \geq f_d^1(t_{ij})$ $\forall i \in \{L, R\}, j \in \{N, Y\}$ with the inequality strict for at least one t_{ij}.*

Notice that because polarisation simply shifts voters preference intensity, any increase in polarisation will leave the respective median voters $\tilde{t}^h_{med}, \tilde{t}^h_{\frac{D+1}{2}}, h \in \{LR, NY\}$ unchanged. As such, polarisation does not alter the *optimal* policy but as the next proposition shows it may change the *implemented* policy.

Proposition 7. *In a legislative election, polarisation will have no effect on voter behaviour, candidate platforms or implemented policies for $k > \bar{k}$. In a single-district election, for any f_d^1 such that $f_d^1(t) > 0 \forall t \in T$ there always exists a distribution with increased polarisation, f_d^2, such that candidate platforms and implemented policies differ from those under f_d^1.*

Proof. See Appendix D □

The fact that polarisation has no impact in a legislative election stems directly from Corollary 2 to Proposition 1 - types t_{ij} and t_{ji} always vote the same way.

In a single-district election, an increase in polarisation will change the composition of swing voters, which in turn changes the platforms candidates campaign on and the policies that are implemented. To take an extreme example, if polarisation is so pronounced that only core party supporters $t_{LN}, t_{LY}, t_{RN}, t_{RY}$ exist - we get multiple equilibria. As there are no more swing voters, choosing pro- or anti-reform policies are both best responses

[37] See Barber and McCarty (2015) for a review of the literature. Gentzkow (2016) argues that it is unclear whether polarisation has in-fact increased in recent times.

for each candidate. More generally, changes in polarisation in a single-district election can lead to any of the four policies being implemented despite no change in $\tilde{t}^{LR}_{med}$, $\tilde{t}^{NY}_{med}$. As the implemented policy in single-district elections may already have been sub-optimal, an increase in polarisation may either increase or decrease the utility of the median voter. In a legislative election, polarisation has no effect on outcomes, so the implemented policy remains that preferred by $\tilde{t}^{LR}_{\frac{D+1}{2}}$ on the left-right dimension and $\tilde{t}^{NY}_{\frac{D+1}{2}}$ on the reform dimension.

Though Proposition 7 considers the specific form of polarisation defined above, the result is more general. For any f^1_d, f^2_d such that $f^1_d(t_L) = f^2_d(t_L)$ and $f^1_d(t_N) = f^2_d(t_N)$, voter behaviour, candidate platforms and implemented policies in a legislative election will not vary when moving from f^1_d to f^2_d. In a single-district election, as before, one can always find a f^2_d which satisfies the constraints but leads to different equilibrium outcomes than f^1_d. This more general result encompasses the case where voter preferences become highly correlated across dimensions, something widely documented among US voters (Barber and McCarty, 2015).

6 Robustness

This section will discuss the robustness of the assumptions made and briefly cover possible extensions.

I model policy-making in the legislature as a sequential process where legislators are committed to voting for their platform on both dimensions. If legislators were not constrained to vote for their platforms, logrolling may be a concern. However, any quid-pro-quo between rival legislator groups would unravel: those who gained votes in the first vote will find it optimal to renege on the deal come the second vote. A trickier issue is if legislators vote on both dimensions simultaneously. If there is a Legislative Condorcet winner, none of my results change. If there is no Legislative Condorcet winner, we have a Condorcet cycle in the legislature, and there is no one-to-one mapping from seat shares to policy outcomes.

I have assumed that candidate platforms are fixed on one dimension. If candidates were free to choose their platform on each dimension, both candidates in each district would choose the majority-preferred position on each dimension. Proposition 4 and Proposition 5 still hold trivially, as the election outcome is predetermined.[38] In a single-district election with flexibility on both dimensions, both candidates choose the majority-preferred position on each dimension only if it is a Condorcet winner. Otherwise, candidates mix over their

[38]In this case, unless $E(S_{-d})$ is pivotal, district d cannot be pivotal, so voter behaviour is not pinned down. Candidate behaviour is determined by the fact their strategies must be best responses to trembles in other candidates' strategies. Given these trembles, district d is pivotal with positive probability ex-ante.

four pure strategies. If a Condorcet loser exists, it is elected with positive probability.

In the model, candidates are purely office-motivated. One might ask; what happens if candidates also care about policy? Might this result in non-majoritarian policies? When $piv^1 = piv(LR)$, reform platforms have no effect on a candidate's win probability, but they do affect candidate payoffs from policy if the district is actually pivotal on the reform dimension. When $piv^1 = piv(NY)$, platform choices always affect a candidate's win probability - because voters vote solely on the reform dimension when platforms differ. In the latter case, the advantaged candidate will have strong incentives to select the majority-preferred position on the reform dimension. In the former case, a candidate who favours a minority-preferred policy will find it optimal to choose such a platform. Overall, as long as $Pr[piv_d^1 = piv(NY)] > Pr[piv(NY)|piv_d^1 = piv(LR)] * Pr[piv_d^1 = piv(LR)]$, the advantaged candidate will choose the majority-preferred policy and the paper's results remain unchanged. As these are equilibrium objects, there is no guarantee that the inequality holds in the current model. It will hold for large enough voter populations if we tweak the model to ensure $Pr[piv_d^1 = piv(NY)] > \epsilon > 0$ for all population sizes.[39] This could be achieved by assuming, for example, that candidates in d believe candidates in other districts play both strategies with positive probability.[40]

I have assumed throughout that there are two binary dimensions of policy, arguing that there are many cases where policy choices are binary (or perceived to be by voters). Do legislative elections still dominate single-district elections if candidate policy choices are non-binary? If there are more than two policies represented in the legislature, it is not guaranteed that one of them has majority support. So, answering our question requires specifying how seat distributions map into policies.[41] Suppose that on each dimension, the platform of the median legislator is implemented. The task for voters is similar to the standard model - they vote based on the most likely pivotal event, which may now be one of several pivotal events on the reform dimension. Suppose the most likely pivotal event is between policies z and $\tilde{z}$ on the reform dimension, and suppose the district median prefers z to $\tilde{z}$. For candidates, equilibrium play is similar to that of Proposition 2: If the share of voters who prefer z to $\tilde{z}$ is greater than the share who prefer the advantaged candidate on the left-right dimension, then both candidates select the district *median*-preferred policy $z_{m,d}$ on the reform dimension. Suppose, instead, the former is smaller than the latter. In that case, the disadvantaged candidate has an incentive to differentiate themselves on the reform

[39]This is because the probability of being pivotal $Pr[piv(NY)|piv_d^1 = piv(LR)]$ goes to zero as the population grows.

[40]This could be because, when selecting their platforms, candidates in d may face uncertainty over $\mathbf{f}_{-d}$ and, therefore, expected platforms in the other $D-1$ districts. Or, more simply, it may be that candidates always have a small chance of making mistakes when choosing their strategy.

[41]Depending on the rule chosen, an equilibrium may not exist Austen-Smith (1984); Osborne (1995).

dimension to decrease their probability of losing while the advantaged candidate favours $z_{m,d}$. In each case, the advantaged candidate selects $z_{m,d}$ with probability going to one. This means the paper's results hold when candidates can choose non-binary policies.[42] We know from Krasa and Polborn (2010) that incentives for candidates to propose non-majority-preferred positions persist in single-district elections when the policy space is non-binary. In a legislative election, as in the baseline binary model, the policy preferred by $t^{NY}_{\frac{D+1}{2}}$ is implemented. If that policy is interior, then moving from binary to a non-binary policy space increases the welfare of $t^{NY}_{\frac{D+1}{2}}$.

The focus on two policy dimensions has been to keep things as tractable as possible. With more dimensions, voters still vote based on the most likely pivotal dimension. The equilibrium play of candidates again mirrors that of Proposition 2. If the majority on a given dimension is greater than the share who prefer the advantaged candidate on the fixed dimension, then both candidates choose the majority-preferred position. If instead, the former is smaller than the latter, candidate behaviour depends on the ranking of pivotal events. Either both choose the majority-preferred position (as above), or they play a matching pennies game as in Proposition 2. In each case, the advantaged candidate selects the majority-preferred position on each dimension with probability going to 1. This means the paper's key propositions all hold when there are multiple binary dimensions on which candidates choose policies.[43] Krasa and Polborn (2010) show that the inefficiencies of single-district elections highlighted in Proposition 8 persist when there are more than two dimensions. As such, the supremacy of legislative elections over single-district elections is robust to including further policy dimensions.

A fundamental assumption driving the paper's main results is that voters are strategic rather than naive. I have already argued that the type of strategic voting here does not require too much mental work from voters. However, if voters in my model were naive, candidates and voters would all behave as if they were in a single-district election. The positive welfare properties laid out in the paper would no longer apply - majority-preferred policies and Condorcet winning policies would no longer be guaranteed. Even so, the paper still sets out the best response of a strategic voter. Regardless of whether others are strategic or naive, she should always cast her ballot based on the dimension where the legislative majority is most likely to change. Notably, as the number of dimensions increases, the utility loss from voting naively rather than strategically grows. Why is this? The only districts with no conflicted voters are those where both candidates choose the same policy

[42]The sketch above is for the case of discrete, non-binary platform choices. A similar logic follows with a continuous policy space; however, pinning down equilibrium play in the median district is more demanding.

[43]A necessary condition is that a more general version of Assumption 1 holds so that candidates believe voters only assess one dimension in equilibrium.

on every dimension they can. With two policy dimensions, this corresponds to two of the four district types. [44] In a model with q dimensions of policy, the share of district types with no conflicted voters is $\frac{1}{2^{q-1}}$.[45] This share goes to zero as the number of policy dimensions increases, showing that the gulf between naive and strategic voting grows as the number of dimensions increases.

7 Conclusion

It is commonly thought that in an election with two parties there can be no strategic voting - your best option is to simply vote for your preferred candidate. This truly is the case in single-district elections with multiple dimensions of policy and multi-district elections with one dimension of policy. In this paper, I showed that when elections have multiple districts and multiple policy dimensions, strategic voting comes to the fore. In contrast to a single-district model, the intensity of a voter's preference on each dimension is irrelevant for her voting decision. Instead, she votes solely based on the dimension most likely to be pivotal in the legislature. Anticipating this behaviour, candidates put forward a different set of policies. I showed that for large elections, the implemented policy bundle: (a) is uniquely pinned down by voter preferences, (b) is the issue-by-issue majority-preferred bundle, (c) is a Condorcet winner if one exists. These properties are seldom guaranteed in a single-district election or if voters are not strategic. I also showed that strong parliamentary systems lead to better policies than strong presidential systems if they have the same median voter. Finally, I established that increased polarisation does not affect policy in a legislative election, while it can significantly change outcomes in a single-district election.

The model has a number of implications for what we should expect to see in real-world elections. First, both voters and candidates in a district will behave differently depending on whether the election is a local (mayoral/gubernatorial) election, or it is part of a legislative election. As a result, the platforms candidates campaign under and the eventual winner may differ between election types even if voter preferences remain constant. This implies we should proceed with caution when interpreting changes in vote share as changes in voter preferences when comparing these different types of elections. Second, the margin of victory in a district will depend on which dimension of policy voters are focusing on. A corollary

[44] (a_{LN}, a_{RN}) and (a_{LY}, a_{RY}) districts have no conflicted voters while (a_{LN}, a_{RY}) and (a_{LY}, a_{RN}) districts do.

