Strategic Voting in Two-Party Legislative Elections

Hughes, Niall

King's Business School, King's College London

30 April 2020

Strategic Voting in Two-Party Legislative Elections

Niall Hughes

King's Business School, King's College London

niall.hughes@kcl.ac.uk

April 2020

Abstract

It is commonly thought that in an election with two parties there can be no strategic voting - voters simply vote for their preferred candidate. In this paper, I show that strategic voting comes to the fore in legislative elections with multiple policy dimensions. In sharp contrast to single-district elections, the intensity of a voter's preference on each dimension is irrelevant for her voting decision. Instead, she votes solely based on the dimension which is most likely to be pivotal in the legislature. Anticipating this behaviour, candidates put forward a different set of policies than they would in a single-district election. For large elections I show that the implemented policy bundle: (a) is uniquely pinned down by voter preferences, (b) is preferred by a majority of districts on each dimension, (c) is a Condorcet winner, if one exists. These properties are not guaranteed in a single-district election. Furthermore, I show that (i) parliamentary systems generate superior policies to presidential systems and (ii) 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 2016 Presidential Election. Should she vote for economically-left, socially-liberal Clinton or economically-right, socially-conservative Trump? In the UK 2019 General Election many traditional working-class voters faced a similar dilemma - 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 Clinton. UK traditional working-class voters who care more about social than economic issues back the Tories, while the others 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 well be that a voter cares a lot more about the economic dimension but, because of how other districts behave, it is more likely that her vote will 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, regardless of how much or little they care about that dimension. In a nutshell, a voter's preference intensity is irrelevant for their 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 voter welfare 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, e.g. US Congress, UK House of Commons, Canadian House of Commons, Indian Lok Sabha. 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 traditional left-right voter divide into a more complex structure where voters care about multiple dimensions and these cut across party allegiances.[1] Thus, while the

[1] See e.g. Buisseret and Van Weelden (2019),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.

difference in voting incentives between single-district and legislative elections may have been purely theoretical in the past, clearly 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 is in the UK, where several tactical voting groups exists. Some, such as Tactical Vote and Turf Out The Tories, want to prevent a Conservative majority while others, such as Best for Britain and Remain United, want to prevent Brexit. Their advice on how to vote often conflicts as their objectives are 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 presidential systems typically choose policy via single-district elections, while parliamentary systems use legislative elections. 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. Two key features of the model are that (i) both of these dimensions are binary, and (ii) candidates have a fixed position on the left-right 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 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 issues are continuous, it is well known that equilibria may not exist.[2] 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. We 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

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

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, Castanheira, and Giani, 2015; Xefteris, 2017).[3] 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 strategic voting. Here, the emphasis has typically been on multi-candidate elections (Myerson, 2002; Myatt, 2007; Patty, Snyder, and Ting, 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.[4] Third, the paper speaks to the vast literature on how political institutions affect policy choices (Persson, 2002; Morelli, 2004; Persson and Tabellini, 2005; Bordignon, Nannicini, and Tabellini, 2016).I show that in majoritarian systems, the type of election affects the congruence between voters and policies: legislative (parliamentary) elections generate superior policies to 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, Troumpounis, and Xefteris (2016); Gentzkow (2016); Canen, Kendall, and Trebbi (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, implications for the real-world, and concludes.

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

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

[4] 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).

anti-reform. Candidates choose their platforms to maximise their probability of being elected in the district. Let $\mu_d = (\mu_L, \mu_R)$ 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. Each legislator is committed to supporting the platform he was elected on. [5]

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:

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

It will often be easier to refer to sets of types t_L, t_R, t_Y, t_N 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.[6] Appendix A summarises several properties of the Poisson model.[7] 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 $f_{d'}$. 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. We say the left candidate is the

[5] 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.

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

[7] 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).

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. I make the following assumption on $\mathbf{f}$.

Assumption 1. *There is at least one district where either $f_d(t_{NR}) > f_d(t_{YL})$ and $f_d(t_{NL}) > f_d(t_{YR})$, or $f_d(t_{NR}) < f_d(t_{YL})$ and $f_d(t_{NL}) < f_d(t_{YR})$ holds.*

This ensures there is a minimal amount a variation in preferences across districts, and is sufficient to guarantee we cannot have every district be a (a_{LN}, a_{RY}) contest or every district be a (a_{LY}, a_{RN}) contest. The details of why this is are Appendix B.[8] If the assumption does not hold, there may be no heterogeneity within parties, in which case voting behaviour in legislative elections and single-district elections coincides.

Let $V \equiv \{v_L, v_R\}$ be the set of actions a voter can take.[9] A voting strategy is $\sigma : T \times D \to \Delta(V)$ where $\sigma_d(t)$ is the probability that a type t voter in district d casts ballot v_L, with $1 - \sigma_d(t)$ being the probability that this type votes v_R. In a Poisson game, all voters of the same type in the same district must follow the same strategy.[10] The expected vote share of the left candidate in the district is

$$\nu_d = \sum_{t \in T} \sigma_d(t) f_d(t) \tag{2}$$

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 $\nu_d < 0.5$.[11] As $k \to \infty$ the probability that the expected winner wins the district goes to one.[12] 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 and Payoffs 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

[8] The preferences in Assumption 1 are those in Case 1 and 2 of Table 4 respectively.

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

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

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

[12] 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.

choice on her district being pivotal.[13] 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. 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.

The event $piv(LR, NY)$ is when the district is pivotal on both dimensions of policy. In this case, depending on the platforms in the district, all four policy outcomes are possible.

Assumption 2. *When deciding how to cast their ballot, voters ignore the event of being jointly pivotal on both dimensions of policy, $piv(LR, NY)$.*

This assumption comes with some loss of generality, which I argue is minimal. 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 $\sum_{d=1}^{\frac{D-1}{2}} 4d$ 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. While the share of double-pivotal events vis a vis single-pivotal events is small, we might worry that, nonetheless, many districts condition on a double-pivotal event. As I show in Appendix C, the restrictions needed for a district to be conditioning on a double-pivotal event are very stringent and unlikely to be generally applicable. Furthermore, Assumption 2 is not required for Lemma 1, Proposition 1, or its first corollary.

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

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

[13]We could additionally have voters condition on their vote being pivotal within their district, but this would not alter any of the results.

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.

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

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.

