

Original citation:

Hughes, Niall. (2016) Voting in legislative elections under plurality rule. *Journal of Economic Theory*, 166. pp. 51-93.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/82165>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

© 2016, Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>

A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Voting In Legislative Elections Under Plurality Rule*

Niall Hughes^{a†}

^aDept of Economics, University of Warwick, United Kingdom

August 2016

Abstract

Models of single district plurality elections show that with three parties anything can happen - extreme policies can win regardless of voter preferences. I show that when single district elections are used to fill a legislature, we get back to a world where the median voter matters. An extreme policy will generally only come about if it is preferred to a more moderate policy by the median voter in a majority of districts. The mere existence of a centrist party can lead to moderate outcomes even if the party itself wins few seats. I also show that, while some voters in a district will not vote for their nationally preferred party, in many equilibria they *will* want the candidate for whom they vote to win that district. This is never the case in single district elections. There, some voters always want the candidate they voted for to lose.

Keywords: Strategic Voting, Legislative Elections, Duverger's Law, Plurality Rule, Poisson Games.

JEL Classification Number: C71, C72, D71, D72, D78.

*I would like to thank the Associate Editor and the three anonymous referees, as well as Gabriele Gratton, Helios Herrera, David K. Levine, Andrea Mattozzi, Nolan McCarty, Massimo Morelli, Salvatore Nunnari, Gerard Roland, and Richard Van Weelden for many useful comments and suggestions.

[†]Correspondence to: Dept of Economics, University of Warwick, Gibbet Hill Rd., CV4 7AL, United Kingdom. e-mail: n.e.hughes@warwick.ac.uk

1 Introduction

Plurality rule (a.k.a. first-past-the-post) is used to elect legislatures in the U.S., U.K., Canada, India, Pakistan, Malaysia as well as a host of other former British colonies - yet we know very little about how it performs in such settings. The literature on single district elections shows that plurality rule performs well when there are only two candidates but poorly when there are more (Myerson (2000, 2002)). Indeed, plurality has recently been deemed the worst voting rule by a panel of voting theorists (Laslier (2012)). However, the objectives of voters differ in single district and legislative elections. In a legislative election, many districts hold simultaneous plurality elections and the winner of each district takes a seat in a legislature. Once all seats are filled, the elected politicians bargain over the formation of government and implement policy. If voters only care about which policy is implemented in the legislature, they will cast their ballots to influence the outcome of the legislative bargaining stage. A voter's preferred candidate will therefore depend on the results in other districts. In contrast, in a single district election - such as a mayoral election - a voter's preference ordering over candidates is fixed, as only the local result matters. These different objectives are at the heart of this paper. I show that when three parties compete for legislative seats and voters care about national policy, several undesirable properties of plurality rule are mitigated.

While there has been some key work on legislative elections with three parties under proportional representation, scant attention has been paid to this setting under plurality rule. Studies of plurality rule have either focused on two-party legislative competition or else on three-party single district elections. In the former case, as voters face a choice of two parties, they have no strategic decision to make - they simply vote for their favourite. However, for almost all countries using plurality rule, with the notable exception of the U.S., politics is not a two-party game: the U.K. has the Conservatives, Labour and the Liberal Democrats; Canada has the Conservatives, Liberals, and New Democrats; India has Congress, BJP and many smaller parties.¹ With a choice of three candidates, voters must consider how others will vote when deciding on their own ballot choice.

In a single plurality election, only one candidate can win. Therefore, when faced with a choice of three options, voters who prefer the candidate expected to come third have an incentive to abandon him and instead vote for their second favourite, so that in equilibrium only two candidates receive votes. These are the only serious candidates. This effect is

¹The recent 2015 UK election witnessed a further fragmentation of the political landscape with the Scottish National Party gaining a large number of seats. Other countries with plurality rule and multiple parties represented in the legislature include: Bangladesh, Botswana, Kenya, Liberia, Malawi, Malaysia, Mongolia, Pakistan, Trinidad & Tobago, Tanzania, Uganda, and Zambia.

known as Duverger’s law.

A vast literature has pointed out two implications of Duverger’s law in single district elections (Palfrey (1989); Myerson and Weber (1993); Cox (1997); Fey (1997); Myerson (2002); Myatt (2007)). First, “*anything goes*”: the equilibrium is completely driven by voters’ beliefs, so *any* of the three candidates could be abandoned, leaving the other two to share the vote. This means that, regardless of voter preferences, there can always be extreme policy outcomes - where a race between the two extreme choices results in an implemented policy far away from the centre. Second, when each of the three choices is preferred by some voter, there will always be *misaligned voting*. That is, some people will vote for an option which is not their most preferred.² Misaligned voting undermines the legitimacy of the elected candidate: one candidate may win a majority simply to “keep out” a more despised opponent, so the winner’s policies may actually be preferred by relatively few voters.

In this paper, I model a legislative election in which each voter casts a ballot in a local district but their utility depends on policy determined in the national parliament. I show that while extreme policy outcomes and misaligned voting are always possible in stand-alone multi-candidate plurality elections, they can both be mitigated in a legislative election setting.

The intuition for the first result is as follows. For an extreme policy to come about, either the left or right party must win a majority of seats. For any party to win a majority of seats in the legislature it must be that they are preferred to some alternative by a majority of voters in a majority of districts. I show that the alternative to a left majority will generally not be a right majority but rather a moderate coalition government. Therefore, for an extreme policy to come about, it must be that a majority of voters in a majority of districts prefers this policy to the moderate coalition policy. This contrasts with single plurality elections where extreme policy outcomes are always possible, regardless of preferences.