[45] If there are q dimensions of policy and each policy is binary, there are 2^q possible policy outcomes. Each district will have one of $2^{2(q-1)}$ platform pairs, and there will be $2^q.q!$ different voter types. There are 2^{q-1} district types where both candidates choose the same platform on all dimensions they can and thus have no conflicted voters. There are $2^{2(q-1)}$ district types where candidates choose different platforms on at least one dimension and thus have conflicted voters.

of Proposition 2 is that if a district has candidates choosing opposing positions on the reform dimension, then the margin of victory for the expected winner will be larger if voters focus on the left-right dimension rather than the reform dimension. Taking this to the real world, we would expect the front-runner to campaign on the core left-right dimension while the underdog tries to woo voters on other dimensions. Third, it is foolish for parties and candidates to ignore an opponent's core voters as a lost cause. Instead, they should target these voters when the left-right dimension is not the relevant one because all voters are potential swing voters in a legislative election. A related implication is that researchers and pollsters should not assume that core supporters of one party never vote for the other party. This may lead to inaccurate predictions when the relevant dimension is non-partisan. An example of this is the 2019 UK General Election where many Labour heartlands elected pro-Brexit Conservative MPs, much to the surprise of political analysts. Fourth, tactical voting groups should update their strategies in light of the fact that there is no conflict of interest between voter groups who prefer the same policies. This should help to prevent scenarios where groups with aligned preferences but different intensities do not cooperate. A notable example where such cooperation failed is in the recent 2019 UK General Election where tactical voting groups wanting to stop Brexit produced conflicting advice to groups trying to prevent a Conservative majority. Similar coordination issues may occur in the US and Canada, where groups also seek to influence voter decisions.[46] A well-meaning voter might have followed the advice on the dimension she cared more about when, in fact, the optimal strategy was the same regardless of which dimension was more important to her. Fifth, polls matter. But the level of information needed is not unrealistic, and the calculations required of voters are not overwhelming. To cast their optimal ballot voters only need to know on which dimension the legislative majority is most likely to change. If they know this, it allows hardline left and right voters to ignore party labels and vote on the dimension where they are more likely to make a difference. This is unlikely to occur in real-world legislative elections unless pollster's projections show not just the likely left-right majority but also the majority on other key dimensions.

[46] In Canada, the website strategicvoting.ca aims to assist voters "to elect as many progressives MPs as possible in federal elections through strategic voting". In the US, tactical voting groups are less common, but newspapers, interest groups and politicians all endorse local candidates with the goal of furthering various national objectives.

Appendix A - Poisson Properties

Single-district Poisson Properties

The number of voters in a district is a Poisson random variable k_d with mean k. The probability of having exactly η voters is $Pr[k_d = \eta] = \frac{e^{-k}k^{\eta}}{\eta!}$. Poisson Voting games exhibit some useful properties. By *environmental equivalence*, from the perspective of a player in the game, the number of other players is also a Poisson random variable k_d with mean k. By the *decomposition property*, the number of voters with type $t \in T$ is Poisson distributed with mean $\sum_{t \in T} k f_d(t)$, and is independent of the number of other types.

The probability of a specific vote profile $x_d = (x_d(L), x_d(R))$ given voter strategies is

$$Pr[x_d|k\nu_d] = \frac{e^{-k\nu_d}(k\nu_d)^{x_d(L)}}{x_d(L)!} \frac{e^{-k(1-\nu_d)}(k(1-\nu_d))^{x_d(R)}}{x_d(R)!} \tag{8}$$

Its associated magnitude is

$$mag[x_d] \equiv \lim_{k\to\infty} \frac{\log(Pr[x_d|k\nu_d])}{k} = \lim_{k\to\infty} \nu_d \psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right) \tag{9}$$

where $\psi(\theta) = \theta(1 - log(\theta)) - 1$.

Magnitude Theorem Let an event A_d be a subset of all possible vote profiles in district d. The magnitude theorem (Myerson (2000)) states that for a large population of size k, the magnitude of an event, $mag[A_d]$, is:

$$mag[A_d] \equiv \lim_{k\to\infty} \frac{\log(Pr[A_d])}{k} = \lim_{k\to\infty} \max_{x_d \in A_d} \nu_d \psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right) \tag{10}$$

That is, as $k \to \infty$, the magnitude of an event A_d is simply the magnitude of the most likely vote profile $x_d \in A_d$. The magnitude $mag[A_d] \in [-1, 0]$ represents the speed at which the probability of the event goes to zero as $k \to \infty$; the more negative its magnitude, the faster that event's probability converges to zero.

Corollary to the Magnitude Theorem If two events A_d and A'_d have $mag[A_d] > mag[A'_d]$, then their probability ratio converges to zero as $k \to \infty$.

$$mag[A'_d] < mag[A_d] \Rightarrow \lim_{k\to\infty} \frac{Pr[A'_d]}{Pr[A_d]} = 0 \tag{11}$$

Suppose we have a 2-candidate election with $\nu_d > 0.5$, so that the left candidate has a higher expected vote share.

Maximising Equation 10 subject to the appropriate constraints we get

$$\begin{aligned} mag[Lwin] &= 0 \\ mag[Rwin] &= 2\sqrt{\nu_d(1-\nu_d)} - 1 \end{aligned} \tag{12}$$

With a magnitude of zero, by the corollary, the probability of the left candidate winning goes to 1 as k gets large.

From equation 3.1 in Myerson (2000), the probability of a tie can be approached by

$$Pr[x_d(L) = x_d(R)] \approx \frac{e^{k(2\sqrt{\nu_d(1-\nu_d)}-1)}}{\pi(k+\frac{1}{3})} \tag{13}$$

Multi-District Poisson Properties

Let $\boldsymbol{x} \equiv (x_1, \ldots, x_d, \ldots, x_D)$ be the realised profile of votes across districts. The probability of a particular profile of votes is

$$Pr[\boldsymbol{x}|k\boldsymbol{\nu}] = \prod_{d\in D} \frac{e^{-k\nu_d}(k\nu_d)^{x_d(L)}}{x_d(L)!} \frac{e^{-k(1-\nu_d)}(k(1-\nu_d))^{x_d(R)}}{x_d(R)!} \tag{14}$$

After some manipulation, taking the log of both sides, and taking the limit as $k \to \infty$ we get the magnitude of this profile of votes

$$mag[\boldsymbol{x}] \equiv \lim_{k\to\infty} \frac{log(Pr[\boldsymbol{x}|k\boldsymbol{\nu}])}{k} = \lim_{k\to\infty} \sum_{d\in D} \nu_d \psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right) = \sum_d mag[x_d] \tag{15}$$

Multi-District Magnitude Theorem

Let $\boldsymbol{A} = (A_1, \ldots, A_d, \ldots, A_D)$ be a multi-district event, where each A_d is a particular district event. Let $\bar{x_d} \in A_d = \underset{x_d}{\text{argmax}}\ \nu_d\psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right)$, that is, $\bar{x_d}$ is the most likely district vote profile in A_d given ν_d . Then, Multi-District Magnitude Theorem Hughes (2016) states that:

$$mag[\boldsymbol{A}] = \sum_{d=1}^{D} mag[A_d] = \sum_{d=1}^{D} mag[\bar{x_d}] = mag[\bar{\boldsymbol{x}}] \tag{16}$$

The first inequality follows from the independence of districts; the second equality follows from the magnitude theorem and the independence of districts; the third equality follows from Equation 15. Together they show that the single-district magnitude theorem extends to a multi-district setting.

Following from this, and using Equation 11, we have that the corollary to the magnitude theorem also extends to the multi-district case. If $mag[\boldsymbol{A'}] < mag[\boldsymbol{A}]$, then

$$\lim_{k\to\infty} \frac{Pr[\boldsymbol{A'}|k\boldsymbol{\nu}]}{Pr[\boldsymbol{A}|k\boldsymbol{\nu}]} = \lim_{k\to\infty} \frac{e^{kmag[\boldsymbol{A'}]}}{e^{kmag[\boldsymbol{A}]}} = \lim_{k\to\infty} e^{k(mag[\bar{\boldsymbol{x}}'] - mag[\bar{\boldsymbol{x}}])} = 0 \tag{17}$$

For district d, let $\tilde{\mathbf{K}}_{\mathbf{d}}(\boldsymbol{A_{-d}})$ be the set of $D-1$ district events with largest magnitude required for $\boldsymbol{A_{-d}}$ to occur. Let $\mathbf{K_d}(\boldsymbol{A_{-d}}) \subset \tilde{\mathbf{K}}_{\mathbf{d}}(\boldsymbol{A_{-d}})$ be the largest subset such that each element of $\mathbf{K_d}(\boldsymbol{A_{-d}})$ has a non-zero magnitude. One can think of $\mathbf{K_d}(\boldsymbol{A_{-d}})$ as the set of electoral upsets with largest magnitude necessary for $\boldsymbol{A_{-d}}$ to occur. It follows that $mag[\mathbf{K_d}(\boldsymbol{A_{-d}})] = mag[\tilde{\mathbf{K}}_{\mathbf{d}}(\boldsymbol{A_{-d}})] = mag[\boldsymbol{A_{-d}}]$. From Lemma 2 in Hughes (2016), we have

$$\mathbf{K_d}(\boldsymbol{A_{-d}}) \subset \mathbf{K_d}(\boldsymbol{A'_{-d}}) \implies mag[\boldsymbol{A_{-d}}] > mag[\boldsymbol{A'_{-d}}] \tag{18}$$

Appendix B - Single-District Model

In this section, I characterise the equilibria of a single-district model first developed by Krasa and Polborn (2010). Two candidates compete in a single district. As in the legislative model, each candidate is constrained on the left-right dimension but free to choose a pro- or anti-reform policy. Voters vote by majority rule, and the winning candidate implements his platform as policy. Most of the results in this section follow directly from Krasa and Polborn (2010). However, they do not analyse the equilibrium properties as the size of the electorate increases.[47] By modelling the setup as a Poisson game and looking for asymptotic equilibria, I can show what happens to chosen platforms and implemented policies as the number of voters increase.

The voting stage is straightforward in a single-district election. There are three differences between voting here and in legislative elections. First, voting can no longer be a strategic choice because voters can only be pivotal between the two policy platforms offered by candidates. For each platform pair a voter faces, her preferred option is pinned down by her type, so her vote choice is deterministic. Table 3 shows the voting behaviour of each type in each of the four scenarios.

District	Vote v_L	Vote v_R
(a_{LN}, a_{RN}) (a_{LY}, a_{RY})	$t_{NL}, t_{LN}, t_{YL}, t_{LY}$	$t_{NR}, t_{RN}, t_{YR}, t_{RY}$
(a_{LN}, a_{RY})	$t_{NL}, t_{LN}, t_{NR}, t_{LY}$	$t_{YL}, t_{RN}, t_{YR}, t_{RY}$
(a_{LY}, a_{RN})	$t_{YL}, t_{LN}, t_{YR}, t_{LY}$	$t_{NL}, t_{RN}, t_{NR}, t_{RY}$

Table 3: Best responses of each voter type

Second, from Table 3, we see that preference intensity does matter here. Types t_{ij} and t_{ji} vote the same way in only three out of the four cases. This creates a conflict of interest between voter groups that want the same policy implemented but disagree about which is the second-best policy. Third, we see that t_{LN} and t_{LY} types always vote v_L while t_{RN} and t_{RY} types always vote v_R. The fact that these voters are never swing voters means candidates can safely ignore them when making their platform choices.

At the candidate competition stage, knowing how each voter type will vote, the left candidate will choose μ_L to maximise the expression below, while the right candidate will

[47] In their model, there is no population uncertainty, so there would be no effect of increasing the size of the electorate.

choose μ_R to minimise it.

$$
\begin{aligned}
Pr[Lwin|\mu_L,\mu_R] &= (\mu_L\mu_R + (1-\mu_L)(1-\mu_R))Pr[Lwin|(a_{LN},a_{RN})] \\
&+ \mu_L(1-\mu_R)Pr[Lwin|(a_{LY},a_{RN})] \\
&+ (1-\mu_L)(\mu_R)Pr[Lwin|(a_{LN},a_{RY})]
\end{aligned} \tag{19}
$$

Letting $\hat{t}_{LY} \equiv t_{LY} \cup t_{YL} \cup t_{YR} \cup t_{LN}$ and $\hat{t}_{LN} \equiv t_{LN} \cup t_{NL} \cup t_{NR} \cup t_{LY}$ I can state the following proposition.