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

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

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 will be omitted from the definition of equilibrium. Equilibrium in this game consists of a voting equilibrium in stage 3, $\boldsymbol{\sigma}^*$, and a candidate equilibrium in stage 2, $\boldsymbol{\mu}^*$. In a voting equilibrium, each voter type in each district best responds to the strategies of all other voters in all districts. In a candidate equilibrium, candidates best respond to the strategy of their local opponent and the strategies of voters in their district. I restrict attention to *responsive* candidate equilibria. A candidate equilibrium $\boldsymbol{\mu}^*$ is non-responsive if it can be sustained for any distribution of voter preferences $\mathbf{f}$. [14] As we are interested in the properties of large national elections, I will analyse asymptotic equilibria. That is, the limit of the set of equilibria as $k \to \infty$. More specifically, following Bouton and Gratton (2015), I restrict attention to asymptotic strictly perfect equilibria. A sequence of Nash equilibria $\{\boldsymbol{\mu}_k^*, \boldsymbol{\sigma}_k^*\}_{k\to\infty}$ is asymptotically strictly perfect if (i) it admits a limit and (ii) as k grows large, the equilibrium is robust to epsilon changes in the strategies of players. This restriction is to rule out knife-edge cases where the equilibrium depends on multiple events having the same probability of being pivotal.

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.

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. □

[14] There are many non-responsive equilibria in a game such as this. For example, any equilibrium where $\frac{D+3}{2}$ districts have platforms (a_{LN}, a_{RN}) is non-responsive. As no district can ever be pivotal on the social reform dimension, any strategy for a candidate is a best response. The same is true for any equilibrium where $\frac{D+3}{2}$ districts have platforms (a_{LY}, a_{RY}).

A conflicted voter has positive and negative terms in her gain function, so her best response depends on the relative size of these terms. Whether the 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. A voter's type determines the former while the latter depends on the strategies of voters and candidates in the other $D-1$ districts. Therefore, conflicted voters are *strategic voters* - their best response depends on the strategies of other players. An unconflicted voter, instead, can ignore (ii) because each term in her gain function has the same sign.

We have 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 most likely pivotal event. Let piv_d^i denote the i-th most likely pivotal event for district d. A property of the Poisson model is that conditional on being pivotal, the most likely pivotal event piv_d^1 is infinitely more likely to occur than any other pivotal event. This greatly simplifies the best response of a conflicted voter.

Proposition 1. *As $k \to \infty$ the best response for a voter is to vote solely based on the most likely pivotal event, piv_d^1. These best responses are represented in Table 1.*

Proof. See Appendix D. □

<table>
<tr><th>District</th><th>Vote v_L</th><th>Vote v_R</th></tr>
<tr><td>(a_{LN}, a_{RN})
(a_{LY}, a_{RY})
(a_{LN}, a_{RY}) with $piv_d^1(LR)$
(a_{LY}, a_{RN}) with $piv_d^1(LR)$</td><td>$t_{NL}, t_{LN}, t_{YL}, t_{LY}$</td><td>$t_{NR}, t_{RN}, t_{YR}, t_{RY}$</td></tr>
<tr><td>(a_{LN}, a_{RY}) with $piv_d^1(NY)$</td><td>$t_{NL}, t_{LN}, t_{NR}, t_{RN}$</td><td>$t_{YL}, t_{LY}, t_{YR}, t_{RY}$</td></tr>
<tr><td>(a_{LY}, a_{RN}) with $piv_d^1(NY)$</td><td>$t_{YL}, t_{LY}, t_{YR}, t_{RY}$</td><td>$t_{NL}, t_{LN}, t_{NR}, t_{RN}$</td></tr>
</table>

Table 1: Best responses of each voter type

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 their district is pivotal in expectation, voters vote based on that pivotal event. If their district is not pivotal in expectation, then, in theory, voters need to calculate the probability of all the various pivotal events. These various pivot probabilities go to zero as the number of voters goes to infinity, but those with smaller probability go to zero at a faster rate. Therefore, conditional on being pivotal, the probability of being in the pivotal event with the largest probability goes to one as the size of the electorate grows. Because of this, voters can safely ignore all but the most likely pivotal event when making their voting decision. 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 $piv_d^1(NY)$, 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 $piv_d^1(NY)$, 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. *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 as $k \to \infty$ voters only consider one dimension when voting - the dimension most likely to change the legislative majority. With only one dimension to focus on, the interested of t_{ij} and t_{ji} types are aligned.

We might worry that by ignoring all but one dimension of policy, voters may select suboptimal 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. When choosing μ, candidates do not know whether voters will either condition on $piv(LR)$ or $piv(NY)$.[15] If voters condition on

[15] By restricting attention to responsive equilibria, we ensure that both pivotal events have positive probability.

$piv(LR)$, the positions of candidates on the reform dimension is irrelevant for voters - any choice of μ is a best response. If voters condition on $piv(NY)$, the expected probability of winning for the left candidate is given by

$$\begin{aligned} Pr[Lwin|piv^1(NY),\mu_L,\mu_R] &= (\mu_L\mu_R+(1-\mu_L)(1-\mu_R))Pr[Lwin|(a_{LN},a_{RN})] \\ &+ \mu_L(1-\mu_R)Pr[Lwin|(a_{LY},a_{RN}),piv^1(NY)] \\ &+ (1-\mu_L)(\mu_R)(1-Pr[Lwin|(a_{LY},a_{RN}),piv^1(NY)]) \end{aligned} \tag{6}$$

As only this case matters, the left candidate will choose μ_L to maximise this expression, while the right candidate will choose μ_R to minimise it. The following proposition characterises all of the equilibria in this candidate competition stage.

Proposition 2. *For any given f_d, the candidate equilibrium is unique. 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 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 $0.5 > f_d(t_N) > f_d(t_L)$	$\mu_L^*\to 0$	$\mu_R^*\to 1$	$\to(a_{LN},a_{RY})$
$f_d(t_L) > 1-f_d(t_N) > 0.5$ or $0.5 > 1-f_d(t_N) > f_d(t_L)$	$\mu_L^*\to 1$	$\mu_R^*\to 0$	$\to(a_{LY},a_{RN})$

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.

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 pinned down by voter preferences. 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 expected winner in each district and the expected implemented policy are unique. The following result follows directly from Proposition 2.

Proposition 3. *As $k \to \infty$ the advantaged candidate in a district is the expected winner with probability $\to 1$.*

Proof. See Appendix D. □

In other words, the probability of the dis-advantaged candidate winning the district goes to zero in large elections. A necessary condition for the dis-advantaged candidate to win in a large election that candidates are playing the matching pennies game and the dis-advantaged candidate happens to choose the majority-preferred reform position while the advantaged candidate happens to choose the minority-preferred position. From Proposition 2, we know that this occurs with probability going to zero as $k \to \infty$. Therefore, with probability going to one, the winner in district d will 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. A small detour is first required to do this. 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, maximise overall voter welfare. In a multi-dimensional setting such as ours, 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 in district d on the reform dimension. This voter is anti-reform if $f_d(t_N) > 0.5$. We can 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.