The misaligned voting result stems from the fact that voters condition their ballots on a wider set of events in my setting. In a standard plurality election, voters condition their vote on the likelihood of being pivotal in their district. However, in a legislative election, voters will condition their ballot choice on their vote being pivotal and their district being decisive in determining the government policy. In many cases a district will be decisive between two policies, even though there are three candidates. For example, a district might be decisive in either granting a majority of seats to a non-centrist party, say the left party, or bringing about a coalition by electing one of the other parties. Under many bargaining rules this coalition policy will be the same regardless of which of the weaker parties is elected. So, voters only face a choice between two policies: that of the left party and that of the coalition.

²Misaligned voting is formally defined in subsection 4.2.

When voters have a choice over two policies there can be no misaligned voting - everyone must be voting for their preferred option of the two.

One technical contribution of the paper is to extend the Poisson games framework of Myerson (2000) to a multi-district setting. I show in Lemma 1 that the Magnitude Theorem and its Corollary can be used to rank the likelihood of various pivotal events across districts. This greatly reduces the complexity of working with multi-district pivotal events and makes the problem much more tractable. Proposition 1 uses this feature to show that multiple (strictly perfect) equilibria always exist. In every equilibrium all districts are duvergerian.

In my benchmark model, parties in the legislature bargain only over policy and do not discount the future. Here, if no party holds a majority of seats, the moderate party's policy will be implemented. While this enormous power of the moderate party may be debatable in reality, the simplicity of bargaining underlines the novelty of the model's voting stage. Two clear predictions emerge from this benchmark model. First, when the moderate party wins at least one seat, extreme policies outcomes are mitigated: the policy of the left or right party can only be implemented if a majority of voters in a majority of districts prefer it to the policy of the moderate party. This result gives us conditions under which a median voter should prefer a parliamentary or a presidential system. Second, if either the left or right party is a serious candidate in less than half of the districts, there can be no misaligned voting. These results change somewhat under different bargaining rules, but their flavour remains the same. When parties bargain over perks of office as well as policy, the first result is generally strengthened - it is even more difficult to have extreme outcomes - while the misaligned voting result is weakened - it can only be ruled out if a non-centrist party is serious in less than a quarter of districts. When I add discounting to the benchmark model (in Appendix C), the power of the moderate party is reduced. Nonetheless, an extreme policy can still only be implemented if it is preferred to a more moderate coalition policy. Here, misaligned voting cannot generally be eliminated in all districts but may be ruled out in a large subset of districts.

This paper contributes primarily to the theoretical literature on strategic voting in legislative elections. The bulk of this work has been on proportional representation. Austen-Smith and Banks (1988) find that, with a minimum share of votes required to enter the legislature, the moderate party will receive just enough votes to ensure representation, with the remainder of its supporters choosing to vote for either the left or right party. Baron and Diermeier (2001) show that, with two dimensions of policy, either minimal-winning, surplus, or consensus governments can form depending on the status quo. On plurality legislative elections Morelli (2004) and Bandyopadhyay, Chatterjee, and Sjöström (2011) show that if parties can make pre-electoral pacts, and candidate entry is endogenous, then voters will not

need to act strategically. My paper nonetheless focuses on strategic voting because in the main countries of interest, the U.K. and Canada, there are generally no pre-electoral pacts and the three main parties compete in almost every district, so strategic candidacy is not present.³

The paper proceeds as follows. In the next section I introduce the benchmark model and define an equilibrium. In Section 3, I solve the model and show conditions which must hold in equilibrium. Section 4 presents the main results on extreme policy outcomes and misaligned voting in the benchmark model. Section 5 adds perks of office to the bargaining stage of the model. Section 6 discusses the assumptions of the model and concludes.

2 Model

Parties Three parties; l , m , and r , contest simultaneous elections in D districts, where D is an odd number. Each district is decided by plurality rule: whichever party receives the most votes in a district $d \in D$ is deemed elected and takes a seat in the legislature.⁴ This gives a distribution of seats in the legislature, $S \equiv (s_l, s_m, s_r)$. Each party $c \in \{l, m, r\}$ has a fixed policy platform $a_c \in [-1, 1]$, with $a_l < a_m < a_r$. Once all the seats in the legislature have been filled, the parties bargain over the formation of government and implement a policy $z \in [-1, 1]$. A party's payoff is $-(z - a_c)^2$, i.e. it wants to minimise the distance between its platform and the final policy.

The benchmark legislative bargaining model I use is that of Baron (1991).⁵ If a party has a majority of seats it chooses a policy unilaterally; if no party wins a majority, we enter a stage of legislative bargaining. First, one party is randomly selected as formateur, where the probability of each party being chosen is equal to its seat share in the legislature. The formateur proposes a policy in $[-1, 1]$. This policy is implemented if a majority of the legislature support it; if not, a new formateur is selected, under the same random recognition rule, and the process repeats itself until agreement is reached.⁶ I assume for now that parties are

³In the 2015 U.K. General Election, the three main parties all contested 631 out of 650 districts (Only the Conservatives contested seats in Northern Ireland), while in the 2015 Canadian Federal Election the three major parties contested all 338 seats.

⁴I abuse notation slightly, letting D be both the number of districts and the set of districts.