Proposition 8. *For any given f_d, the candidate equilibrium in a single-district election is unique. The equilibrium strategies are represented in Table 4.*

Proof. See Appendix D. □

Case	District Preferences	μ_L^*	μ_R^*	a_d^*
1	$f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$	0	0	(a_{LN}, a_{RN})
2	$f_d(\hat{t}_{LY}) > f_d(t_L) > f_d(\hat{t}_{LN})$	1	1	(a_{LY}, a_{RY})
	$f_d(\hat{t}_{LN}), f_d(\hat{t}_{LY}) > f_d(t_L)$			
3a	$f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L), 0.5$	$\bar{\mu}_L \to 0.5$	$\bar{\mu}_R \to 0.5$	mixed
3b	$0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$	$\bar{\mu}_L \to 0$	$\bar{\mu}_R \to 1$	(a_{LN}, a_{RY})
3c	$0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$	$\bar{\mu}_L \to 1$	$\bar{\mu}_R \to 0$	(a_{LY}, a_{RN})
	$f_d(t_L) > f_d(\hat{t}_{LN}), f_d(\hat{t}_{LY})$			
4a	$0.5, f_d(t_L) > f_d(\hat{t}_{LN}), f_d(\hat{t}_{LY})$	$\bar{\mu}_L \to 0.5$	$\bar{\mu}_R \to 0.5$	mixed
4b	$f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$	$\bar{\mu}_L \to 1$	$\bar{\mu}_R \to 0$	(a_{LY}, a_{RN})
4c	$f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$	$\bar{\mu}_L \to 0$	$\bar{\mu}_R \to 1$	(a_{LN}, a_{RY})

Table 4: Equilibrium platforms in the single-district candidate competition game. As $k \to \infty$, mixed strategy equilibria converge as shown.

Examining Table 4, a number of differences stand out from the legislative model in Table 2.

First, there are more cases to consider in a single-district election. This is because in a legislative election the relative size of $f_d(t_L)$ and $f_d(t_N)$ pins down the candidate equilibrium, while in a single-district election the relative sizes of $f_d(t_L)$, $f_d(\hat{t}_{LN})$ and $f_d(\hat{t}_{LY})$ all matter. This difference stems from the fact that in a single-district election only a subset of voters are swing voters, while in a legislative election all voters are.

Second, we see that the mapping from district preferences into policy pairs differs between the two tables. That is, the same district preferences may lead to one candidate equilibrium in a single-district election and a different one in a legislative election.

Third, even as $k \to \infty$ we can have equilibria in mixed strategies which are not degenerate in a single-district election. For the district preferences in case 3a and 4a the realisation of platforms and the eventual winning policy is random. Furthermore, the advantaged candidate does not win with probability $\to 1$. This contrasts with the legislative election case where as $k \to \infty$ voter preferences pin down policy and the advantaged candidate wins with probability $\to 1$.

Finally, from Table 4 we can see that a Condorcet winner (if it exists) may not win a single-district election. This can happen in two different ways. First, it may be that there is a Condorcet winner but preferences correspond to case 3a or 4a so that candidates randomise equally over platforms. Here, with probability 0.5, the Condorcet winner policy will not be on the ballot. Second, in cases 1 and 2, a Condorcet winner policy may exist but not be chosen by a candidate. For example, this occurs in case 1 if $f_d(t_Y) > f_d(t_N)$ so that a_{LY} or a_{RY} is a Condorcet winner, but the equilibrium platforms are be (a_{LN}, a_{RN}).

Appendix C - Double-Pivotal Events

Alternatives to Assumption 1

Assumption 1 states that candidates believes that voters never believe $piv_d^1 = piv(LR, NY)$. Its importance is that it reduces

$$\begin{aligned} Pr[Lwin] = Pr[Lwin|piv_d^1 = piv(NY), \mu_L, \mu_R] * Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] \\ +Pr[Lwin|piv_d^1 = piv(LR)] * Pr[piv_d^1 = piv(LR)|\boldsymbol{\mu_{-d}}] \\ +Pr[Lwin|piv_d^1 = piv(LR, NY), \mu_L, \mu_R] * Pr[piv_d^1 = piv(LR, NY)|\boldsymbol{\mu_{-d}}] \end{aligned} \tag{20}$$

to Equation 6.

Choosing μ_L to maximise Equation 6 is equivalent to doing so for Equation 7, as $Pr[Lwin|piv_d^1 = piv(LR)]$ is independent of (μ_L, μ_R). Therefore, any assumption that reduces Equation 20 to Equation 6 will leave the paper's results unchanged. Specifically, either (or both) of the following assumptions can be substituted in place of Assumption 1:

- **Assumption 1a:** Voters believe $piv(LR, NY)$ is never the most likely pivotal event.
- **Assumption 1b:** Candidates believe $piv(LR, NY)$ is never the most likely pivotal event.

Under each assumption, candidates ignore the possibility that voters condition on double-pivotal events.

Role of Assumption 2

While Assumption 1 is required for Proposition 2 and subsequent results, Assumption 2 is a restriction on district preferences which makes invoking Assumption 1 more natural. Specifically, Assumption 2 guarantees that the following lemma holds.

Lemma 2. *For $k > \bar{k}$ we have $Pr[piv_d^1 = piv(LR, NY)] \to 0$ as $k \to \infty$.*

Proof. If $E(S_{-d})$ is such that $E(s_{LN}) \neq E(s_{RY})$ and $E(s_{LY}) \neq E(s_{RN})$, then either $\mathbf{K_d}(piv(LR)) \subset \mathbf{K_d}(piv(LR, NY))$ or $\mathbf{K_d}(piv(NY)) \subset \mathbf{K_d}(piv(LR, NY))$. It follows from Equation 18 that $mag[piv(LR, NY)] < \max\{mag[piv(LR)], mag[piv(NY)]\}$. Therefore, we cannot have $piv_d^1 = piv(LR, NY)$ if $|E(s_{LN}) - E(s_{RY})| > 1$ and $|E(s_{LY}) - E(s_{RN})| > 1$. From Proposition 2 and Proposition 3 we know that $Pr(E(s_i = D_i) \to 1$ as $k \to \infty$. Therefore, as $k \to \infty$, Assumption 2 implies $Pr[|E(s_{LN}) - E(s_{RY})| > 1]$ and $Pr[|E(s_{LY}) - E(s_{RN})| > 1]$ go to one. In turn, this gives the desired result of $Pr[piv_d^1 = piv(LR, NY)] \to 0$ as $k \to \infty$. □

Conditioning on $piv(LR, NY)$

If Assumption 2 does not hold, it is possible for a district to have $piv_d^1 = piv(LR, NY)$ even for large electorates. A necessary but not sufficient condition is that $E(S_{-d})$ is such that $E(s_{LN}) = E(s_{RY})$ or $E(s_{LY}) = E(s_{RN})$.

There are two cases to consider:

(i) If $E(S_{-d})$ is such that $E(s_{LN}) = E(s_{RY})$ and $E(s_{LY}) = E(s_{RN})$, then d is pivotal on both dimensions in expectation. Clearly $piv_d^1 = piv(LR, NY)$.

(ii) If $E(S_{-d})$ is such that $E(s_{LN}) = E(s_{RY})$ or $E(s_{LY}) = E(s_{RN})$, then $piv_d^1 = piv(LR, NY)$ if and only if $mag[\mathbf{K_d}(piv(LR, NY))] > \max\{mag[\mathbf{K_d}(piv(LR))], mag[\mathbf{K_d}(piv(NY))]\}$.

The condition in (ii) is not easy to satisfy. I will show the case where $E(s_L) = E(s_Y) = \frac{D-1}{2} + x$, where $x > 0$, as the other three cases are equivalent. We can think of $\mathbf{K_d}(piv(LR, NY))$ as the chain of x electoral upsets necessary for $piv(LR, NY)$ to occur. Let $\underline{x} \in \mathbf{K_d}(piv(LR, NY))$ be the upset with smallest magnitude. For $mag[\mathbf{K_d}(piv(LR, NY))] > \max\{mag[\mathbf{K_d}(piv(LR))], mag[\mathbf{K_d}(piv(NY))]\}$ to hold, it must be that the magnitude of $\underline{x}$ is greater than the magnitudes of certain upsets in the other $D-1-x$ districts. Specifically, the magnitude of $\underline{x}$ must be greater than (a) the magnitude of the expected winner losing in any (a_{LN}, a_{RY}) district, and (b) the magnitude of an advantaged left candidate losing in any (a_{LN}, a_{RN}) or (a_{LY}, a_{RY}) district. That is, to generate $piv_d^1 = piv(LR, NY)$, we not only need a very specific expected seat distribution; we also need a very exact set of upsets to occur.[48]

[48]Furthermore,in each of the x districts: the platforms are (a_{LY}, a_{RN}), the left candidate is the expected winner, and voters condition on the right candidate winning.

Appendix D - Proofs

Proof of Lemma 1

Proof. **Step 1:** *Unconflicted voters have a dominant strategy for any* $\boldsymbol{a}$*.*

By definition, the utility differential terms in the gain function of an unconflicted voter all have the same sign. As each term is multiplied by a non-negative pivot probability, an unconflicted voter with positive utility differentials must have a non-negative gain function. Therefore, v_L is a weakly dominant strategy for such a voter. An unconflicted voter with negative utility differentials must have a non-positive gain function, making v_R a weakly dominant strategy for such a voter.

Step 2: *The best response of a conflicted voter depends on the strategies of voters and candidates in the other* $D-1$ *districts.*

To show that a conflicted voter is a strategic voter, we need to show that for some profile of candidate and voter strategies in the other $D-1$ districts, her gain function is positive, while for some other profile of strategies, her gain function is negative. Let district d be a (a_{LY}, a_{RN}) district. The conflicted types here are $t_{LN}, t_{NL}, t_{RY}, t_{YR}$. Let strategies in the other $D-1$ districts be as follows: in each district, both candidates choose anti-reform platforms; in $\frac{D-1}{2}$ districts all voter types vote v_L, while in the other $\frac{D-1}{2}$ all types vote v_R. This gives $Pr[piv_d(LR|N)] = 1$. The gain function is positive for t_{LN}, t_{NL} but negative for t_{RY}, t_{YR}. Now, instead, let strategies in the other $D-1$ districts be as follows: in each district, all types vote v_L; in $\frac{D-1}{2}$ districts both candidates choose anti-reform platforms, while in the other $\frac{D-1}{2}$ both candidates choose pro-reform platforms. This gives $Pr[piv_d(NY|L)] = 1$. The gain function is now negative for t_{LN}, t_{NL} but positive for t_{RY}, t_{YR}. Using these same two strategy profiles, it is clear that all other conflicted voters - types $t_{LY}, t_{YL}, t_{RN}, t_{NR}$ in (a_{LY}, a_{RN}) districts - are also strategic voters.

□

Proof of Proposition 1

Proof. **Step 1:** *If voters focus on* piv_d^1 *alone, the best response for each voter type is given by Table 1.*

This comes immediately from each voter's gain function.

Step 2: *Unconflicted voters behave as if they condition on* piv_d^1*.*

By Lemma 1, unconflicted voters have a dominant strategy. Therefore, it is without loss of generality to assume they condition on piv_d^1.