Proposition 4. *As $k \to \infty$ the unique expected policy $E(z)$ is that which is preferred by $\tilde{t}_{\frac{D+1}{2}}^{LR}$ on the left-right dimension and by $\tilde{t}_{\frac{D+1}{2}}^{NY}$ on the reform dimension.*

Proof. See Appendix D. □

The focus on expected policies and voters above is because of the random nature of the model. It is always possible for the realised population of voters to differ from the expected population sufficiently that $E(z) \neq z$. However, as $k \rightarrow \infty$, this probability goes to zero. The logic of the proposition is quite straightforward: We know that the expected winner in each district runs on a policy which is majority-preferred on both dimensions, and wins the seat with probability $\rightarrow 1$ as $k \rightarrow \infty$. We also know that the implemented policy, z, requires a majority of seats on each dimension. 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, we are guaranteed a unique expected 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.

We can go further in exploring voter welfare in our setting, but, we 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. We can now state the following result:

Proposition 5. *If a Condorcet Winner policy exists in a district, it is the expected winner. If a Legislative Condorcet winner exists, it is the expected implemented policy $E(z)$.*

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. 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 candidates.[16] 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. 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.

Strategic voting in our setup has several features which distinguish it from strategic voting in multi-candidate elections (Myerson and Weber, 1993; Patty, Snyder, and Ting, 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 our 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 frontrunners 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

[16] 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.

larger than in our 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![17] In our setting, we get the reverse: strategic voting leads to a unique policy outcome, while this is not guaranteed with naive voters. Moreover, this unique policy is the most representative policy possible. Strategic voting in our setting is not just individually rational - it is welfare maximising. A popular argument against strategic voting is that election outcomes will not accurately reflect aggregate voter preferences[18]. My analysis shows not only that this is false, but that the opposite is true - naive voting does not accurately reflect voter preferences.

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 that 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 Equation 15 we know that preference intensity is irrelevant - types t_{ij} and t_{ji} always vote the same way. They 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

[17]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.

[18]See for example: https://www.electoral-reform.org.uk/what-is-tactical-voting-and-why-is-it-bad-for-democracy/ and https://www.theguardian.com/commentisfree/2015/may/06/wrong-vote-tactically-general-election.

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 that is not majority-preferred on each dimension; and (iii) a Condorcet winner policy is not chosen by either candidate.[19] None of these hold true for the same district in a legislative election.

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 use non-degenerate mixed strategies. This means that candidate's 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.

In terms of representativeness, 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 is 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

[19] 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.

Coate (2008) and finds empirical support in Bouton, Conconi, Zanardi, and Pino (2020). Besley and Coate (2008) show that citizens initiative 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.

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.[20] In these models the differences between the two systems stem from the executive being subject to a confidence vote in a parliamentary system but not in a presidential system. Here, I focus on another less-studied difference between the two systems - the executive is chosen by different types of elections in each case. A president is typically selected by a direct national election while a prime minister is chosen by a majority within the legislature.[21] In other words, a president is elected in a single-district election while a prime minister is elected by a legislative election. 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, we 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 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 $\tilde{t}^{NY}_{\frac{D+1}{2}} \in \{t_N, t_Y\}$. [22] 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

[20] For an overview see Persson and Tabellini (2005).

[21] A notable exception to this is the US Presidential election and its Electoral College system.

[22] $\tilde{t}^{LR}_{med} \in \{t_L, t_R\}$ 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.

median voter on the left-right dimension is a Democrat but the the median voter in the median district (Illinois's 13th District) is a Republican.[23] 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. Cases such as this 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 voters are better off under a parliamentary system than a presidential one.

Proposition 6. *For any distribution of voter preferences and any distribution of voters across districts such that* $\tilde{t}^j_{med} = \tilde{t}^j_{\frac{D+1}{2}} \forall j \in \{LR, NY\}$*, the utility of* $\tilde{t}^j_{med}$ *is weakly greater under a parliamentary system than a presidential system.*

Proof. See Appendix D. □

From Proposition 4, we know that the expected 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 leader yet it is parliamentary systems with their indirect election of a leader that is more representative of voters preferences. This difference occurs because there is more than one dimension of 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 suboptimal 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.[24] In this section, I ask whether increased polarisation can affect candidate platforms and implemented policies; and whether the impact differs

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

[24] 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.

between single-district and legislative elections. I define an increase in polarisation as an increase in the share of voters who care primarily about the left-right dimension keeping the overall share of voters who prefer policy ij fixed. For example, while the share of voters who support a left, pro-reform policy remains the same - within that group, there is an increase in t_{LY} voters and a corresponding decrease in t_{YL} voters. No voter changes which policies they prefer but some voters become more partisan. More formally:

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 $\tilde{t}^j_{med}, \tilde{t}^j_{\frac{D+1}{2}}, j \in \{LR, NY\}$ 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 behavior, candidate platforms or implemented policies. 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 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 suboptimal, 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.

6 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 parliamentary systems lead to better policies than 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. Second, the margin of victory in a district will depend on which dimension of policy voters are focusing on. A corollary 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 frontrunner 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. 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. It 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.

In the remainder of this section, I will discuss the robustness of the assumptions made and briefly cover possible extensions. A fundamental assumption that driving the main results of the paper 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 instead 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 polices 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. 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 binary by voters). But what if one or both dimensions are non-binary? This would not cause any problems in analysing individual districts; however, we would have to include a coalition formation stage in the legislature as any given policy on a dimension is no longer guaranteed a majority of votes.

Similarly, the focus on two dimensions of policy has been to keep things as tractable as possible. With more than two dimensions, the task for voters would be the same - vote based on the most likely pivotal dimension. However, 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 on every dimension they can. With two dimensions of policy, this corresponds to two of the four district types. [25] It can be shown that in a model with q dimensions of policy, the share of district types with no conflicted voters is $\frac{1}{2^{q-1}}$.[26] This share goes to zero as the number of policy dimensions increase, showing that the gulf between naive and strategic voting grows as the number of dimensions increase. For candidates, including more dimensions of policy complicates the platform choice game as there are more ways their platforms can differ. A fully generalised version of the model is something I leave for future work.

[25] (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

[26] 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.