⁵A large literature has grown from legislative bargaining model of Baron and Ferejohn (1989), in which legislators bargain over the division of a dollar. See Baron (1991), Banks and Duggan (2000), Baron and Diermeier (2001), Jackson and Moselle (2002), Eraslan, Diermeier, and Merlo (2003), Kalandrakis (2004), Banks and Duggan (2006), and many others. Morelli (1999) introduces a different approach to legislative bargaining whereby potential coalition partners make demands to an endogenously chosen formateur. In contrast with the Baron and Ferejohn (1989) setup, the formateur does not capture a disproportionate share of the payoffs.

⁶My results hold whether we use a random recognition rule or a fixed-order rule, another common way of modelling the bargaining process. Under the fixed order rule, the party with the largest number of seats

perfectly patient, $\delta = 1$, but this is relaxed in Appendix C. As is standard in the bargaining literature, I restrict attention to stationary strategies.⁷ A party's strategy specifies which policy to propose if formateur, and which policies to accept or reject otherwise.⁸

Voters A voter's type, $t \in [-1, 1]$, is simply his position on the policy line; his utility is $u_t(z) = -(z - t)^2$. As such, a voter does not care who wins his district *per se*, nor does he care which parties form government. All that matters is the final policy, z , decided in the legislature. Following Myerson (2000, 2002), the number of voters in each district is not fully known but rather is a random variable n_d , which follows a Poisson distribution and has mean n .⁹ Appendix A summarises several properties of the Poisson model. 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.¹⁰ The *actual* population of voters in d consists of n_d independent and identically distributed (i.i.d.) draws from a distribution f_d that admits a density and has full support over the type space $[-1, 1]$.¹¹ The draws in d from f_d are independent of the draws in any other district d' from $f_{d'}$. A voter knows his own type, the distribution from which he was drawn, and the distribution functions of the other districts, $\mathbf{f} \equiv (f_1, \dots, f_d, \dots, f_D)$, but he does not know the actual distribution of voters that is drawn in any district. Let $V \equiv \{v_l, v_m, v_r\}$ be the set of actions an individual can take, with v_c indicating a vote for party c . Voting is costless; thus, there will be no abstention.

A voting strategy is $\sigma : [-1, 1] \times D \rightarrow \Delta(V)$ where $\sigma_{t,d}(v_c)$ is the probability that a type t voter in district d casts ballot v_c .¹² The usual constraints apply: $\sigma_{t,d}(v_c) \geq 0, \forall c$ and $\sum_c \sigma_{t,d}(v_c) = 1, \forall (t, d)$. In a Poisson game, all voters of the same type in the same district

proposes a policy in $[-1, 1]$, which is implemented if a majority of the legislature support it; if not, the second largest party proposes a policy. If this second policy does not gain majority support, the smallest party proposes a policy, and if still there is no agreement, a new round of bargaining begins with the largest party again first to move. See Baron (1991) for more on random and fixed order bargaining rules.

⁷A party's strategy is stationary if its continuation strategy is the same at the beginning of any period, regardless of the history of play.

⁸More formally, a stationary strategy for party c consists of a proposal $y_c \in [-1, 1]$ offered any time c is recognised and a measurable decision rule $r_c : [-1, 1] \rightarrow \{\text{accept}, \text{reject}\}$.

⁹The probability that there are exactly k voters in a district is $Pr[n_d = k] = \frac{e^{-n} n^k}{k!}$.

¹⁰Krishna and Morgan (2011) use a Poisson model to show that in large elections, voluntary voting dominates compulsory voting when voting is costly and voters have preferences over ideology and candidate quality. Bouton and Castanheira (2012) use a Poisson model to show that when a divided majority need to aggregate information as well as coordinate their voting behaviour, approval voting serves to bring about the first-best outcome in a large election. Furthermore, Bouton (2013) uses a Poisson model to analyse the properties of runoff elections.

¹¹Thus the measure of voters who are indifferent between two distinct policies is zero.

¹²More precisely, the strategy $\sigma_{t,d}(v_c)$ is the marginal distribution of Milgrom-Weber distributional strategies (Milgrom and Weber (1985)).

must follow the same strategy.¹³ Given the various $\sigma_{t,d}$'s, the expected vote share of party c in the district is

$$\tau_d(c) = \int_{-1}^1 \sigma_{t,d}(v_c) f_d(t) dt \quad (1)$$

which can also be interpreted as the probability of a randomly selected voter playing v_c . The expected distribution of party vote shares in d is $\tau_d \equiv (\tau_d(l), \tau_d(m), \tau_d(r))$. The realised profile of votes is $x_d \equiv (x_d(l), x_d(m), x_d(r))$, but this is uncertain *ex ante*. As the population of voters is made up of n_d independent draws from f_d , where $E(n_d) = n$, the expected number of ballots for candidate c is $E(x_d(c)|\sigma_d) = n\tau_d(c)$. In the event of a two-way tie, a coin toss determines the winner, while in a three-way tie each party wins with probability $\frac{1}{3}$. If nobody votes, I assume that party m wins the seat.¹⁴ Let $\sigma \equiv (\sigma_1, \dots, \sigma_d, \dots, \sigma_D)$ denote the profile of voter strategies across districts and let σ_{-d} be that profile with σ_d omitted. Let $\tau \equiv (\tau_1, \dots, \tau_d, \dots, \tau_D)$ denote the profile of expected party vote share distributions and let τ_{-d} be that profile with τ_d omitted. Thus, we have $\tau(\sigma, \mathbf{f})$.