Step 3: *For any $k > \bar{k}$, all conflicted voters condition on piv_d^1 if $mag[piv_d^1] > mag[piv_d^2]$.*

This part of the proof follows the same steps for any two most likely pivotal events and for either (a_{LY}, a_{RN}) or (a_{LN}, a_{RY}). Without loss of generality, take a (a_{LY}, a_{RN}) district and suppose we have $piv_d^1 = piv_d(LR|Y)$, and $piv_d^2 = piv_d(NY|R)$.The relevant gain function is Equation 4. We can divide both sides by $Pr[piv_d(LR|Y)]$ to get

$$\begin{aligned}\frac{G_{t,d}(v_L|(a_{LY}, a_{RN}), k\boldsymbol{\nu}, \mathbf{a_{-d}})}{Pr[piv_d(LR|Y)]} &= U(z_{LY}|t) - U(z_{RY}|t) \\ &+\frac{Pr[piv_d(LR|N)]}{Pr[piv_d(LR|Y)]}(U(z_{LN}|t) - U(z_{RN}|t)) \\ &+\frac{Pr[piv_d(NY|L))]}{Pr[piv_d(LR|Y)]}(U(z_{LY}|t) - U(z_{LN}|t)) \\ &+\frac{Pr[piv_d(NY|R)]}{Pr[piv_d(LR|Y)]}(U(z_{RY}|t) - U(z_{RN}|t)) \\ &+\frac{Pr[piv_d(LR, NY)]}{Pr[piv_d(LR|Y)]}(U(z_{LY}|t) - U(z_{RN}|t))\end{aligned} \tag{21}$$

A voter need only condition their vote on the pivotal event $piv_d(LR|Y)$ if

$$\begin{aligned}|U(z_{LY}|t) - U(z_{RY}|t)| > \; &|\frac{Pr[piv_d(LR|N)]}{Pr[piv_d(LR|Y)]}(U(z_{LN}|t) - U(z_{RN}|t)) \\ &+\frac{Pr[piv_d(NY|L))]}{Pr[piv_d(LR|Y)]}(U(z_{LY}|t) - U(z_{LN}|t)) \\ &+\frac{Pr[piv_d(NY|R)]}{Pr[piv_d(LR|Y)]}(U(z_{RY}|t) - U(z_{RN}|t)) \\ &+\frac{Pr[piv_d(LR, NY)]}{Pr[piv_d(LR|Y)]}(U(z_{LY}|t) - U(z_{RN}|t))|\end{aligned} \tag{22}$$

By the triangle inequality, and as $Pr[piv_d(NY|R)] > Pr[piv_d(NY|L))], Pr[piv_d(LR|N)], Pr[piv_d(LR, NY)]$, Equation 22 must hold if

$$\begin{aligned}|U(z_{LY}|t) - U(z_{RY}|t)| > \; &\frac{Pr[piv_d(NY|R)]}{Pr[piv_d(LR|Y)]}\Big[(|U(z_{LN}|t) - U(z_{RN}|t)|) \\ &+(|U(z_{LY}|t) - U(z_{LN}|t)|) \\ &+(|U(z_{RY}|t) - U(z_{RN}|t)|) \\ &+(|U(z_{LY}|t) - U(z_{RN}|t)|)\Big]\end{aligned}$$

It follows that voters will condition only on $piv_d(LR|Y)$ if

$$\frac{\underline{U}}{3\overline{U}} > \frac{Pr[piv_d(NY|R)]}{Pr[piv_d(LR|Y)]} \tag{23}$$

Where the denominator on the left hand side come from the fact that a conflicted voter can have at most three utility differentials of the same sign. By the corollary to the multi-district magnitude theorem in Section 7, $mag[piv_d(LR|Y)] > mag[piv_d(NY|R)]$ implies

$$\lim_{k\to\infty} \frac{Pr[piv_d(NY|R)]}{Pr[piv_d(LR|Y)]} = 0$$

As $0 < \underline{U} < \overline{U} < \infty$, there must exist a $0 < \bar{k} < \infty$ such that for all $k > \bar{k}$, Equation 23 holds.

Step 4: *In any strictly perfect equilibrium with $k > \bar{k}$, a district with conflicted voters has a unique most likely pivotal event.*

This comes directly Lemma 3.

□

Lemma 3

Lemma 3. *In any strictly perfect equilibrium with $k > \bar{k}$, a district with conflicted voters cannot have $mag[piv_d^1] = mag[piv_d^2]$.*

Proof. Suppose, to the contrary, we have $mag[piv_d^1] = mag[piv_d^2]$. Clearly, $E(S_{-d}|piv_d^1) \neq E(S_{-d}|piv_d^2)$ as no two pivotal events in P can stem from the same seat distribution. Therefore in these two pivotal events, the expected winner must differ in at least one district; call this district d'. Suppose, without loss of generality, that we have $\nu_{d'} > 0.5$ and the expected winner in d' conditional on piv_d^1 is L but the expected winner conditional on piv_d^2 is R. Therefore, by Equation 12 we must have $mag[Lwin(d')] = 0$ and $mag[Rwin(d')] = 2\sqrt{\nu_{d'}(1-\nu_{d'})} - 1$. That is, even though the overall magnitudes $mag[piv_d^1]$ and $mag[piv_d^1]$ are equal, the magnitudes and expected vote shares *in each district* are not - they differ in at least district d'.

Recall that an equilibrium $\boldsymbol{\sigma}^*$ is strictly perfect if and only if there exists $\epsilon > 0$ such that if $\forall \tilde{\nu_d} \in \Delta V : |\tilde{\nu_d} - \nu_d(\boldsymbol{\sigma}^*, \mathbf{f})| < \epsilon$, then $\sigma^*_{t,d} \in BR_{t,d}(k\tilde{\boldsymbol{\nu}})$ for all $t \in T$. Choose a tremble such that $\tilde{\nu_{d'}} > \nu_{d'} > 0.5$ but $\tilde{\nu_{-d'}} = \nu_{-d'}$. This gives $mag[Rwin(d')|\tilde{\nu_{d'}}] < mag[Rwin(d')|\nu_{d'}] < mag[Lwin(d')|\tilde{\nu_{d'}}] = mag[Lwin(d')|\nu_{d'}]$. As the magnitudes in all other districts remain unchanged, we have $mag[piv_d^1|\tilde{\nu_{d'}}] > mag[piv_d^2|\tilde{\nu_{d'}}]$. Under this tremble, by Step 3 of Proposition 1, conflicted voters focus exclusively on piv_d^1 with their best responses given in Table 1. Now choose another tremble such that $\nu_{d'} > \tilde{\nu_{d'}} > 0.5$ but $\tilde{\nu_{-d'}} = \nu_{-d'}$. This gives

$mag[Rwin(d')|\nu_{d'}] < mag[Rwin(d')|\tilde{\nu_{d'}}] < mag[Lwin(d')|\tilde{\nu_{d'}}] = mag[Lwin(d')|\nu_{d'}]$. As the magnitudes in all other districts remain unchanged, we have $mag[piv_d^1|\tilde{\nu_{d'}}] < mag[piv_d^2|\tilde{\nu_{d'}}]$. Under this tremble, by Step 3 of Proposition 1, conflicted voters focus exclusively on piv_d^2 with their best responses given in Table 1. As the best responses of conflicted voters differ depending on whether they condition on piv_d^1) or piv_d^2, it is clear that any tremble in $\boldsymbol{\nu}$ at $mag[piv_d^1] = mag[piv_d^2]$ causes the best response of conflicted voters to change discontinuously. Therefore, any equilibrium with conflicted voters, $k > \bar{k}$ and $mag[piv_d^1] = mag[piv_d^2]$ cannot be a strictly perfect equilibrium.

□

Proof of Proposition 2

Proof. The proof proceeds in five steps. First, I show that if voters in a district are due to focus on the left-right dimension, any combination of platforms is a mutual best response. Second, I show that if voters in a district are due to focus on the reform dimension there is a unique mutual best response given in Table 2 which depends only on f_d. Third, I show that as long as there is positive probability that voters will focus on the reform dimension, candidates will behave as if voters focus entirely on the reform dimension. Fourth, I show that in any strictly perfect equilibrium, candidates will behave as if voters focus entirely on the reform dimension - even if equilibrium strategies are such that $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] = 0$. Fifth, for the candidate equilibria given in Table 2, I show that as the number of voters goes to infinity, all mixed strategy equilibria converge to pure strategy equilibria.

Step 1: *If $k > \bar{k}$ and $piv_d^1 = piv(LR)$, then any strategy profile (μ_L, μ_R) is a mutual best response.*

By Proposition 1, when $k > \bar{k}$ and $piv_d^1 = piv(LR)$, voters vote exclusively on the left-right dimension. As such, $Pr[Lwin|piv_d^1 = piv(LR)]$ is not a function of μ_d. Therefore, any (μ_L, μ_R) constitutes a mutual best response.

Step 2: *If $k > \bar{k}$ and $piv_d^1 = piv(NY)$, there is a unique strategy profile (μ_L, μ_R) which is a mutual best response. These best responses are given by Table 2.*

By Proposition 1, when $k > \bar{k}$ and $piv_d^1 = piv(NY)$, voters vote exclusively on the reform dimension. The expected probability of winning for the left candidate is given by Equation 7. The left candidate chooses μ_L to maximise it, while the right candidate chooses μ_R to minimise it. This yields the following best response correspondences:

$$BR_L(\mu_R) = \begin{cases} 1 & \text{if } \mu_R > \tilde{\mu}_R \\ [0,1] & \text{if } \mu_R = \tilde{\mu}_R \\ 0 & \text{if } \mu_R < \tilde{\mu}_R \end{cases} \tag{24}$$

$$BR_R(\mu_L) = \begin{cases} 1 & \text{if } \mu_L < \tilde{\mu}_L \\ [0,1] & \text{if } \mu_L = \tilde{\mu}_L \\ 0 & \text{if } \mu_L > \tilde{\mu}_L \end{cases} \tag{25}$$

where $\tilde{\mu}_L$ and $\tilde{\mu}_R$ are given by:

$$\tilde{\mu}_L \equiv \frac{Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] + Pr[Lwin|(a_{LN}, a_{RN})] - 1}{2Pr[Lwin|(a_{LN}, a_{RN})] - 1} \tag{26}$$
$$\tilde{\mu}_R \equiv 1 - \tilde{\mu}_L$$

I divide districts into 6 cases and analyse equilibrium strategies in each:

- **Case 1:** If $f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$, then $\mu_L = \mu_R = 0$. I prove this for the case of $f_d(t_N) > f_d(t_L) > 0.5$, as the other case is identical. $f_d(t_N) > f_d(t_L) > 0.5$ and voter strategies given in Table 1 imply that $1 - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > Pr[Lwin|(a_{LN}, a_{RN}) > 0.5$. This in turn means that $\tilde{\mu}_L < 0$ and $\tilde{\mu}_R > 1$ and that both are playing a best response at $(\mu_L, \mu_R) = (0, 0)$

- **Case 2:** If $f_d(t_Y) > f_d(t_L), 1 - f_d(t_L)$, then $\mu_L = \mu_R = 1$. I prove this for the case of $f_d(t_Y) > f_d(t_L) > 0.5$, as the other case is identical. $f_d(t_Y) > f_d(t_L) > 0.5$ and voter strategies given in Table 1 imply that $Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > Pr[Lwin|(a_{LN}, a_{RN}) > 0.5$. This in turn means that $\tilde{\mu}_L > 1$ and $\tilde{\mu}_R < 0$ and that both are playing a best response at $(\mu_L, \mu_R) = (1, 1)$

- **Case 3a & 3b:** If $f_d(t_L) > f_d(t_N) > 0.5$ or $f_d(t_R) > f_d(t_Y) > 0.5$, then $(\mu_L, \mu_R) = (\tilde{\mu}_L, \tilde{\mu}_R)$. I prove this for the case of $f_d(t_L) > f_d(t_N) > 0.5$, as the other is identical. $f_d(t_L) > f_d(t_N) > 0.5$ and voter strategies given in Table 1 imply that $Pr[Lwin|(a_{LN}, a_{RN}) > 1 - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$. This in turn means that $\tilde{\mu}_L \in (0, 0.5)$ and $\tilde{\mu}_R \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu}_L, \tilde{\mu}_R)$.

- **Cases 4a & 4b:** If $f_d(t_L) > f_d(t_Y) > 0.5$ or $f_d(t_R) > f_d(t_N) > 0.5$, then $(\mu_L, \mu_R) = (\tilde{\mu}_L, \tilde{\mu}_R)$. I prove this for the case of $f_d(t_L) > f_d(t_Y) > 0.5$, as the other is identical. $f_d(t_L) > f_d(t_Y) > 0.5$ and voter strategies given in Table 1 imply that

$Pr[Lwin|(a_{LN}, a_{RN}) > Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$. This in turn means that $\tilde{\mu}_L \in (0.5, 1)$ and $\tilde{\mu}_R \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu}_L, \tilde{\mu}_R)$.