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 vote profile $x_d = (x_d(L), x_d(R))$ given voter strategies is

$$Pr[x_d | k\nu_d] = \frac{e^{-k\nu_d}(k\nu_d)^{x_d(L)}}{x_d(L)!} \frac{e^{-k(1-\nu_d)}(k(1-\nu_d))^{x_d(R)}}{x_d(R)!} \tag{7}$$

Its associated magnitude is

$$mag(x_d) \equiv \lim_{k\to\infty} \frac{\log(Pr[x_d|k\nu_d])}{k} = \lim_{k\to\infty} \nu_d \psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right) \tag{8}$$

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:

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

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 $k \to \infty$.

$$mag(A'_d) < mag(A_d) \Rightarrow \lim_{k\to\infty} \frac{Pr[A'_d]}{Pr[A_d]} = 0 \tag{10}$$

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 9 subject to the appropriate constraints we get

$$\begin{aligned} mag(Lwin) &= 0 \\ mag(Rwin) &= 2\sqrt{\nu_d(1-\nu_d)} - 1 \end{aligned} \tag{11}$$

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

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

$$Pr[\boldsymbol{x}|k\boldsymbol{\nu}] = \prod_{d \in D} \frac{e^{-k\nu_d}(k\nu_d)^{x_d(L)}}{x_d(L)!} \frac{e^{-k(1-\nu_d)}(k(1-\nu_d))^{x_d(R)}}{x_d(R)!} \tag{12}$$

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

$$mag(\boldsymbol{x}) \equiv \lim_{k\to\infty} \frac{log(Pr[\boldsymbol{x}|k\boldsymbol{\nu}])}{k} = \lim_{k\to\infty} \sum_{d \in D} \nu_d \psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right) = \sum_d mag(x_d) \tag{13}$$

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 $\bar{x_d} \in A_d = \underset{x_d}{\operatorname{argmax}}\ \nu_d\psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right)$, 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:

$$mag(\boldsymbol{A}) = \sum_{d=1}^{D} mag(A_d) = \sum_{d=1}^{D} mag(\bar{x_d}) = mag(\bar{\boldsymbol{x}}) \tag{14}$$

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 13. Together they show that the single-district magnitude theorem extends to a multi-district setting.

Following from this, and using Equation 10, we have that the corollary to the magnitude theorem also extends to the multi-district case. If $mag(\boldsymbol{A'}) < mag(\boldsymbol{A})$, then

$$\lim_{k\to\infty} \frac{Pr[\boldsymbol{A'}|k\boldsymbol{\nu}]}{Pr[\boldsymbol{A}|k\boldsymbol{\nu}]} = \lim_{k\to\infty} \frac{e^{kmag(\boldsymbol{A'})}}{e^{kmag(\boldsymbol{A})}} = \lim_{k\to\infty} e^{k(mag(\bar{\boldsymbol{x}}')-mag(\bar{\boldsymbol{x}}))} = 0 \tag{15}$$

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.[27] 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

[27] 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.

$$\begin{aligned} Pr[Lwin|\mu_L,\mu_R] &= (\mu_L\mu_R + (1-\mu_L)(1-\mu_R))Pr[Lwin|(a_{LN},a_{RN})] \\ &+ \mu_L(1-\mu_R)Pr[Lwin|(a_{LY},a_{RN})] \\ &+ (1-\mu_L)(\mu_R)Pr[Lwin|(a_{LN},a_{RY})] \end{aligned} \tag{16}$$

Letting $\hat{t}_{LY} \equiv t_{LY} \cup t_{YL} \cup t_{YR} \cup t_{LN}$ and $\hat{t}_{LN} \equiv t_{LN} \cup t_{NL} \cup t_{NR} \cup t_{LY}$ we 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 as $k \to \infty$ 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	$f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L), 0.5$	$\mu_L^* \to 0.5$	$\mu_R^* \to 0.5$	mixed
3b	$0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$	$\mu_L^* \to 1$	$\mu_R^* \to 0$	(a_{LY}, a_{RN})
3c	$0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$	$\mu_L^* \to 0$	$\mu_R^* \to 1$	(a_{LN}, a_{RY})
	$f_d(t_L) > f_d(\hat{t}_{LN}), f_d(\hat{t}_{LY})$			
4a	$0.5, f_d(t_L) > f_d(\hat{t}_{LN}), f_d(\hat{t}_{LY})$	$\mu_L^* \to 0.5$	$\mu_R^* \to 0.5$	mixed
4b	$f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$	$\mu_L^* \to 1$	$\mu_R^* \to 0$	(a_{LY}, a_{RN})
4c	$f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$	$\mu_L^* \to 0$	$\mu_R^* \to 1$	(a_{LN}, a_{RY})

Table 4: Equilibrium platforms in the single-district candidate competition game as $k \to \infty$.

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

In a legislature with D districts where each district could potentially have one of 4 candidates $\{LN, LY, RN, RY\}$ elected there are $\frac{(D+1)(D+2)(D+3)}{6}$ possible seat distributions S.[28] From the point of view of a voter in district d - she needs to consider the expected seat distribution in the other $D-1$ districts when casting her vote. Call this $E(S_{-d})$. There are $\frac{(D)(D+1)(D+2)}{6}$ possible seat distributions $E(S_{-d})$. Of these, $\frac{D+1}{2}$ will be double-pivotal events. Those are cases where both $s_{LN} = s_{RY}$ and $s_{LN} = s_{RY}$.

There are $\sum_{n=1}^{\frac{D-1}{2}} 4n$ seat distributions which are single-pivotal events, that is where $s_{LN} + s_{LY} = s_{RN} + s_{RY}$ or $s_{LN} + s_{RN} = s_{LY} + s_{RY}$. There are the same number of events where either $s_{LN} = s_{RY}$ or $s_{LN} = s_{RY}$ holds but both do not hold simultaneously, I call these *diagonal* seat distributions. This set of expected seat distributions is important, as only with a diagonal seat distribution is it possible for a district to condition on the double-pivotal event. Combining all of the above, we can express a lower bound on the share of seat distributions in which no district conditions on a double-pivotal event as

$$1 - \frac{(D+1) + \sum_{n=1}^{\frac{D-1}{2}} 24n}{(D)(D+1)(D+2)} \tag{17}$$

For $D = 101$ it is ≈ 0.971 , for the US House of Representatives with $D = 435$ it is ≈ 0.993 and for the UK House of Commons with $D = 649$ it is ≈ 0.9954. Thus, in the vast majority of possible seat distributions, it is not possible for any district to condition on a double-pivotal event.

We can go one step further and examine diagonal seat distributions - those seat distributions which in theory could allow a district to condition on a double-pivotal event. Suppose a district d has a diagonal seat distribution $E(S_{-d})$. When can it condition on a double-pivotal event? If district d faces (a_{LY}, a_{RY}) or (a_{LN}, a_{RN}) it cannot condition on a double-pivotal event as it can only ever be pivotal on the left-right dimension. For district d to be conditioning on a double-pivotal event it must be that it faces either (a_{LY}, a_{RN}) or (a_{LN}, a_{RY}) <u>and</u> its most likely pivotal event is $piv_d(LR, NY)$.

We can thus ask - conditional on $E(S_{-d})$ being a diagonal seat distribution, what is the probability that the most likely pivotal event for district d is $piv_d(LR, NY)$? To do this, we look at every diagonal event and calculate the share of double-pivotal events out of all pivotal events for district d given that distribution $E(S_{-d})$. I make two assumptions in order to be able to do so. First, I assume that for any expected seat distribution $E(S_{-d})$, the expected runner-up in each district is equally likely to be pro- or anti-reform. For example,

[28] This is the $(D+1)$th tetrahedral or triangular pyramidal number.

if we have $D = 11$ and $E(S_{-d}) = (s_{LN}, s_{LY}, s_{RN}, s_{RY}) = (0, 0, 0, 10)$ it could be that all ten of those s_{RY} seats come from (a_{LN}, a_{RY}) districts. Or it could be that nine of them come from (a_{LN}, a_{RY}) districts and one comes from a (a_{LY}, a_{RY}) district, or it could be that eight come from (a_{LN}, a_{RY}) districts and two comes from (a_{LY}, a_{RY}) districts, and so on. There is only one way the first event can happen, 9 ways the second event can happen and 36 ways the third event can happen. Each subsequent term is the next column in Pascal's triangle for $D - 2$.