Pivotality, Decisiveness and Payoffs A single vote is *pivotal* if it makes or breaks a tie for first place in the district. A district is *decisive* if the policy outcome z depends on which candidate that district elects. When deciding on his strategy, a voter need only consider cases in which his vote affects the policy outcome. Therefore, he will condition his vote choice on being *pivotal* in his district and on the district being *decisive*.

Let $piv_d(c, c')$ denote when, in district d , a vote for party c' is pivotal against c . This occurs when $x_d(c) = x_d(c') \geq x_d(c'')$ – so that an extra vote for c' means it wins the seat – or when $x_d(c) = x_d(c') + 1 \geq x_d(c'')$ – so that an extra vote for c' forces a tie. Let λ_d denote an event in which district d is decisive in determining which policy z is implemented; and let λ_d^i denote the i -th most likely decisive event for district d . Here the decision of district d will lead to one of three final policy outcomes, thus, we can write each decisive event as $\lambda_d^i(z^i)$ or $\lambda_d^i(z_l^i, z_m^i, z_r^i)$ where each z_c^i is the policy outcome of the legislative bargaining stage when the decisive district elects party c . Note that these policies need not correspond to the announced platforms of the parties – typically coalition bargaining will lead to compromised policies. Two decisive events λ_d^i and λ_d^j are *distinct* if $(z_l^i, z_m^i, z_r^i) \neq (z_l^j, z_m^j, z_r^j)$. Let Λ be the set of all decisive events in which each element in the set is distinct from every other element in the set¹⁵; this set consists of I elements. As we will see, the number and type of

¹³This stems from the very nature of population uncertainty. See Myerson (1998) pg. 377 for more detail.

¹⁴The probability of zero turnout in a district is e^{-n} .

¹⁵That is, for any two elements λ_d^i and λ_d^j in the set Λ , we have $(z_l^i, z_m^i, z_r^i) \neq (z_l^j, z_m^j, z_r^j)$.

decisive events in the set Λ depends on the legislative bargaining rule used.

It is useful for the following sections to classify decisive events into three categories. Let $\lambda(3)$ be a decisive event where all three policies z_l^i , z_m^i , and z_r^i are different points on the policy line; let $\lambda(2)$ be a case where two of the three policies are identical.¹⁶ Finally, let $\lambda(2')$ be an event where there are three different policies but one of them is the preferred choice of no voter.¹⁷ We can now turn to voter payoffs.

Let $G_{t,d}(v_c|n\tau)$ denote the expected gain for a voter of type t in district d of voting for party c , given the strategies of all other players in the game – this includes players in his own district as well as those in the other $D - 1$ districts. The expected gain of voting v_l is given by

$$G_{t,d}(v_l|n\tau) = \sum_{i=1}^I Pr[\lambda_d^i] \left(Pr[piv_d(m, l)] \left(u_t(z_l^i) - u_t(z_m^i) \right) + Pr[piv_d(r, l)] \left(u_t(z_l^i) - u_t(z_r^i) \right) \right) \quad (2)$$

with the gain of voting v_m and v_r similarly defined. The probability of being pivotal between two candidates, $Pr[piv_d(c, c')]$, depends on the strategies and distribution of player types in that district, summarised by τ_d , while the probability of district d being decisive depends on the strategies and distributions of player types in the other $D - 1$ districts, τ_{-d} . The best response correspondence of a type t in district d to a strategy profile and distribution of types given by τ is

$$BR_{t,d}(n\tau) \equiv \operatorname{argmax}_{\sigma_{t,d}} \sum_{v_c \in V} \sigma_{t,d}(v_c) G_{t,d}(v_c|n\tau) \quad (3)$$

Equilibrium Concept The equilibrium of this game consists of a *voting equilibrium* and a *bargaining equilibrium*. In a *bargaining equilibrium*, each party's strategy is a best response to the strategies of the other two parties. I restrict attention to stationary bargaining equilibria, as is standard in such games.¹⁸

As we are interested in the properties of large national elections, at the voting stage I will analyse asymptotic equilibria.¹⁹ That is, the limit of the set of equilibria as $n \rightarrow \infty$. More specifically, following Bouton and Gratton (2015), I restrict attention to asymptotic strictly

¹⁶Obviously, $\lambda(1)$ events cannot exist; if electing any of the three parties gives the same policy, it is not a decisive event.

¹⁷For this to be the case, the universally disliked policy must be a lottery over two or more policies.

¹⁸An equilibrium is stationary if it is a subgame perfect equilibrium and each party's strategy is the same at the beginning of each bargaining period, regardless of the history of play.

¹⁹Let $\hat{\Gamma} \equiv \{\Gamma_n\}_{n \rightarrow \infty}$ be a sequence of games. A strategy profile σ^* is an asymptotic equilibrium of $\hat{\Gamma}$ if there exists a sequence of Nash equilibria $\{\sigma_n^*\}_{n \rightarrow \infty}$ of Γ_n such that $\sigma_{t,d,n}^* \rightarrow \sigma_{t,d}^*$ for almost all t in every d .

perfect equilibria.²⁰ A sequence of Nash equilibria $\{\sigma_n^*\}_{n \rightarrow \infty}$ is asymptotically strictly perfect if (i) it admits a limit and (ii) as n grows large, $Pr[t \in [-1, 1] : \sigma_{t,d}^* \notin BR_{t,d}(n\tilde{\tau}_n)] \rightarrow 0$ in every district for any $\tilde{\tau}_n$ sufficiently close to $\tau(\sigma_n^*, \mathbf{f})$.²¹ That is, the equilibrium must be robust to epsilon changes in the strategies of players. Bouton and Gratton (2015) argue that restricting attention to such equilibria in multi-candidate Poisson games is appropriate because it rules out unstable and undesirable equilibria identified by Fey (1997). If, instead, Bayesian Nash equilibrium is used there may be knife-edge equilibria in which voters expect two or more candidates to get exactly the same number of votes. Bouton and Gratton (2015) also note that requiring strict perfection is equivalent to robustness to heterogenous beliefs about the distribution of preferences.