Step 3: *If $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] > 0$, the left candidate chooses μ_L to maximise Equation 7, while the right candidate will choose μ_R to minimise it. The candidate mutual best responses are the same as those with $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] = 1$ given by Table 2.*

This follows directly from Steps 1 and 2. If any strategy is a best response when $Pr[piv_d^1 = piv(LR)|\boldsymbol{\mu_{-d}}] = 1$ and mutual best responses are given by Table 2 when $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] = 1$; then it must be that mutual best responses are given by Table 2 whenever $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] > 0$.

Step 4: *For $k > \bar{k}$, in any strictly perfect equilibrium, it must be that candidates' best responses in each district are as in Table 2.*

To prove this step, I must show that even if $Pr[piv_d^1 = piv(LR)|\boldsymbol{\mu_{-d}}] = 1$, the only equilibria which are strictly perfect, are those identified in Table 2. Recall that a Nash equilibrium $\boldsymbol{\mu^*}$ is a strictly perfect equilibrium of Γ if for any arbitrary sequence of perturbations α^η such that $0 < \alpha_{d,i}^\eta < \epsilon^\eta$ and $\epsilon^\eta \to 0$ as $\eta \to \infty$, there exists a sequence of totally mixed strategy Nash equilibria $\boldsymbol{\mu}^\eta$ of the game $\Gamma(\epsilon^\eta)$ such that $\boldsymbol{\mu}^\eta \to \boldsymbol{\mu}^*$ as $\eta \to \infty$. In any perturbed game where players play completely mixed strategies we must have $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] > 0$. To see this, note that if the realised platforms are (a_{LN}, a_{RN}) in $\frac{D-1}{2}$ districts and (a_{LY}, a_{RY}) in $\frac{D-1}{2}$, then we have $Pr[piv_d(NY)] = 1$ for the remaining district. This realisation of platforms is always possible given completely mixed strategies. From Step 3, if $Pr[piv_d^1 = piv(NY)|\boldsymbol{\mu_{-d}}] > 0$, the left candidate chooses μ_L to maximise Equation 7, while the right candidate will choose μ_R to minimise it.

- **Case 1:** If $f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$, then $(\mu_L, \mu_R) = (\alpha_{d,L}^\eta(Y), \alpha_{d,R}^\eta(Y))$. Equation 7 is strictly decreasing in μ_L and strictly increasing in μ_R when $f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$. Given the constraint of the perturbed game that $\alpha_{d,i}(Y) < \mu_i < 1 - \alpha_{d,i}(N)$ for $i \in \{L, R\}$, mutual best responses are given by $(\mu_L, \mu_R) = (\alpha_{d,L}^\eta(Y), \alpha_{d,R}^\eta(Y))$. As $\eta \to \infty$, we have $(\alpha_{d,L}^\eta(Y), \alpha_{d,R}^\eta(Y)) \to (0, 0)$

- **Case 2:** If $f_d(t_Y) > f_d(t_L), 1 - f_d(t_L)$, then $(\mu_L, \mu_R) = (1 - \alpha_{d,L}(N), 1 - \alpha_{d,R}(N))$. Equation 7 is strictly increasing in μ_L and strictly decreasing in μ_R when $f_d(t_Y) > f_d(t_L), 1 - f_d(t_L)$. Given the constraint of the perturbed game that $\alpha_{d,i}(Y) < \mu_i < 1 - \alpha_{d,i}(N)$ for $i \in \{L, R\}$, mutual best responses are given by $(\mu_L, \mu_R) = (1 - \alpha_{d,L}^\eta(N), 1 - \alpha_{d,R}^\eta(N))$. As $\eta \to \infty$, we have $(1 - \alpha_{d,L}^\eta(N), 1 - \alpha_{d,R}^\eta(N)) \to (1, 1)$

- **Cases 3a, 3b, 4a & 4b:** The left candidate chooses μ_L to maximise Equation 7, while the right candidate chooses μ_R to minimise it. As can be seen from Equation 26, the restriction on strategy space does not affect $\tilde{\mu}_L$. The mutual best responses $(\tilde{\mu}_L, \tilde{\mu}_R)$ found in Step 2 above remain optimal in any perturbed game. Furthermore, as $(\tilde{\mu}_L, \tilde{\mu}_R)$ are completely mixed strategies, they are feasible in any perturbed game.

Step 5: *As $k \to \infty$, the mixed strategy equilibrium identified in Table 2 converge to pure strategy equilibria in which the advantaged candidate chooses the majority-preferred reform position and the disadvantaged candidate chooses the minority reform position.*

Next, I show that in cases 3a, 3b, 4a, 4b the mixed strategy probabilities converge to degenerate probabilities as $k \to \infty$.

Let $\nu_d[(a_{LN}, a_{RN})]$ represent the expected vote share of the left candidate if voters vote on the left-right dimension and let $\nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]$ represent the expected vote share of the left candidate if he is pro-reform while the right candidate is anti-reform and voters vote on the reform dimension. $\nu_d[(a_{LN}, a_{RY}), piv_d^1 = piv(NY)] = 1 - \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]$ is similarly defined.

From Appendix A we know that if $\nu_d[(a_{LN}, a_{RN})] > 0.5$ then $mag[Lwin|(a_{LN}, a_{RN})] = 0$ and $mag[Rwin|(a_{LN}, a_{RN})] = 2\sqrt{\nu_d[(a_{LN}, a_{RN})](1 - \nu_d[(a_{LN}, a_{RN})])} - 1$. Similarly if $\nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$ then $mag[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] = 0$ and $mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] = 2\sqrt{\nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)](1 - \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)])} - 1$.

Furthermore, if $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$ then we have $mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Rwin|(a_{LN}, a_{RN})]$. From the corollary to the magnitude theorem we know that if $mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Rwin|(a_{LN}, a_{RN})]$ then $\lim_{k\to\infty} \frac{Pr[Rwin|(a_{LN}, a_{RN})]}{Pr[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]} = 0$. We can now examine that happens to $(\tilde{\mu}_L, \tilde{\mu}_R)$ as $k \to \infty$ in cases 3a, 3b, 4a, 4b.

- **Case 3a:** If $f_d(t_L) > f_d(t_N) > 0.5$, voter preferences and voter strategies given in Table 1 imply $\nu_d[(a_{LN}, a_{RN})] > 1 - \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$. This implies $mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] = mag[Lwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can re-write

$$\tilde{\mu}_L \equiv \frac{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]}{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LN}, a_{RN})]}$$

  and divide top and bottom by $Pr[Lwin|(a_{LN}, a_{RN})]$ to get

$$\tilde{\mu}_L \equiv \frac{\frac{1 - Pr[Lwin|(a_{LN}, a_{RN})]}{Pr[Lwin|(a_{LN}, a_{RN})]} - \frac{Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]}{Pr[Lwin|(a_{LN}, a_{RN})]}}{\frac{1 - Pr[Lwin|(a_{LN}, a_{RN})]}{Pr[Lwin|(a_{LN}, a_{RN})]} - 1}$$

As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty}(\tilde{\mu}_L, \tilde{\mu}_R) = (0, 1)$$

- **Case 3b:** If $f_d(t_R) > f_d(t_Y) > 0.5$, voter preferences and voter strategies given in Table 1 imply $1 - \nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$. This implies $mag[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] = mag[Rwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Lwin|(a_{LN}, a_{RN})] > -1$.

  We can re-write

$$\tilde{\mu}_L \equiv \frac{1 - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] - Pr[Lwin|(a_{LN}, a_{RN})]}{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LN}, a_{RN})]}$$

  and divide top and bottom by $1 - Pr[Lwin|(a_{LN}, a_{RN})]$ to get

$$\tilde{\mu}_L \equiv \frac{\frac{1-Pr[Lwin|(a_{LY},a_{RN}),piv_d^1=piv(NY)]}{1-Pr[Lwin|(a_{LN},a_{RN})]} - \frac{Pr[Lwin|(a_{LN},a_{RN})]}{1-Pr[Lwin|(a_{LN},a_{RN})]}}{1 - \frac{Pr[Lwin|(a_{LN},a_{RN})]}{1-Pr[Lwin|(a_{LN},a_{RN})]}}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty}(\tilde{\mu}_L, \tilde{\mu}_R) = (0, 1)$$

- **Case 4a:** If $f_d(t_L) > f_d(t_Y) > 0.5$, voter preferences and voter strategies given in Table 1 imply $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$. This implies $mag[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] = mag[Lwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can re-write

$$\tilde{\mu}_L \equiv \frac{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]}{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LN}, a_{RN})]}$$

  and divide top and bottom by $Pr[Lwin|(a_{LN}, a_{RN})]$ to get

$$\tilde{\mu}_L \equiv \frac{\frac{1-Pr[Lwin|(a_{LY},a_{RN}),piv_d^1=piv(NY)]}{Pr[Lwin|(a_{LN},a_{RN})]} - 1}{\frac{1-Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RN})]} - 1}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty} (\tilde{\mu}_L, \tilde{\mu}_R) = (1, 0)$$

- **Case 4b:** If $f_d(t_R) > f_d(t_N) > 0.5$, voter preferences and voter strategies given in Table 1 imply $1 - \nu_d[(a_{LN}, a_{RN})] > 1 - \nu_d[(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > 0.5$. This implies $mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] = mag[Rwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Lwin|(a_{LN}, a_{RN})] > -1$.

  We can re-write

$$\tilde{\mu}_L \equiv \frac{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]}{1 - Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LN}, a_{RN})]}$$

  and divide top and bottom by $1 - Pr[Lwin|(a_{LN}, a_{RN})]$ to get

$$\tilde{\mu}_L \equiv \frac{1 - \frac{Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)]}{1 - Pr[Lwin|(a_{LN}, a_{RN})]}}{1 - \frac{Pr[Lwin|(a_{LN}, a_{RN})]}{1 - Pr[Lwin|(a_{LN}, a_{RN})]}}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty} (\tilde{\mu}_L, \tilde{\mu}_R) = (1, 0)$$

□

Proof of Proposition 3

Proof. **Step 1:** For $k > \bar{k}$, if $f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$ or $f_d(t_L), 1 - f_d(t_L) > f_d(t_N)$, the advantaged candidate is the expected winner.

For $k > \bar{k}$, if $f_d(t_N) > f_d(t_L) > 1 - f_d(t_L)$, we have $(\mu_L^*, \mu_R^*) = (0, 0)$ from Table 2. Therefore, by Table 1, voters vote based on the left-right dimension alone. This gives $\nu_d = f_d(t_L) > 0.5$, meaning advantaged candidate is the expected winner. An identical argument holds for $f_d(t_N) > 1 - f_d(t_L) > f_d(t_L)$ as well as $f_d(t_L), 1 - f_d(t_L) > f_d(t_N)$.

Step 2: For $k > \bar{k}$, and $f_d(t_N) > f_d(t_L), 1 - f_d(t_L)$ or $f_d(t_L), 1 - f_d(t_L) > f_d(t_N)$, we have $Pr[Lwin] \to 1$ if $f_d(t_L) > 0.5$ and $Pr[Lwin] \to 0$ if $f_d(t_L) < 0.5$ as $k \to \infty$.

For $k > \bar{k}$ and $f_d(t_N) > f_d(t_L) > 1 - f_d(t_L)$ we have $\nu_d > 0.5$ and thus $mag[Lwin|(a_{LN}, a_{RN})] = 0 > mag[Rwin|(a_{LN}, a_{RN})] = 2\sqrt{\nu_d(1-\nu_d)} - 1$. By the corollary to the magnitude theorem this gives $\frac{Pr[Rwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RN})]} \to 0$ as $k \to \infty$. Given that $Pr[Rwin|(a_{LN}, a_{RN})] = 1 - Pr[Lwin|a_{LN}, a_{RN})]$, this implies $Pr[Lwin|a_{LN}, a_{RN})] \to 1$ as $k \to \infty$. An identical argument holds for $f_d(t_N) > 1 - f_d(t_L) > f_d(t_L)$ as well as $f_d(t_L), 1 - f_d(t_L) > f_d(t_N)$.