For any seat distribution $E(S_d)$ we can rank the districts in terms of safeness - where the safer the seat, the higher the probability of the expected winner winning. The second assumption I make is that each district is equally likely to be in every position in the safeness ranking. Returning to the example of $E(S_d) = (s_{LN}, s_{LY}, s_{RN}, s_{RY}) = (0, 0, 0, 10)$, this means that the probability of a district being in the one of the 5 "safest" districts is $\frac{5}{10}$. The probability of two particular districts being in the top 5 safest districts is $\frac{5}{10} \times \frac{4}{9}$ and so on. Given these two assumptions, the share of cases where a district d facing a diagonal distribution $E(S_{-d})$ conditions on a double-pivotal event is given by

$$\frac{\sum_{k=0}^{\frac{D-3}{2}} 2^k \left[1 + \sum_{i=1}^{\frac{D-3}{2}-k} \sum_{x=0}^{i} \binom{i}{x} \frac{(D-1-2i-k)!}{(\frac{D-1}{2}-i)!} \frac{(\frac{D-1}{2}+x)!}{(D-1-i-k+x)!}\right]}{\binom{\frac{D+1}{2}}{2}(2^{\frac{D-1}{2}})}$$

For a legislature of $D = 101$ the share is ≈ 0.0045, for the US House of Representatives with $D = 435$ it is ≈ 0.0015 for the UK House of Commons with $D = 649$ it is ≈ 0.0007. Note that this is the share of diagonal seat distributions where a district is conditioning on a double-pivotal event, not the share of *all* seat distributions. As we saw above, the share of distributions that are diagonal are themselves tiny.

Appendix D - Proofs

Proof of Lemma 1

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

In (a_{LN}, a_{RN}) and (a_{LY}, a_{RY}) districts the relevant gain function is Equation 3. This gain function will always be positive for t_L and negative for t_R types regardless of the values of the pivot probabilities. In (a_{LY}, a_{RN}) districts the relevant gain function is Equation 4. This gain function will always be positive for t_{LY} and t_{YL} types but negative for t_{RN} and t_{NR} types regardless of the values of the pivot probabilities. In (a_{LN}, a_{RY}) districts the relevant gain function is Equation 5. This gain function will always be positive for t_{LN} and t_{NL} types but negative for t_{RY} and t_{YR} types regardless of the values of the pivot probabilities.

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

In (a_{LY}, a_{RN}) districts every pivot probability is positive if all other $D-1$ districts are not (a_{LY}, a_{RN}) districts, which is guaranteed by Assumption 1. As types $t_{LN}, t_{NL}, t_{RY}, t_{YR}$ have both positive and negative terms in their gain functions, their best response will depend on the values of the various pivot probabilities. These pivotal probabilities depend on the strategies of voters in the other $D-1$ districts; therefore, the best response of conflicted voters in d depends on the strategies of voters in the other $D-1$ districts. By the same argument, the best response of types $t_{LY}, t_{YL}, t_{RN}, t_{NR}$ in (a_{LN}, a_{RY}) districts depends on the strategies of voters in the other $D-1$ districts.

□

Proof of Proposition 1

Proof. Take a (a_{LY}, a_{RN}) district and suppose there is a pivotal event which has a larger magnitude than all others. Without loss of generality let this be $piv_d(LR|Y)$. We can then divide both sides of the relevant gain function Equation 4 by the probability of this pivotal event to get

$$\frac{G_{t,d}(v_L|(a_{LY}, a_{RN}), k\boldsymbol{\nu}, \mathbf{a_{-d}})}{Pr[piv_d(LR|Y)]} = (U(z_{LY}|t) - U(z_{RY}|t) \\ + \frac{Pr[piv_d(LR|N)]}{Pr[piv_d(LR|Y)]}(U(z_{LN}|t) - U(z_{RN}|t)) \\ + \frac{Pr[piv_d(NY|L))]}{Pr[piv_d(LR|Y)]}(U(z_{LY}|t) - U(z_{LN}|t)) \\ + \frac{Pr[piv_d(NY|R)]}{Pr[piv_d(LR|Y)]}(U(z_{RY}|t) - U(z_{RN}|t)) \\ + \frac{Pr[piv_d(LR, NY)]}{Pr[piv_d(LR|Y)]}(U(z_{LY}|t) - U(z_{RN}|t)) \tag{18}$$

Taking limits and applying the corollary to the magnitude theorem we get

$$\lim_{k\to\infty} \frac{G_{t,d}(v_L|(a_{LY}, a_{RN}), k\boldsymbol{\nu}, \mathbf{a_{-d}})}{Pr[piv_d(LR|Y)]} = U(z_{LY}|t) - U(z_{RY}|t) \tag{19}$$

Therefore, if there is a unique pivotal event with largest magnitude, voters in d should base their decision only on this event.

It remains to show that there is always a unique pivotal event with largest magnitude. To see this, notice that the magnitude of a pivotal event is determined by the expected vote share of the left candidate across districts $\boldsymbol{\nu}$. Recall that a strictly perfect equilibrium must be robust to epsilon changes in the strategies of players. Therefore, in any strictly perfect equilibrium, the pivotal event with largest magnitude must be unique. □

Proof of Proposition 2

Proof. If voters condition on $piv(LR)$, the positions of candidates on the reform dimension is irrelevant for voters - any choice of μ is a best response. If voters condition on $piv(NY)$, the expected probability of winning for the left candidate is given by Equation 6

The left candidate will choose μ_L to maximise Equation 6, while the right candidate will choose μ_R to minimise it. This yields the following best response correspondences:

$$BR_L(\mu_R) = \begin{cases} 1 & \text{if } \mu_R > \tilde{\mu_R} \\ [0, 1] & \text{if } \mu_R = \tilde{\mu_R} \\ 0 & \text{if } \mu_R < \tilde{\mu_R} \end{cases} \tag{20}$$

$$BR_R(\mu_L) = \begin{cases} 1 & \text{if } \mu_L < \tilde{\mu_L} \\ [0,1] & \text{if } \mu_L = \tilde{\mu_L} \\ 0 & \text{if } \mu_L > \tilde{\mu_L} \end{cases} \tag{21}$$

where $\tilde{\mu_L}$ and $\tilde{\mu_R}$ are given by:

$$\tilde{\mu_L} \equiv \frac{Pr[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] + Pr[Lwin|(a_{LN}, a_{RN})] - 1}{2Pr[Lwin|(a_{LN}, a_{RN})] - 1}$$
$$\tilde{\mu_R} \equiv 1 - \tilde{\mu_L}$$

We can divide districts into 8 possible preference cases, We analyse each in turn:

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

- **Cases 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 $1 - Pr[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] > Pr[Lwin|(a_{LN}, a_{RN}) > 0.5$. 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 $Pr[Lwin|(a_{LN}, a_{RN}) > 1 - Pr[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$. 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 $Pr[Lwin|(a_{LN}, a_{RN}) > Pr[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$. 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 $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.