3 Equilibrium

Bargaining Equilibrium The following proposition shows the policy outcome of legislative bargaining for any distribution of seats.

Proposition 0 (Jackson and Moselle (2002)). *The equilibrium policy outcome of the legislative bargaining stage is*

$$z = \begin{cases} a_l & \text{if } s_l > \frac{D-1}{2} \\ a_r & \text{if } s_r > \frac{D-1}{2} \\ a_m & \text{otherwise} \end{cases} \quad (4)$$

When no party has a majority of seats and $\delta = 1$, in any stationary bargaining equilibrium $z = a_m$ is proposed and eventually passed with probability one.²² To see this, note that if any other policy is proposed, a majority of legislators will find it worthwhile to wait until a_m is offered (which will occur when party m is eventually chosen as formateur). The result is regardless of whether the protocol is fixed order or random.

Every feasible seat distribution is mapped into a policy outcome, so, voters can fully anticipate which policy will arise from a given seat distribution. The set of distinct decisive events is given by

$$\Lambda = \{\lambda(a_l, a_m, a_m), \lambda(a_m, a_m, a_r), \lambda(a_l, a_m, a_r)\} \quad (5)$$

²⁰The original formulation of strictly perfect equilibrium was for games with a finite number of players; Bouton and Gratton (2015) in their Appendix C extend this to Poisson games and define its asymptotic version.

²¹More formally, An asymptotic equilibrium σ^* is *asymptotically strictly perfect* if there exists a sequence of Nash equilibria $\{\sigma_n^*\} \rightarrow \sigma^*$ such that for any $\delta > 0$ there exist $N \in \mathbb{N}$ and $\epsilon > 0$ such that for any $n > N$, if $\forall \tilde{\tau}_d \in \Delta V : |\tilde{\tau}_d - \tau_d(\sigma^*, \mathbf{f})| < \epsilon$, then $Pr[t \in [-1, 1] : \sigma_{t,d}^* \in BR_{t,d}(n\tilde{\tau})] < \delta$ for almost all $t \in [-1, 1]$.

²²For a discussion see Baron (1991) and for a proof see Jackson and Moselle (2002).

These cases correspond to when party l is one seat short of a majority, party r is one seat short of a majority, and when both parties l and r are one seat short of a majority respectively.²³

Voting Equilibrium A slight detour on how voters optimally cast their vote is in order before describing the voting equilibrium. The results derived in this subsection do not rely on the specific bargaining protocol and party preferences. I shall use this fact later in Section 5. We know from the Magnitude Theorem (Myerson (2000), see Appendix A) that as $n \rightarrow \infty$ voters in a single district election need only condition their choice on the most likely vote profile in which their vote is pivotal. The following lemma extends this result to the case of multi-district elections considered here.

Lemma 1. *As $n \rightarrow \infty$, if there is a unique most likely event in which his vote is both pivotal and decisive over two distinct policy outcomes z and z' , then a voter who is not indifferent between z and z' need only condition his choice on this single event.*

Proof. See Appendix B. □

Indifferent voters are of measure zero, so can safely be ignored. The lemma greatly simplifies the decision process for a voter - he can ignore all but the most likely (i.e. largest magnitude) case in which his vote is pivotal and decisive. This is because as the size of the electorate increases, the overall probability of being pivotal and decisive goes to zero but, conditional on being pivotal and decisive, almost all the probability falls on the most likely event in that set of events. The lemma states that this is the case if there is a unique pivotal and decisive event with largest magnitude. However, we might worry that more than one event has the largest magnitude. The proof of Proposition 1 below shows that in any asymptotic strictly perfect equilibrium, the pivotal and decisive event with largest magnitude is unique. This uniqueness ensures that districts are duvergerian.

Proposition 1. *For any majoritarian legislative bargaining rule and any distribution of voter preferences, $\mathbf{f} \equiv (f_1, \dots, f_d, \dots, f_D)$, there are multiple equilibria; in every equilibrium districts are duvergerian.*

Proof. See Appendix B. □

It is perhaps unsurprising that there are multiple equilibria and districts are always duvergerian, especially given the findings of the extensive literature on single district plurality elections. The logic as to why races are duvergerian is similar to the single district case; voters

²³The final case can only occur when party m has no seats.

condition their ballot choice on the most likely case in which they are pivotal, decisive and not indifferent over outcomes. This greatly simplifies the decision process of voters and means they need only consider the two frontrunners in their district.²⁴ The following corollary and the properties which follow it will be useful for the analysis in the remainder of the paper.

Corollary to Proposition 1. *For any majoritarian legislative bargaining rule and any distribution of voter preferences, there always exists an equilibrium in which party m wins at least one seat in the legislature.*

Proof. See Appendix B. □

While we cannot pin down which equilibrium will be played, the following properties will always hold.

1. In each district only two candidates receive votes; call these *serious candidates*.
2. If $\tau_d(c) > \tau_d(c') > \tau_d(c'') = 0$, candidate c is the expected winner and his probability of winning goes to one as $n \rightarrow \infty$. Let d_c denote such a district and let D_c be the set of districts with c as the expected winner.
3. The expected seat distribution is $E(S) = E(s_l, s_m, s_r) = (\#D_l, \#D_m, \#D_r)$.
4. A district with c and c' as serious candidates will condition on the most likely decisive event $\lambda^i \in \Lambda$ such that $z_c^i \neq z_{c'}^i$.