Step 3: For $k > \bar{k}$, if $f_d(t_L) > f_d(t_N), 1 - f_d(t_N)$ or $f_d(t_L) < f_d(t_N), 1 - f_d(t_N)$, the probability that the advantaged candidate is the expected winner goes to 1 as $k \to \infty$.

For $k > \bar{k}$, if $f_d(t_L) > f_d(t_N) > 0.5$, we have $(\mu_L^*, \mu_R^*) = (\tilde{\mu}_L, \tilde{\mu}_R)$ from Table 2. The advantaged candidate is not the expected winner if $\nu_d < 0.5$. This occurs only if (a_{LY}, a_{RN}) and $piv_d^1 = piv(NY)$. The realised platforms are (a_{LY}, a_{RN}) with probability $\tilde{\mu}_L(1 - \tilde{\mu}_R) = \tilde{\mu}_L^2$. From Case 3a in the proof of Proposition 2, we have that $\tilde{\mu}_L \to 0$ as $k \to \infty$. Therefore $Pr[\nu_d < 0.5] \to 0$. An identical argument holds for the case of $0.5 > f_d(t_N) > f_d(t_L)$ as well as $f_d(t_L) > 1 - f_d(t_N) > 0.5$ and $0.5 > 1 - f_d(t_N) > f_d(t_L)$.

Step 4: For $k > \bar{k}$ and $f_d(t_L) > f_d(t_N), 1 - f_d(t_N)$ or $f_d(t_L) < f_d(t_N), 1 - f_d(t_N)$, we have $Pr[Lwin] \to 1$ if $f_d(t_L) > 0.5$ and $Pr[Lwin] \to 0$ if $f_d(t_L) < 0.5$ as $k \to \infty$.

Take $f_d(t_L) > f_d(t_N) > 0.5$ and examine the probability of the left candidate winning in district d, given by Equation 6. Clearly $Pr[Lwin|piv_d^1 = piv(LR)] = Pr[Lwin|(a_{LN}, a_{RN})]$. From Step 2, if $f_d(t_L) > 0.5$, then $Pr[Lwin|(a_{LN}, a_{RN})] \to 1$ as $k \to \infty$. Therefore, a sufficient condition for $Pr[Lwin] \to 1$ as $k \to \infty$ is that Equation 7 converges to 1. From Case 3a in the proof of proposition 2 we have that $\lim_{k\to\infty}(\tilde{\mu}_L, \tilde{\mu}_R) = (0, 1)$. Substituting these values into Equation 7 we get

$$Pr[Lwin|piv_d^1 = piv(NY), \tilde{\mu}_L, \tilde{\mu}_R] = (1 - Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)])$$

From $f_d(t_L) > f_d(t_N) > 0.5$ we have $\nu_d < 0.5$ when (a_{LY}, a_{RN}), $piv_d^1 = piv(NY)$ and $k > \bar{k}$. This implies $mag[Rwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] > mag[Lwin|(a_{LN}, a_{RN}), piv_d^1 = piv(NY)]$ and, by the corollary to the magnitude theorem, $\frac{Pr[Lwin|(a_{LY},a_{RN}),piv_d^1=piv(NY)}{Pr[Rwin|(a_{LY},a_{RN}),piv_d^1=piv(NY)]} \to 0$ as $k \to \infty$. This implies $Pr[Lwin|(a_{LY}, a_{RN}), piv_d^1 = piv(NY)] \to 0$ as $k \to \infty$ and therefore $Pr[Lwin] \to 1$ as $k \to \infty$. An identical argument holds for the case of $0.5 > f_d(t_N) > f_d(t_L)$ as well as $f_d(t_L) > 1 - f_d(t_N) > 0.5$ and $0.5 > 1 - f_d(t_N) > f_d(t_L)$.

□

Proof of Proposition 4

Proof. I prove the case of z_{LN} as the other cases as identical. Let $\mathbf{f}$ be such that $D_{LN}+D_{LY} > \frac{D+1}{2}$ and $D_{LN} + D_{RN} > \frac{D+1}{2}$.

Step 1: $Pr[z = z_{LN}] \to 1$ *as* $k \to \infty$.

By Proposition 2, as $k \to \infty$ the advantaged candidate in each district chooses the majority-preferred reform position with probability going to one. By Proposition 3, as $k \to \infty$ the advantaged candidate in each district wins with probability going to one. Therefore, $Pr[D_j = s_j] \to 1$ for all $j \in \{LN, LY, RN, RY\}$ as $k \to \infty$. It follows form majority voting in the legislature that $Pr[z = z_{LN}] \to 1$ as $k \to \infty$.

Step 2: $Pr[t^{LR}_{\frac{D+1}{2}} \in \{t_L\}] \to 1$ *and* $Pr[t^{NY}_{\frac{D+1}{2}} \in \{t_N\}] \to 1$ *as* $k \to \infty$.

$D_{LN} + D_{LY} > \frac{D+1}{2}$ and $D_{LN} + D_{RN} > \frac{D+1}{2}$ implies $\tilde{t}^{LR}_{\frac{D+1}{2}} \in \{t_L\}$ and $\tilde{t}^{NY}_{\frac{D+1}{2}} \in \{t_N\}$. By the Law of Large Numbers, we know that $(t^{LR}_{\frac{D+1}{2}}, t^{NY}_{\frac{D+1}{2}})$ converges to $(\tilde{t}^{LR}_{\frac{D+1}{2}}, \tilde{t}^{NY}_{\frac{D+1}{2}})$ as $k \to \infty$.

□

Proof of Proposition 5

Proof. I prove the proposition for z_{LN} as the proof is equivalent for all other policies. **Step 1:** For $k > \bar{k}$, if a Condorcet winner policy exists in a district, it wins with probability $\to 1$ as $k \to \infty$.

In a given district, z_{LN} is a Condorcet winner if $f_d(t_L), f_d(t_N), f_d(\hat{t}_{LY}) > 0.5$ holds. From Proposition 3 we have that if $f_d(t_L), f_d(t_N) > 0.5$, with probability $\to 1$ the winner is the left candidate with platform a_{LN} as $k \to \infty$. Therefore, if z_{LN} is a Condorcet winner in district d, a candidate with platform a_{LN} wins the seat with probability $\to 1$ as $k \to \infty$.

Step 2: For $k > \bar{k}$, if a Legislative Condorcet winner exists, it is the implemented policy z with probability $\to 1$ as $k \to \infty$.

Suppose z_{LN} is a Legislative Condorcet winner. It must be that (i) $f_d(t_L) > 0.5$ in at least $\frac{D+1}{2}$ districts, (ii) $f_d(t_N) > 0.5$ in at least $\frac{D+1}{2}$ districts, and (iii) $f_d(\hat{t}_{LY}) > 0.5$ in at least $\frac{D+1}{2}$ districts. This means $\tilde{t}^{LR}_{\frac{D+1}{2}} \in \{t_L\}$ and $\tilde{t}^{NY}_{\frac{D+1}{2}} \in \{t_N\}$. By the Law of Large Numbers, $(t^{LR}_{\frac{D+1}{2}}, t^{NY}_{\frac{D+1}{2}})$ converges to $(\tilde{t}^{LR}_{\frac{D+1}{2}}, \tilde{t}^{NY}_{\frac{D+1}{2}})$ as $k \to \infty$. Therefore is must be that $Pr[t^{LR}_{\frac{D+1}{2}} \in \{t_L\}] \to 1$ and $Pr[t^{NY}_{\frac{D+1}{2}} \in \{t_N\}] \to 1$ as $k \to \infty$. Then, by Proposition 4, the implemented policy is z_{LN} with probability $\to 1$ as $k \to \infty$.

□

Proof of Proposition 6

Proof. By adding $\tilde{t}^h_{med} = \tilde{t}^h_{\frac{D+1}{2}} \forall h \in \{LR, NY\}$ to Proposition 4, we know that as $k \to \infty$ then with probability going to one, the implemented policy in a parliamentary system is that which is preferred by $\tilde{t}^h_{med}$ on dimension $h \in \{LR, NY\}$. All that remains to show is that utility of $\tilde{t}^h_{med}$ under a presidential system is the same for some voter distributions but strictly lower for the rest. As a presidential election is simply a national single-district election, the mapping from preferences to policy outcomes is given byTable 4.

- **Case 1:** Suppose $f_d(t_L) > 0.5$ and $f_d(t_N) > 0.5$ so that $\tilde{t}^{LR}_{med} \in t_L$, $\tilde{t}^{NY}_{med} \in t_N$. Depending of the relative share of voter types, a single district will be in one of the following equilibria 1, 2, 3a, 4a, 4b, 4c in Table 4. If $f_d(\hat{t}_{LN}) > 0.5$, the implemented policy is z_{LN} only in equilibria 1 and 4c. If $f_d(\hat{t}_{LN}) < 0.5$ the implemented policy is z_{LN} only in equilibrium 1. In all other equilibria - 2, 3a, 4a, 4b, the expected implemented policy does not maximise the utility of $\tilde{t}^h_{med}$.

- **Case 2:** Suppose $f_d(t_L) > 0.5$ and $f_d(t_N) < 0.5$ so that $\tilde{t}^{LR}_{med} \in t_L$, $\tilde{t}^{NY}_{med} \in t_Y$. Depending of the relative share of voter types, a single district will be in one of the following equilibria 1, 2, 3a, 4a, 4b, 4c in Table 4. If $f_d(\hat{t}_{LY}) > 0.5$, the implemented policy is z_{LY} only in equilibria 2 and 4b. If $f_d(\hat{t}_{LY}) < 0.5$ the implemented policy is z_{LY} only in equilibrium 2. In all other equilibria - 1, 3a, 4a, 4c the expected implemented policy does not maximise the utility of $\tilde{t}^h_{med}$.

- **Case 3:** Suppose $f_d(t_L) < 0.5$ and $f_d(t_N) > 0.5$ so that $\tilde{t}^{LR}_{med} \in t_R$, $\tilde{t}^{NY}_{med} \in t_N$. Depending of the relative share of voter types, a single district will be in one of the following equilibria 1, 2, 3a, 3b, 3c, 4a in Table 4. If $f_d(\hat{t}_{LN}) > 0.5$, the implemented policy is z_{RN} only in equilibria 1 and 3c. If $f_d(\hat{t}_{LN}) < 0.5$ the implemented policy is z_{RN} only in equilibrium 1. In all other equilibria - 2, 3a, 3b, 4a, the expected implemented policy does not maximise the utility of $\tilde{t}^h_{med}$.

- **Case 4:** Suppose $f_d(t_L) < 0.5$ and $f_d(t_N) < 0.5$ so that $\tilde{t}^{LR}_{med} \in t_R$, $\tilde{t}^{NY}_{med} \in t_Y$. Depending of the relative share of voter types, a single district will be in one of the following equilibria 1, 2, 3a, 3b, 3c, 4a in Table 4. If $f_d(\hat{t}_{LY}) > 0.5$, the implemented policy is z_{RY} only in equilibria 2 and 3b. If $f_d(\hat{t}_{LY}) < 0.5$ the implemented policy is z_{RY} only in equilibrium 2. In all other equilibria - 1, 3a, 3c, 4a, the expected implemented policy does not maximise the utility of $\tilde{t}^h_{med}$.

□

Proof of Proposition 7

Proof. The fact the polarisation will have no effect on voter behaviour, candidate platforms or implemented policies in legislative elections follows directly from Corollary 2 to Proposition 1.