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 $\nu_d[(a_{LY}, a_{RN}), piv^1(NY)]$ represent the expected vote share of the left candidate if he is anti-reform while the right candidate is pro-reform and voters vote on the reform dimension. $\nu_d[(a_{LN}, a_{RR}), piv^1(NY)] = 1 - \nu_d[(a_{LY}, a_{RN}), piv^1(NY)]$ 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 $\nu_d[(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$ then $mag[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] = 0$ and $mag[Rwin|(a_{LY}, a_{RN}), piv^1(NY)] = 2\sqrt{\nu_d[(a_{LY}, a_{RN}), piv^1(NY)](1 - \nu_d[(a_{LY}, a_{RN}), piv^1(NY)])} - 1$.

Furthermore, if $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$ then we have $mag[Rwin|(a_{LY}, a_{RN}), piv^1(NY)] > mag[Rwin|(a_{LN}, a_{RN})]$. From the corollary to the magnitude theorem we know that if $mag[Rwin|(a_{LY}, a_{RN}), piv^1(NY)] > mag[Rwin|(a_{LN}, a_{RN})]$ then $\lim_{k\to\infty} \frac{Pr[Rwin|(a_{LN}, a_{RN})]}{Pr[Rwin|(a_{LY}, a_{RN}), piv^1(NY)]} = 0$. We can now examine that happens to $(\tilde{\mu_L}, \tilde{\mu_R})$ as $k \to \infty$ in cases 5-8.

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

  We can re-write

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

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

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

  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 $1 - \nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$. This implies $mag[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] = mag[Rwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Rwin|(a_{LY}, a_{RN}), piv^1(NY)] > mag[Lwin|(a_{LN}, a_{RN})] > -1$.

We can re-write

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

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

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

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 $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$. This implies $mag[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] = mag[Lwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Rwin|(a_{LY}, a_{RN}), piv^1(NY)] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can re-write

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

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

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

  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 prefernеces and voter strategies given in Table 1 imply $1 - \nu_d[(a_{LN}, a_{RN})] > 1 - \nu_d[(a_{LY}, a_{RN}), piv^1(NY)] > 0.5$. This implies $mag[Rwin|(a_{LY}, a_{RN}), piv^1(NY)] = mag[Rwin|(a_{LN}, a_{RN})] = 0$ and $0 > mag[Lwin|(a_{LY}, a_{RN}), piv^1(NY)] > mag[Lwin|(a_{LN}, a_{RN})] > -1$.

We can re-write

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

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

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

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. I prove this proposition by showing the probability of the dis-advantaged candidate being the expected winner goes to zero as $k \to \infty$. Let $f_d(t_L), f_d(t_N) > 0.5$ so that the left candidate is advantaged and the majority of voters oppose the reform. The right candidate is the expected winner only if the platforms are (a_{LY}, a_{RN}) and we have $piv^1(NY)$. The realised platforms are (a_{LY}, a_{RN}) with probability $\mu_L(1 - \mu_R)$. However, from Table 2 we see that when $f_d(t_L), f_d(t_N) > 0.5$ we have $\mu_L \to 0$ and $\mu_R \to 1$. Therefore, as $k \to \infty$, the probability of the right candidate being the expected winner, even conditional on the district having $piv^1(NY)$, goes to zero. □

Proof of Proposition 4

Proof. First, I show that $E(z)$ is unique and pinned down by $\mathbf{f}$. From Proposition 2 we know that: (i) within the set of responsive equilibria, $\mathbf{a}$ is pinned down by $\mathbf{f}$; and (ii) as $k \to \infty$ all mixing probabilities go to 0 or 1. Therefore, the realisation of platforms is completely pinned down by $\mathbf{f}$ and voter behaviour is described by Table 1. There may be multiple equilibria in the voting game, but from Proposition 3 we know that in each equilibrium the advantaged candidate chooses the majority-preferred reform policy and wins the seat with probability going to one. The legislative voting rule means that each elected candidate votes for the policies they campaigned under. Thus, there is a direct mapping from $\mathbf{f}$ to the expected policy $E(z)$.

Next, I show that $E(z)$ must be preferred by $\tilde{t}^{LR}_{\frac{n+1}{2}}$ on the left-right dimension and by $\tilde{t}^{NY}_{\frac{n+1}{2}}$ on the reform dimension. Without loss of generality let z_{LN} be the expected policy. It must then be that $E(s_{LN}) + E(s_{LY}) > \frac{D+1}{2}$ and $E(s_{LN}) + E(s_{RN}) > \frac{D+1}{2}$. We know from Proposition 3 that the expected winner has the policy preferred by a majority of voters in d on each dimension. Therefore is must be that $\tilde{t}^{LR}_{\frac{n+1}{2}} \in \{t_L\}$ and $\tilde{t}^{LR}_{\frac{n+1}{2}} \in \{t_N\}$.

□

Proof of Proposition 5

Proof. 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. In a given district, z_{LN} is a Condorcet winner if $f_d(t_L) > 0.5$, $f_d(t_N) > 0.5$, and $f_d(t_{LY} \cup t_{LN} \cup t_{NL} \cup t_{NR}) > 0.5$ all hold. We know from Proposition 3 that if $f_d(t_L), f_d(t_N) > 0.5$, the expected winner is the left candidate with platform a_{LN}, and he wins with probability going to 1 as $k \to \infty$. Therefore, if z_{LN} is a Condorcet winner in district d, a candidate with platform a_{LN} will win the seat with probability going to 1.

Now, suppose z_{LN} is a Legislative Condorcet winner. It must therefore 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. From Proposition 3 we know that in the districts where (i) holds the left candidate is the expected winner and in districts where (ii) holds the expected winner has an anti-reform platform. Therefore there must be a majority of districts where the left candidate is the expected winner, and a majority of districts where an anti-reform candidate is the expected winner. Given how seats map into policy in the legislature, it must therefore be that the expected implemented policy $E(z) = z_{LN}$.