The fourth property says: if a district's most likely decisive event, λ_d^1 , is of type $\lambda(3)$ or $\lambda(2')$, then voters must be conditioning on this event; if λ_d^1 is of type $\lambda(2)$, voters will be conditioning on it only if they are not indifferent between the two serious candidates.

4 Analysis of Benchmark Model

4.1 Extreme Policy Outcomes

The model sustains multiple equilibria for any distribution of preferences. In some of these equilibria, the moderate party m wins no seats. Nonetheless, the available empirical evidence suggests that a moderate party almost always wins at least one seat. In UK and Canadian elections over the past hundred years this has always been the case.²⁵ These

²⁴By restricting attention to strictly perfect equilibria we rule out knife edge cases where candidates are expected to get exactly the same share of votes.

²⁵Possible reasons might be that the moderate party has historically been a serious candidate in a given district or perhaps it has many non-strategic supporters there.

examples suggests that it might be of greater practical importance to focus on equilibria that feature this characteristic. In this sub-section, I focus on equilibria in which the moderate party always wins a seat. Notice that by the Corollary to Proposition 1 this does not restrict the set of elections I consider.

Much attention in the U.S. has focused on how a system with two polarised parties has led to policies which are far away from the median voter’s preferred point.²⁶ An open question is the degree to which policy outcomes reflect the preferences of voters in a legislative election with three parties. Let \tilde{t}_d be the expected position of the median voter in district d , and order the districts so that $\tilde{t}_1 < \dots < \tilde{t}_{\frac{D+1}{2}} < \dots < \tilde{t}_D$. Then, $\tilde{t}_{\frac{D+1}{2}}$ is the expected median voter in the median district. Furthermore, let $a_{lm} \equiv \frac{a_l + a_m}{2}$ and $a_{mr} \equiv \frac{a_m + a_r}{2}$. While extreme outcomes can always occur in stand-alone multi-candidate plurality elections, the proposition below gives a sufficient condition for the policy outcome to be the platform of the moderate party, a_m .

Proposition 2. *In any equilibrium in which the moderate party is expected to win at least one seat, if*

$$a_{lm} < \tilde{t}_{\frac{D+1}{2}} < a_{mr}$$

then the expected outcome is the moderate party’s platform a_m .

Proof. See Appendix B. □

The proposition stands in stark contrast to single district plurality elections. In a stand-alone plurality election it can always be that l and r are the serious candidates, so either a_l or a_r will be implemented. In a legislative election it takes much more to get non-moderate policies. For example, for a_l to come about it must be that (a) voters act as if their vote is pivotal in deciding between an l majority government and a coalition, and (b) a majority of voters in a majority of districts prefer policy a_l to the coalition policy, a_m .

This result gives a novel insight into multiparty legislative elections under plurality. In the U.K., until recently, a vote for the centrist Liberal Democrats (Lib Dems) has typically been considered a “wasted vote”.²⁷ The popular belief was that the Lib Dems were not a legitimate contender for government and so, even if they took a number of seats in parliament, they would not influence policy. As a result, centrist voters instead voted for either the Conservatives (right-wing) or Labour (left-wing).²⁸ My model shows that electing a Lib

²⁶See McCarty, Poole, and Rosenthal (2008) as well as Polborn and Krasa (2015).

²⁷See “What Future for the Liberal Democrats” by Lord Ashcroft, 2010.

²⁸My assumption that the Lib Dems are the centrist party is supported by polling data. Since 2006 YouGov have asked survey respondents to place the three main parties on a policy line. The Lib Dems have always been named as the centrist party. See <https://yougov.co.uk/news/2014/07/23/britains-changing-political-spectrum>

Dem candidate is far from a waste. Electing just one member of the Lib Dems to the legislature will be enough to moderate extreme policies unless voter preferences favour one of the non-centrist parties very much. In this benchmark case moderation leads to any coalition implementing Lib Dem policies in full. In Section 5 and Appendix C we'll see that a coalition doesn't implement the exact policy of the moderate party - nonetheless, coalition policies remain quite centrist and extreme outcomes are still mitigated. This result suggests that concerns about the average voter not being adequately represented in the U.K. or Canada are misplaced. If the Conservatives win a majority in parliament it must be that a majority of voters in a majority of districts prefer their policy to that of a moderate coalition. On the other hand, a moderate coalition can come about for any distribution of voter preferences.²⁹ Supporters of centrist parties - the Lib Dems in the U.K. and the Liberals in Canada - are therefore hugely advantaged by the current electoral system; it is the supporters of the non-centrist parties who are disadvantaged.