To show the result for single-district elections, take a f_d^1 such that $f_d^1(t) > 0 \forall t \in T$ and suppose without loss of generality that $f_d(t_L), f_d(t_N) > 0.5$ so that $\tilde{t}_{med}^{LR} = t_L$ and $\tilde{t}_{med}^{NY} = t_N$. Without any further restriction on f_d the equilibrium may be any one of 6 equilibrium cases laid out in Table 4 : {1, 2, 3a, 4a, 4b, 4c}. Which case is in fact the equilibrium depends on the relative sizes of $f_d(\hat{t}_{LN}), f_d(t_L)$ and $f_d(\hat{t}_{LY})$. For the new distribution f_d^2 we must have $f_d^2(t_L) = f_d^1(t_L) = f_d(t_L) > 0.5$ as polarisation does not affect the total share of t_L voters. I will proceed by showing that whichever of the 6 equilibria occurs under f^1, a more polarised distribution f^2 can change the equilibrium to any one of the other equilibria.

- **If $f_d^1(\hat{t}_{LN}) > f_d^1(t_L)$, then there exists $f_d^2(\hat{t}_{LN}) < f_d^2(t_L)$.** If $f_d^1(\hat{t}_{LN}) > f_d^1(t_L)$, it must be that $f_d^1(t_{NR}) > f_d^1(t_{YL}) > 0$ where the last inequality is by assumption that $f_d^1(t) > 0 \forall t \in T$. There exists a f^2 such that $f_d^2(t_{NR}) \in (0, f_d^1(t_{NR}))$. Therefore, there exists a f^2 such that $f_d^2(t_{NR}) < f_d^1(t_{YL}) = f_d^2(t_{YL})$, which in turn gives $f_d^2(\hat{t}_{LN}) < f_d^2(t_L)$.

- **If $f_d^1(\hat{t}_{LN}) < f_d^1(t_L)$, then there exists $f_d^2(\hat{t}_{LN}) > f_d^2(t_L)$.** If $f_d^1(\hat{t}_{LN}) < f_d^1(t_L)$, it must be that $f_d^1(t_{YL}) > f_d^1(t_{NR}) > 0$ where the last inequality is by assumption that $f_d^1(t) > 0 \forall t \in T$. There exists a f^2 such that $f_d^2(t_{YL}) \in (0, f_d^1(t_{YL}))$. Therefore, there exists a f^2 such that $f_d^2(t_{YL}) < f_d^1(t_{NR}) = f_d^2(t_{NR})$, which in turn gives $f_d^2(\hat{t}_{LN}) > f_d^2(t_L)$.

- **If $f_d^1(\hat{t}_{LY}) > f_d^1(t_L)$, then there exists $f_d^2(\hat{t}_{LY}) < f_d^2(t_L)$.** If $f_d^1(\hat{t}_{LY}) > f_d^1(t_L)$, it must be that $f_d^1(t_{YR}) > f_d^1(t_{NL}) > 0$ where the last inequality is by assumption that $f_d^1(t) > 0 \forall t \in T$. There exists a f^2 such that $f_d^2(t_{YR}) \in (0, f_d^1(t_{YR}))$. Therefore, there exists a f^2 such that $f_d^2(t_{YR}) < f_d^1(t_{NL}) = f_d^2(t_{NL})$, which in turn gives $f_d^2(\hat{t}_{LY}) < f_d^2(t_L)$.

- **If $f_d^1(\hat{t}_{LY}) < f_d^1(t_L)$, then there exists $f_d^2(\hat{t}_{LY}) > f_d^2(t_L)$.** If $f_d^1(\hat{t}_{LY}) < f_d^1(t_L)$, it must be that $f_d^1(t_{NL}) > f_d^1(t_{YR}) > 0$ where the last inequality is by assumption that $f_d^1(t) > 0 \forall t \in T$. There exists a f^2 such that $f_d^2(t_{NL}) \in (0, f_d^1(t_{NL}))$. Therefore, there exists a f^2 such that $f_d^2(t_{NL}) < f_d^1(t_{YR}) = f_d^2(t_{YR})$, which in turn gives $f_d^2(\hat{t}_{LY}) > f_d^2(t_L)$.

Notice that each of the four cases above we change the distribution of a different voter type. Any change to $f_d(t_{YL})$ or $f_d(t_{NR})$ has no effect on the ordering of $f_d(\hat{t}_{LY}), f_d(t_L)$. Any change to $f_d(t_{YR})$ or $f_d(t_{NL})$ has no effect on the ordering of $f_d(\hat{t}_{LN}), f_d(t_L)$. Regardless of the initial ordering of $f_d^1(\hat{t}_{LY}), f_d^1(\hat{t}_{LN}), f_d^1(t_L)$, and thus any initial equilibrium, any re-ordering can be attained by choosing the appropriate f_d^2. That is, any for any equilibrium in the set {1, 2, 3a, 4a, 4b, 4c} in Table 4, a more polarised distribution can move us to any other equilibrium in the set.

The proof for the case of $f_d(t_L), f_d(t_Y) > 0.5$ is identical, while in the case of $f_d(t_R) > 0.5$ the relevant equilibria are {1, 2, 3a, 3b, 3c, 4a}. Using the same approach as above, one can show that an increase in polarisation can move a single-district election from any one equilibrium to any other.

□

Proof of Proposition 8

Proof. The proof proceeds in two steps. First, I characterise the equilibrium strategies for a fixed k. Then, I compute what happens to these equilibrium strategies as $k \to \infty$.

Step 1: Equilibrium strategies.

The left candidate will choose μ_L to maximise Equation 19, while the right candidate will choose μ_R to minimise it. This yields the following best response correspondences:

$$BR_L(\mu_R) = \begin{cases} 1 & \text{if } \mu_R > \bar{\mu}_R \\ [0,1] & \text{if } \mu_R = \bar{\mu}_R \\ 0 & \text{if } \mu_R < \bar{\mu}_R \end{cases} \tag{27}$$

$$BR_R(\mu_L) = \begin{cases} 1 & \text{if } \mu_L < \bar{\mu}_L \\ [0,1] & \text{if } \mu_L = \bar{\mu}_L \\ 0 & \text{if } \mu_L > \bar{\mu}_L \end{cases} \tag{28}$$

where $\bar{\mu}_L$ and $\bar{\mu}_R$ are given by:

$$\bar{\mu}_L \equiv \frac{Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN})]}{2Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN})] - Pr[Lwin|(a_{LN}, a_{RY})]} \tag{29}$$

$$\bar{\mu}_R \equiv 1 - \bar{\mu}_L$$

We can also write Equation 29 as

$$\bar{\mu}_L \equiv \frac{Pr[Rwin|(a_{LN}, a_{RN})] - Pr[Rwin|(a_{LY}, a_{RN})]}{2Pr[Rwin|(a_{LN}, a_{RN})] - Pr[Rwin|(a_{LY}, a_{RN})] - Pr[Rwin|(a_{LN}, a_{RY})]} \tag{30}$$

We can divide districts into eight possible preference cases. We analyse each case in turn:

- **Case 1:** If $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$, voter strategies given in Table 3 imply that $Pr[Lwin|(a_{LN}, a_{RY})] > Pr[Lwin|(a_{LN}, a_{RN})] > Pr[Lwin|(a_{LY}, a_{RN})]$. Examining Equation 19 we have a corner solution - both are playing a best response at

$(\mu_L, \mu_R) = (0, 0)$.

- **Case 2:** If $f_d(\hat{t}_{LY}) > f_d(t_L) > f_d(\hat{t}_{LN})$, voter strategies given in Table 3 imply that $Pr[Lwin|(a_{LY}, a_{RN})] > Pr[Lwin|(a_{LN}, a_{RN})] > Pr[Lwin|(a_{LN}, a_{RY})]$. Examining Equation 19 we have a corner solution - both are playing a best response at $(\mu_L, \mu_R) = (1, 1)$.

- **Case 3:** If $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L)$, then $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$. There are 3 sub-cases to consider here:

  - **Case 3a:** With $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L), 0.5$ and voter strategies given in Table 3 it must be that $Pr[Lwin|(a_{LY}, a_{RN})], Pr[Lwin|(a_{LN}, a_{RY})] > Pr[Lwin|(a_{LN}, a_{RN})], 0.5$. This in turn means that $\bar{\mu}_L \in (0, 1)$ and $\bar{\mu}_R \in (0, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 3b:** With $0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$ and voter strategies given in Table 3 it must be that $Pr[Lwin|(a_{LN}, a_{RY})], 0.5 > Pr[Lwin|(a_{LY}, a_{RN})] > Pr[Lwin|(a_{LN}, a_{RN})]$. This in turn means that $\bar{\mu}_L \in (0, 0.5)$ and $\bar{\mu}_R \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 3c:** With $0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$ and voter strategies given in Table 3 it must be that $Pr[Lwin|(a_{LY}, a_{RN})], 0.5 > Pr[Lwin|(a_{LN}, a_{RY})] > Pr[Lwin|(a_{LN}, a_{RN})]$. This in turn means that $\bar{\mu}_L \in (0.5, 1)$ and $\bar{\mu}_R \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.

- **Case 4:** If $f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$, then $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$. There are 3 sub-cases to consider here:

  - **Case 4a:** With $0.5, f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$ and voter strategies given in Table 3 it must be that $Pr[Lwin|(a_{LN}, a_{RN})], 0.5 > Pr[Lwin|(a_{LY}, a_{RN})], Pr[Lwin|(a_{LN}, a_{RY})]$. This in turn means that $\bar{\mu}_L \in (0, 1)$ and $\bar{\mu}_R \in (0, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 4b:** With $f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$ and voter strategies given in Table 3 it must be that $Pr[Lwin|(a_{LN}, a_{RN})] > Pr[Lwin|(a_{LN}, a_{RY})] > Pr[Lwin|(a_{LY}, a_{RN})], 0.5$. This in turn means that $\bar{\mu}_L \in (0.5, 1)$ and $\bar{\mu}_R \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 4c:** With $f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$ and voter strategies given in Table 3 it must be that $Pr[Lwin|(a_{LN}, a_{RN})] > Pr[Lwin|(a_{LY}, a_{RN})] > Pr[Lwin|(a_{LN}, a_{RY})], 0.5$. This in turn means that $\bar{\mu}_L \in (0, 0.5)$ and $\bar{\mu}_R \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.

Step 2: Equilibrium strategies as $k \to \infty$.

We can now examine that happens to $(\bar{\mu}_L, \bar{\mu}_R)$ as $k \to \infty$ in cases 3 and 4.

- **Case 3a:** With $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L), 0.5$ and voter strategies given in Table 3 it must be that $\nu_d[(a_{LY}, a_{RN})], \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LN}, a_{RN})], 0.5$. This implies $mag[Lwin|(a_{LY}, a_{RN})] = mag[Lwin|(a_{LN}, a_{RY})] = 0$ and $mag[Rwin|(a_{LN}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RY})], mag[Rwin|(a_{LY}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation 30 by $Pr[Rwin|(a_{LN}, a_{RN})]$ to get

$$\bar{\mu}_L \equiv \frac{1 - \frac{Pr[Rwin|(a_{LY},a_{RN})]}{Pr[Rwin|(a_{LN},a_{RN})]}}{2 - \frac{Pr[Rwin|(a_{LY},a_{RN})]}{Pr[Rwin|(a_{LN},a_{RN})]} - \frac{Pr[Rwin|(a_{LN},a_{RY})]}{Pr[Rwin|(a_{LN},a_{RN})]}}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k \to \infty} (\bar{\mu}_L, \bar{\mu}_R) = (0.5, 0.5)$$

- **Case 3b:** With $0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$ and voter strategies given in Table 3 it must be that $0.5, \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RN})]$. This implies $mag[Lwin|(a_{LN}, a_{RY})] > mag[Lwin|(a_{LY}, a_{RN})] > mag[Lwin|(a_{LN}, a_{RN})]$.