□

Proof of Proposition 6

Proof. By adding $\tilde{t}^j_{med} = \tilde{t}^j_{\frac{D+1}{2}} \forall j \in \{LR, NY\}$ to Proposition 4, we know that the unique expected policy $E(z)$ in parliamentary system is that which is preferred by $\tilde{t}^j_{med}$ on dimension $j \in \{LR, NY\}$. All that remains to show is that utility of $\tilde{t}^j_{med}$ is weakly lower under a presidential system.

To prove that the utility of $\tilde{t}^j_{med} \forall j \in \{LR, NY\}$ is weakly lower under a presidential system, I show for each case in Table 4 that the implemented policy z on dimension j need not coincide with that preferred by $\tilde{t}^j_{med}$.

- **Case 1:** Here we will have $z = a_{LN}$ if $\tilde{t}^{LR}_{med} = t_L$ and $z = a_{RN}$ otherwise. However, given $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$, it is possible that $\tilde{t}^{NY}_{med} = t_Y$ in which case z does not coincide with the preferred option of the median on each dimension.

- **Case 2:** Here we will have $z = a_{LY}$ if $\tilde{t}^{LR}_{med} = t_L$ and $z = a_{RY}$ otherwise. However, given $f_d(\hat{t}_{LY}) > f_d(t_L) > f_d(\hat{t}_{LN})$, it is possible that $\tilde{t}^{NY}_{med} = t_N$ in which case z does not coincide with the preferred option of the median on each dimension.

- **Case 3a:**There are two sub-cases to consider here. First, suppose $\tilde{t}^{LR}_{med} = t_L$ - then, with probability 0.5 we will have $z = a_{LN}$ and with probability 0.5 we will have $z = a_{LY}$. In each case, with probability 0.5 z does not coincide with the preferred option of the median on each dimension. Second, suppose $\tilde{t}^{LR}_{med} = t_R$ - each of the four policies will be implemented with probability 0.25. Therefore, with probability 0.75, z does not coincide with the preferred option of the median on each dimension.

- **Case 3b:** Here we will have $z = a_{RN}$. Given $0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$ we must have $\tilde{t}^{LR}_{med} = t_R$ but we can have $\tilde{t}^{NY}_{med} = t_Y$ in which case z does not coincide with the preferred option of the median on each dimension.

- **Case 3c:** Here we will have $z = a_{RY}$. Given $0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$ we must have $\tilde{t}^{LR}_{med} = t_R$ but we can have $\tilde{t}^{NY}_{med} = t_N$ in which case z does not coincide with the preferred option of the median on each dimension.

- **Case 4a:**There are two sub-cases to consider here. First, suppose $\tilde{t}^{LR}_{med} = t_R$ - then, with probability 0.5 we will have $z = a_{RN}$ and with probability 0.5 we will have $z = a_{RY}$. In each case, with probability 0.5 z does not coincide with the preferred option of the median on each dimension. Second, suppose $\tilde{t}^{LR}_{med} = t_L$ - each of the four policies will be implemented with probability 0.25. Therefore, with probability 0.75, z does not coincide with the preferred option of the median on each dimension.

- **Case 4b:** Here we will have $z = a_{LY}$. Given $f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$ we must have $\tilde{t}^{LR}_{med} = t_L$ but we can have $\tilde{t}^{NY}_{med} = t_N$ in which case z does not coincide with the preferred option of the median on each dimension.

- **Case 4c:** Here we will have $z = a_{LN}$. Given $f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$ we must have $\tilde{t}^{LR}_{med} = t_L$ but we can have $\tilde{t}^{NY}_{med} = t_Y$ in which case z does not coincide with the preferred option of the median on each dimension.

$\square$

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 one of the other equilibria.

- **From Case 1 to Case 4a/4b/4c:** Suppose that under f_d^1 we had $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$. This corresponds to Case 1 in Table 4 where the equilibrium platforms are (a_{LN}, a_{RN}) and the expected policy is $E(z) = a_{LN}$

  Now, take a distribution such that polarisation increases - specifically such that $f_d^1(\hat{t}_{LN}) > f_d(t_L) > f_d^2(\hat{t}_{LN})$ (This would be the case if e.g. $f_d^2(t_{RN}) = f_d^1(t_{RN}) + f_d^1(t_{NR})$ and $f_d^2(t_{YL}) = f_d^1(t_{YL})$). The increase in polarisation from f^1 to f^2 gives a re-ordering of the size of voter groups such that now $f_d(t_L) > f_d(\hat{t}_{LN}), f_d(\hat{t}_{LN})$. This corresponds to either Case 4a if $f_d(\hat{t}_{LN}), f_d(\hat{t}_{LY}) < 0.5$, Case 4b if $f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$ and Case 4c if $f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$. In case 4a each candidate mixes with equal probability and the implemented policy is any of the four policies with equal probability. In Case 4b the equilibrium platforms are (a_{LY}, a_{RN}) and the expected policy is $E(z) = a_{LY}$. In Case 4c the equilibrium platforms are (a_{LN}, a_{RY}) and the expected policy is $E(z) = a_{LN}$.

- **From Case 4a/4b/4c to Case 2:** Take a distribution such that polarisation increases - specifically such that $f_d^2(\hat{t}_{LY}) > f_d(t_L) > f_d^1(\hat{t}_{LY})$ (This would be the case if e.g. $f_d^2(t_{LN}) = f_d^1(t_{LN}) + f_d^1(t_{NL})$ and $f_d^2(t_{YR}) = f_d^1(t_{YR})$). The increase in polarisation from f^1 to f^2 gives a re-ordering of the size of voter groups such that now $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LN})$. This corresponds to either Case 2 in Table 4. The equilibrium platforms are then (a_{LY}, a_{RY}) and the expected policy is $E(z) = a_{LY}$.

- **From Case 2 to Case 3a:** Take a distribution such that polarisation increases - specifically such that $f_d^2(\hat{t}_{LN}) > f_d(t_L) > f_d^1(\hat{t}_{LN})$ (This would be the case if e.g. $f_d^2(t_{LY}) = f_d^1(t_{LY}) + f_d^1(t_{YL})$ and $f_d^2(t_{NR}) = f_d^1(t_{NR})$). The increase in polarisation from f^1 to f^2 gives a re-ordering of the size of voter groups such that now $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) >$

$f_d(t_L)$. This corresponds to either Case 3a in Table 4. Each candidate mixes with equal probability and the implemented policy is z_{LN} or z_{LY} with equal probability.

- **From Case 3a to Case 1:** Take a distribution such that polarisation increases - specifically such that $f_d^1(\hat{t}_{LY}) > f_d(t_L) > f_d^2(\hat{t}_{LY})$ (This would be the case if e.g. $f_d^2(t_{RY}) = f_d^1(t_{RY}) + f_d^1(t_{YR})$ and $f_d^2(t_{NL}) = f_d^1(t_{NL})$). The increase in polarisation from f^1 to f^2 gives a re-ordering of the size of voter groups such that now $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$. This corresponds to either Case 1 in Table 4. The equilibrium platforms are then (a_{LN}, a_{RN}) and the expected policy is $E(z) = a_{LN}$.