There are two concerns one might have about the proposition. The first is that it is phrased in terms of expected outcomes. The focus on expected policies and voters 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 $n \rightarrow \infty$ this probability goes to zero. The second concern is that it places a condition on the expected number of seats the moderate party wins, an endogenous variable. It is always possible that l and r are the serious candidates in every district, m wins no seats, and the final policy is extreme. These types of multiple equilibria are ever-present in voting games. However, even if $E(s_m) = 0$, the proposition still holds so long as m is a serious candidate in one relatively "close" district.³⁰

The median voter in the median district, $\tilde{t}_{\frac{D+1}{2}}$, is not the same as the median voter in the overall population, \tilde{t}_{med} . By definition \tilde{t}_{med} is the individual at the 50th percentile of the whole population distribution over $[-1,1]$. The "median of medians", $\tilde{t}_{\frac{D+1}{2}}$, instead will be somewhere between the 25th and 75th percentiles.³¹ Where exactly $\tilde{t}_{\frac{D+1}{2}}$ is in the interval depends on \mathbf{f} .³² Thus, we can have $\tilde{t}_{\frac{D+1}{2}} < \tilde{t}_{med}$ or $\tilde{t}_{\frac{D+1}{2}} \geq \tilde{t}_{med}$. The following proposition

²⁹For any \mathbf{f} , there are always equilibria where $z = a_m$ is the expected outcome. If support is strong for party r then an equilibrium in which districts focus on a $\lambda(a_l, a_m, a_m)$ decisive event will give $z = a_m$. Likewise, if l is popular then a focus on $\lambda(a_m, a_m, a_r)$ will give $z = a_m$.

³⁰ $E(s_m) > 0$ is a sufficient but not necessary condition for the proposition to hold. As can be seen in the proof of the proposition, all that is required is that $\mathbf{K}_{d_l}(\lambda(a_l, a_m, a_m))$ - the sequence of pivotal events with largest magnitude resulting in $\lambda_{d_l} = \lambda(a_l, a_m, a_m)$ - is a subset of $\mathbf{K}_{d_l}(\lambda(a_l, a_m, a_r))$.

³¹More specifically the lower and upper bounds are the $\frac{(D+1)25}{D}$ and $\frac{(3D-1)25}{D}$ percentiles. These bounds converge to the 25th and 75th percentiles as the number of districts increases.

³²The lower bound is the case where \mathbf{f} such that the bottom quartile of voter types are all situated in the same $D + 1$ districts, and within each of those districts these types are just barely in the majority. The upper bound corresponds to a \mathbf{f}' where the bottom three quartiles of voter types make up 100% of voters in

gives conditions under which \tilde{t}_{med} would be better off under a legislative election or single district election. This is equivalent to asking whether the median voter would be better off under a parliamentary or presidential system.

Proposition 3. *In any equilibrium in which the moderate party is expected to win at least one seat, we have the following:*

1. *If $\tilde{t}_{\frac{D+1}{2}}, \tilde{t}_{med} \in (a_{lm}, a_{mr})$, the utility of \tilde{t}_{med} is weakly higher in a legislative election than in a single district election.*
2. *If $\tilde{t}_{med} < a_{lm} < \tilde{t}_{\frac{D+1}{2}}$ or if $\tilde{t}_{\frac{D+1}{2}} < a_{mr} < \tilde{t}_{med}$, the utility of \tilde{t}_{med} is weakly higher in a single district election than in a legislative election.*
3. *In all other cases \tilde{t}_{med} may be better off under single district or legislative elections depending on which equilibrium is played.*

Proof. See Appendix B. □

The proposition can give us insights into when we should have a parliamentary or a presidential system. In countries where the median voter and the “median of medians” are both relatively moderate, a parliamentary system does a better job of representing the median voter. In countries where the median voter is extreme in one direction while the “median of medians” is either moderate or extreme in the other direction, then a presidential system would be better for the median voter. The reason is that legislative elections produce more moderate policy outcomes. This is desirable if the median voter is also moderate, but may not be best if he is extreme.

4.2 Misaligned Voting

When deciding which candidate to vote for, a voter must first consider (a) how the election of each candidate would map into his utility - so as to come up with a ranking over ballots - and then (b) whether to vote for his top-ranked alternative. This is true both in stand-alone district elections and in legislative elections. In stand-alone district elections (a) is simple - if party c is elected, policy a_c is implemented - it doesn't depend on the strategies of other players. In a legislative election (a) is more complicated. Here, a voter's preference over candidates in d depends on which candidates win the other $D - 1$ districts (because this determines the policy outcome of legislative bargaining). That is, preferences over ballots depend on the strategies of voters in other districts. While the calculation of

$D - 1$ districts, and in each of the remaining $D + 1$ districts they are just barely in the minority.

(a) differs in stand-alone and legislative elections, in both cases voters still face (b): whether to vote for their top-ranked alternative or not. A voter in d casts a *misaligned vote* (see Kawai and Watanabe (2013)) if he does not vote for his top-ranked alternative. Of course, in a legislative election, what constitutes a voter's top-ranked alternative will depend on the equilibrium played. More formally,

Definition. *An equilibrium with voting strategies σ^* exhibits **misaligned voting** in district d if there exists a voter type (t, d) and parties c and c' such that $\sigma_{t,d}^*(v_c) = 1$ and a victory in district d for party c' would give voters with type (t, d) a higher continuation expected payoff than a victory of party c .*

If a voter casts a misaligned vote, he is essentially giving up on his preferred candidate due to Duverger's law. In a stand-alone district election, the decision on whether to vote for his top-ranked candidate depends on whether that candidate is serious or not. With candidates l, m and r , whichever candidate is least likely to be pivotal will be abandoned by his supporters, leading to a two-party race. Every equilibrium will exhibit misaligned voting, either all types with $t < a_{lm}$, all types with $t > a_{mr}$, or those in the interval between will cast a misaligned vote. In a legislative election, a voter's preference within a district depends upon what he expects to happen nationally. Only by conditioning on the most likely decisive events can he know which local candidate he prefers. The decision on whether to vote for this top-ranked alternative is determined, as in the stand-alone case, by which are the serious candidates in the district. Nonetheless, Proposition 4 below shows that there are many equilibria of the legislative election in which there is no misaligned voting in any district.