  We can divide the top and bottom of Equation 29 by $Pr[Lwin|(a_{LN}, a_{RY})]$ to get

$$\bar{\mu}_L \equiv \frac{\frac{Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]} - \frac{Pr[Lwin|(a_{LY},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]}}{2\frac{Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]} - \frac{Pr[Lwin|(a_{LY},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]} - 1}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k \to \infty} (\bar{\mu}_L, \bar{\mu}_R) = (0, 1)$$

- **Case 3c:** With $0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$ and voter strategies given in Table 3 it must be that $0.5, \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LN}, a_{RN})]$. This implies $mag[Lwin|(a_{LY}, a_{RN})] > mag[Lwin|(a_{LN}, a_{RY})] > mag[Lwin|(a_{LN}, a_{RN})]$.

  We can divide the top and bottom of Equation 29 by $Pr[Lwin|(a_{LY}, a_{RN})]$ to get

$$\bar{\mu}_L \equiv \frac{\frac{Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LY},a_{RN})]} - 1}{2\frac{Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LY},a_{RN})]} - 1 - \frac{Pr[Lwin|(a_{LN},a_{RY})]}{Pr[Lwin|(a_{LY},a_{RN})]}}$$

As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty} (\bar{\mu}_L, \bar{\mu}_R) = (1, 0)$$

- **Case 4a:** With $0.5, f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$ and voter strategies given in Table 3 it must be that $0.5, \nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN})], \nu_d[(a_{LN}, a_{RY})]$. This implies $mag[Lwin|(a_{LN}, a_{RN})] > mag[Lwin|(a_{LN}, a_{RY})], mag[Lwin|(a_{LY}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation 29 by $Pr[Lwin|(a_{LN}, a_{RN})]$ to get

$$\bar{\mu}_L \equiv \frac{1 - \frac{Pr[Lwin|(a_{LY}, a_{RN})]}{Pr[Lwin|(a_{LN}, a_{RN})]}}{2 - \frac{Pr[Lwin|(a_{LN}, a_{RY})]}{Pr[Lwin|(a_{LN}, a_{RN})]} - \frac{Pr[Lwin|(a_{LY}, a_{RN})]}{Pr[Lwin|(a_{LN}, a_{RN})]}}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty} (\bar{\mu}_L, \bar{\mu}_R) = (0.5, 0.5)$$

- **Case 4b:** With $f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$ and voter strategies given in Table 3 it must be that $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LY}, a_{RN})], 0.5$. This implies $mag[Rwin|(a_{LY}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RY})] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation 30 by $Pr[Rwin|(a_{LY}, a_{RN})]$ to get

$$\bar{\mu}_L \equiv \frac{\frac{Pr[Rwin|(a_{LN}, a_{RN})]}{Pr[Rwin|(a_{LY}, a_{RN})]} - 1}{2\frac{Pr[Rwin|(a_{LN}, a_{RN})]}{Pr[Rwin|(a_{LY}, a_{RN})]} - 1 - \frac{Pr[Rwin|(a_{LN}, a_{RY})]}{Pr[Rwin|(a_{LY}, a_{RN})]}}$$

  As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k\to\infty} (\bar{\mu}_L, \bar{\mu}_R) = (1, 0)$$

- **Case 4c:** With $f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$ and voter strategies given in Table 3 it must be that $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})], 0.5$. This implies $mag[Rwin|(a_{LN}, a_{RY})] > mag[Rwin|(a_{LY}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

We can divide the top and bottom of Equation 30 by $Pr[Rwin|(a_{LN}, a_{RY})]$ to get

$$\bar{\mu}_L \equiv \frac{\frac{Pr[Rwin|(a_{LN},a_{RN})]}{Pr[Rwin|(a_{LN},a_{RY})]} - \frac{Pr[Rwin|(a_{LY},a_{RN})]}{Pr[Rwin|(a_{LN},a_{RY})]}}{2\frac{Pr[Rwin|(a_{LN},a_{RN})]}{Pr[Rwin|(a_{LN},a_{RY})]} - \frac{Pr[Rwin|(a_{LY},a_{RN})]}{Pr[Rwin|(a_{LN},a_{RY})]} - 1}$$

As in each case the denominator has a larger magnitude than the numerator, each probability ratio goes to zero as $k \to \infty$ and we get

$$\lim_{k \to \infty} (\bar{\mu}_L, \bar{\mu}_R) = (0, 1)$$

$\square$

References

Alesina, A. and Rosenthal, H. (1995). *Partisan Politics, Divided Government, and the Economy*. Cambridge University Press.

— and — (2000). Polarized platforms and moderate policies with checks and balances. *Journal of Public Economics*, **75** (1), 1–20.

Aragonès, E., Castanheira, M. and Giani, M. (2015). Electoral competition through issue selection. *American Journal of Political Science*, **59** (1), 71–90.

Austen-Smith, D. (1984). Two-party competition constituencies. *Mathematical Social Sciences*, **7**, 177–198.

Barber, M. and McCarty, N. (2015). Causes and Consequences of Polarization. In *Political Negotiation: A Handbook*.

Besley, T. and Coate, S. (2008). Issue unbundling via citizens initiatives. *Quarterly Journal of Political Science*, **3** (4), 379–397.

Bhattacharya, S. (2013). Preference Monotonicity and Information Aggregation in Elections. *Econometrica*, **81** (3), 1229–1247.

Bol, D. and Verthé, T. (2019). Strategic Voting Versus Sincere Voting. *Oxford Research Encyclopedia of Politics*, (October), 1–20.

Bonomi, G., Gennaioli, N. and Tabellini, G. (2021). Identity, beliefs, and political conflict. *Quarterly Journal of Economics*, **136** (4), 2371–2411.

Bordignon, M., Nannicini, T. and Tabellini, G. (2016). Moderating Political Extremism : Single Round versus Runoff Elections under Plurality Rule. *American Economic Review*, **106** (8), 2349–2370.

Bouton, L. (2013). A Theory of Strategic Voting in Runoff Elections. *American Economic Review*, **103** (4), 1248–1288.

— and Castanheira, M. (2012). One Person, Many Votes: Divided Majority and Information Aggregation. *Econometrica*, **80** (1), 43–87.

—, Conconi, P., Zanardi, M. and Pino, F. (2020). The Tyranny of the Single Minded : Guns, Environment and Abortion. *Review of Economics and Statistics*.

— and Gratton, G. (2015). Majority Runoff Elections: Strategic Voting and Duverger's Hypothesis. *Theoretical Economics*, **10** (2), 283–314.

Buisseret, P. and Van Weelden, R. (2019). Crashing the Party? Elites, Outsiders, and Elections. *American Journal of Political Science*.

— and Van Weelden, R. (forthcoming). Polarization, valence, and policy competition. *American Economic Review: Insights*.

Callander, S. (2005). Electoral Competition in Heterogeneous Districts. *Journal of Political Economy*, **113** (5), 1116–1145.

Canen, N., Kendall, C. and Trebbi, F. (2020). Unbundling Polarization. *Econometrica*, **88** (3), 1197–1233.

Cantoni, E. and Pons, V. (2022). Does context outweigh individual characteristics in driving voting behavior? evidence from relocations within the united states. *American Economic Review*, **112** (4), 1226–72.

De Sinopoli, F., Meroni, C. and Pimienta, C. (2014). Strategic stability in Poisson games. *Journal of Economic Theory*.

— and Pimienta, C. (2009). Undominated (and) perfect equilibria in Poisson games. *Games and Economic Behavior*, **66** (2), 775–784.

Durand, F., Macé, A. and Nunez, M. (2021). Analysis of Approval Voting in Poisson Games.

Eggers, A. C. and Vivyan, N. (2020). Who Votes More Strategically? *American Political Science Review*.

Eyster, E. and Kittsteiner, T. (2007). Party platforms in electoral competition with heterogeneous constituencies. *Theoretical Economics*, **2** (1), 41–70.

Feddersen, T. and Pesendorfer, W. (1997). Voting Behavior and Information Aggregation in Elections With Private Information. *Econometrica*, **65** (5), 1029–1058.

Fisher, S. D. and Myatt, D. P. (2017). Strategic Voting in Plurality Rule Elections. pp. 1–40.

Fourati, M., Gratton, G. and Grosjean, P. (2019). Render unto Caesar: Taxes, charity, and political Islam. *European Economic Review*, **119**, 114–146.

Gentzkow, M. (2016). Polarization in 2016. *Toulouse Network of Information Technology whitepaper*, pp. 1–23.

Gethin, A., Martínez-Toledano, C. and Piketty, T. (2021). Brahmin Left Versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948-2020. *The Quarterly Journal of Economics*.

Hinich, M. J. and Ordeshook, P. C. (1974). The electoral college: a spatial analysis. *Political Methodology*, **1** (3), 1–29.

Hughes, N. (2016). Voting in Legislative Elections under Plurality Rule. *Journal of Economic Theory*, **166** (November), 51–93.

Kramer, G. (1972). Strategic Voting. *Journal of Mathematical Sociology*, **6**.

Krasa, S. and Polborn, M. (2010). The binary policy model. *Journal of Economic Theory*, **145** (2), 661–688.

— and Polborn, M. K. (2018). Political Competition in Legislative Elections. *American Political Science Review*, **112** (4), 809–825.

Krishna, V. and Morgan, J. (2011). Overcoming Ideological Bias in Elections. *Journal of Political Economy*, **119** (2), 183–211.

Llavador, H. (2006). Electoral platforms, implemented policies, and abstention. *Social Choice and Welfare*, **27** (1), 55–81.

Matakos, K., Troumpounis, O. and Xefteris, D. (2016). Electoral Rule Disproportionality and Platform Polarization. *American Journal of Political Science*, **60** (4), 1026–1043.

McMurray, J. (2013). Aggregating Information by Voting: The Wisdom of the Experts versus the Wisdom of the Masses. *Review of Economic Studies*, **80** (1), 1–50.

Morelli, M. (2004). Party Formation and Policy Outcomes under Different Electoral Systems. *Review of Economic Studies*, **71** (3), 829–853.

Myatt, D. P. (2007). On the Theory of Strategic Voting. *Review of Economic Studies*, **74** (1), 255–281.

— (2017). A Theory of Protest Voting. *The Economic Journal*, (July 2012), 1–40.

Myerson, R. B. (1998). Population uncertainty and Poisson games. *International Journal of Game Theory*, **27** (3), 375–392.

— (2000). Large Poisson Games. *Journal of Economic Theory*, **94** (1), 7–45.

— (2002). Comparison of Scoring Rules in Poisson Voting Games. *Journal of Economic Theory*, **103** (1), 219–251.

— and Weber, R. J. (1993). A Theory of Voting Equilibria. *American Political Science Review*, **87** (1), 102–114.

Okada, A. (1981). On stability of perfect equilibrium points. *International Journal of Game Theory*, **10** (2), 67–73.

Osborne, M. J. (1995). Spatial Models of Political Competition under Plurality Rule - a Survey of Some Explanations of the Number of Candidates and the Positions They Take. *Canadian Journal of Economics*, **28** (2), 261–301.

Patty, J. W., Snyder, J. M. and Ting, M. M. (2009). Two's Company, Three's an Equilibrium: Strategic Voting and Multicandidate Elections. *Quarterly Journal of Political Science*, **4** (3), 251–278.

Persson, T. (2002). Do Political Institutions Shape Economic Policy? *Econometrica*, **70** (3), 883–905.

—, Roland, G. and Tabellini, G. (1997). Separation of powers and political accountability. *Quarterly Journal of Economics*, **112** (November), 1163–1202.

—, — and — (2000). Comparative Politics and Public Finance. *Journal of Political Economy*, **108** (6), 1121–1161.

— and Tabellini, G. (2005). *The economic effects of constitutions*, vol. 54. The MIT Press.

Polborn, M. K. and Snyder, J. M. (2017). Party Polarization in Legislatures with Office-Motivated Candidates. *Quarterly Journal of Economics*, **132** (3), 1509–1550.

Razin, R. (2003). Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates. *Econometrica*, **71** (4), 1083–1119.

Shugart, M. S., Carey, J. M. *et al.* (1992). *Presidents and assemblies: Constitutional design and electoral dynamics*. Cambridge University Press.

Xefteris, D. (2017). Multidimensional electoral competition between differentiated candidates. *Games and Economic Behavior*, **105**, 112–121.