The above cycle of increases in polarisation shows that regardless of which equilibrium we being with given f_d^1, an increase in polarisation can move us to any of the other five feasible equilibria. The proof for the case of $f_d(t_L), f_d(t_N) > 0.5$ 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 left candidate will choose μ_L to maximise the expression, while the right candidate will choose μ_R to minimise it. This yields the following best response correspondences:

$$BR_L(\mu_R) = \begin{cases} 1 & \text{if } \mu_R > \tilde{\mu_R} \\ [0,1] & \text{if } \mu_R = \tilde{\mu_R} \\ 0 & \text{if } \mu_R < \tilde{\mu_R} \end{cases} \tag{22}$$

$$BR_R(\mu_L) = \begin{cases} 1 & \text{if } \mu_L < \tilde{\mu_L} \\ [0,1] & \text{if } \mu_L = \tilde{\mu_L} \\ 0 & \text{if } \mu_L > \tilde{\mu_L} \end{cases} \tag{23}$$

where $\tilde{\mu_L}$ and $\tilde{\mu_R}$ are given by:

$$\tilde{\mu_L} \equiv \frac{Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN})]}{2Pr[Lwin|(a_{LN}, a_{RN})] - Pr[Lwin|(a_{LY}, a_{RN})] - Pr[Lwin|(a_{LN}, a_{RR})]} \tag{24}$$

$$\tilde{\mu_R} \equiv 1 - \tilde{\mu_L}$$

We can also write Equation 24 as

$$\tilde{\mu_L} \equiv \frac{Pr[Rwin|(a_{LN}, a_{RN})] - Pr[Rwin|(a_{LY}, a_{RN})]}{2Pr[Rwin|(a_{LN}, a_{RN})] - Pr[Rwin|(a_{LY}, a_{RN})] - Pr[Rwin|(a_{LN}, a_{RR})]} \tag{25}$$

We can divide districts into 8 possible preference cases, We analyse each in turn:

- **Case 1:** If $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$, then $\mu_L = \mu_R = 0$. $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$ and voter strategies given in Table 3 imply that $Pr[Lwin|(a_{LN}, a_{RY})] > Pr[Lwin|(a_{LN}, a_{RN}) > Pr[Lwin|(a_{LY}, a_{RN})$. 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 $f_d(\hat{t}_{LY}) > f_d(t_L) > f_d(\hat{t}_{LN})$, then $\mu_L = \mu_R = 1$. $f_d(\hat{t}_{LY}) > f_d(t_L) > f_d(\hat{t}_{LN})$ and voter strategies given in Table 3 imply that $Pr[Lwin|(a_{LY}, a_{RN})] > Pr[Lwin|(a_{LN}, a_{RN}) > Pr[Lwin|(a_{LN}, a_{RY})$. 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 3:** If $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L)$, then $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$. 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 $\tilde{\mu_L} \in (0, 1)$ and $\tilde{\mu_R} \in (0, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.
  - **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 $\tilde{\mu_L} \in (0, 0.5)$ and $\tilde{\mu_R} \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.
  - **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 $\tilde{\mu_L} \in (0.5, 1)$ and $\tilde{\mu_R} \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.

- **Case 4:** If $f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$, then $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$. 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 $\tilde{\mu_L} \in (0, 1)$ and $\tilde{\mu_R} \in (0, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.

– **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 $\tilde{\mu_L} \in (0.5, 1)$ and $\tilde{\mu_R} \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.

– **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 $\tilde{\mu_L} \in (0, 0.5)$ and $\tilde{\mu_R} \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\tilde{\mu_L}, \tilde{\mu_R})$.

We can now examine that happens to $(\tilde{\mu_L}, \tilde{\mu_R})$ 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 25 by $Pr[Rwin|(a_{LN}, a_{RN})]$ to get

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

  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

$$\lim_{k\to\infty} (\tilde{\mu_L}, \tilde{\mu_R}) = (0.5, 0.5)$$

- **Case 3b:** With $0.5, f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY})$ 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 24 by $Pr[Lwin|(a_{LN}, a_{RY})]$ to get

$$\tilde{\mu_L} \equiv \frac{\frac{Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]} - \frac{Pr[Lwin|(a_{LY},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]}}{2\frac{Pr[Rwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]} - \frac{Pr[Lwin|(a_{LY},a_{RN})]}{Pr[Lwin|(a_{LN},a_{RY})]} - 1}$$

  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

$$\lim_{k \to \infty} (\tilde{\mu_L}, \tilde{\mu_R}) = (0, 1)$$

- **Case 3c:** 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 $0.5, \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LN}, a_{RN})]$. 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 24 by $Pr[Lwin|(a_{LY}, a_{RN})]$ to get

$$\tilde{\mu_L} \equiv \frac{\frac{Pr[Lwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LY},a_{RN})]} - 1}{2\frac{Pr[Rwin|(a_{LN},a_{RN})]}{Pr[Lwin|(a_{LY},a_{RN})]} - 1 - \frac{Pr[Lwin|(a_{LN},a_{RY})]}{Pr[Lwin|(a_{LY},a_{RN})]}}$$

  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

$$\lim_{k \to \infty} (\tilde{\mu_L}, \tilde{\mu_R}) = (1, 0)$$

- **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 24 by $Pr[Lwin|(a_{LN}, a_{RN})]$ to get

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

  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

$$\lim_{k \to \infty} (\tilde{\mu_L}, \tilde{\mu_R}) = (0.5, 0.5)$$

- **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 $mag[\text{R-win}|(a_{LY}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RY})] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

We can divide the top and bottom of Equation 25 by $Pr[Rwin|(a_{LY}, a_{RN})]$ to get

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

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

$$\lim_{k\to\infty} (\tilde{\mu_L}, \tilde{\mu_R}) = (1, 0)$$

- **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 $mag[\text{R-win}|(a_{LN}, a_{RY})] > mag[Rwin|(a_{LY}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation 25 by $Pr[Rwin|(a_{LN}, a_{RY})]$ to get

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

  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

$$\lim_{k\to\infty} (\tilde{\mu_L}, \tilde{\mu_R}) = (0, 1)$$

$\square$

References

Aragonès, E., M. Castanheira, and M. Giani (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 N. McCarty (2015): "Causes and Consequences of Polarization," in *Political negotiation: A handbook*.

Besley, T., and S. Coate (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 T. Verthé (2019): "Strategic Voting Versus Sincere Voting," *Oxford Research Encyclopedia of Politics*, (October), 1–20.

Bordignon, M., T. Nannicini, and G. Tabellini (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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Matakos, K., O. Troumpounis, and D. Xefteris (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.

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

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., J. M. Snyder, and M. M. Ting (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.

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

Polborn, M. K., and J. M. Snyder (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.

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