Proposition 4. *For any distribution of voter preferences, there always exist equilibria that do not exhibit misaligned voting in any district. In any such equilibrium, either l or r is a serious candidate in fewer than $\frac{D-1}{2}$ districts.*

Proof. See Appendix B. □

The proposition can be understood as follows: A voter with $t > a_{lm}$ conditioning on a $\lambda(a_l, a_m, a_m)$ event will vote for whichever of m or r is serious in his district. However, his vote will be misaligned unless he is voting for his preferred alternative. For m and r to both be his preferred alternative he must be indifferent between them in *all* possible decisive events. He would not be indifferent in $\lambda(a_m, a_m, a_r)$ or $\lambda(a_l, a_m, a_r)$ events, for example. If party r is serious in fewer than $\frac{D-1}{2}$ districts then only $\lambda(a_l, a_m, a_m)$ events exist, which in turn means voters with $t > a_{lm}$ are indifferent between m and r . Voters with $t < a_{lm}$ are

not misaligned either - they simply vote for their preferred candidate, l . As we will see, this result is a special case of Proposition 7, where λ^1 is a $\lambda(2)$ event and no λ^2 event exists.

The caveat to Proposition 2 also applies here. That is, there always exist equilibria where party l or r is serious in more than half the districts. Nonetheless, the proposition gives us a clear prediction on when there will and will not be misaligned voting with three parties competing in a legislative election. It shows that the conventional wisdom - no misaligned with two candidates, always misaligned with three - is wide of the mark; whether there is misaligned voting or not depends on the strength of the non-centrist parties.

The proposition also has implications for the study of third-party entry into a two-party system. Suppose, as is plausible, that a newly formed party cannot become focal in many districts - maybe because they have limited resources, or because voters do not yet consider them a serious alternative. Either way, an entering third-party is likely to be weaker than the two established parties. Proposition 4 tells us that if a third party enters on the flanks of the two established parties, then there will be no misaligned voting and no effect on the policy outcome as long as this party is serious in fewer than half the districts. On the other hand, if a third party enters at a policy point in between the two established parties, this can shake up the political landscape. First of all, there will necessarily be misaligned voting. Second of all, the policy outcome will now depend on which equilibrium voters focus on - either a_l or a_m if $\tilde{t}_{\frac{D+1}{2}} < a_{lr}$ and a_m or a_r if $\tilde{t}_{\frac{D+1}{2}} > a_{lr}$. Success for the new party in just one district can radically change the policy outcome. The implication is that parties in a two-party system should be less concerned about the entry of fringe parties and more concerned about potential centrist parties stealing the middle ground.

5 Legislative Bargaining over Policy and Perks

While the model of bargaining over policy in the previous section is tractable, it lacks one of the key features of the government formation process: parties often bargain over perks of office such as ministerial positions as well as over policy. Here, as parties can trade off losses in the policy dimension for gains in the perks dimension, and vice versa, a larger set of policy outcomes are possible. This section will show that, nonetheless, the results of the benchmark model extend broadly to the case of bargaining over policy and perks.

The following legislative bargain model is due to Austen-Smith and Banks (1988) (henceforth ASB).³³ The party winning the most seats of the three begins the process by offering a policy outcome $y^1 \in [-1, 1]$ and a distribution of a fixed amount of transferable private

³³Other papers with bargaining over policy and perks include Diermeier and Merlo (2000) and Bandyopadhyay and Oak (2008).

benefits across the parties, $b^1 = (b_l^1, b_m^1, b_r^1) \in [0, B]^3$. It is assumed that B is large enough so that any possible governments can form, i.e. l can offer enough benefits to party r so as to overcome their ideological differences.³⁴ If the first proposal is rejected, the party with the second largest number of seats gets to propose (y^2, b^2) . If this is rejected, the smallest party proposes (y^3, b^3) . If no agreement has been reached after the third period, a caretaker government implements (y^0, b^0) , which gives zero utility to all parties.³⁵

At its turn to make a proposal, party c solves

$$\begin{aligned} & \max_{b_{c'}, y} B - b_{c'} - (y - a_c)^2 \\ & \text{subject to } b_{c'} - (y - a_{c'})^2 \geq W_{c'} \end{aligned} \quad (6)$$

where $W_{c'}$ is the continuation value of party c' and $W_{c''} + (y - a_{c''})^2 > W_{c'} + (y - a_{c'})^2$, so that the formateur makes the offer to whichever party is cheaper. Solving the game by backward induction, Austen-Smith and Banks (1988) show that a coalition government will always be made up of the largest party and the smallest party.³⁶ They solve for the equilibrium policy outcome, for any possible distance between a_l, a_m and a_r .

Table 1 shows the policy outcome for each seat distribution and distance between parties, where $\Delta_l \equiv a_m - a_l$ and $\Delta_r \equiv a_r - a_m$. I assume if two parties have exactly the same number of seats, a coin is tossed before the bargaining game begins to decide the order of play. So, if $s_l = s_r > s_m$, then with probability one-half, the game will play out as when $s_l > s_r > s_m$ and otherwise as $s_r > s_l > s_m$. The set of possible policy outcomes depends on the number of seats in the legislature, and on the distance between party policies. While there are many more decisive cases than when $B = 0$, they can be grouped into the three categories defined previously: $\lambda(2)$, $\lambda(2')$ and $\lambda(3)$ events.

| Seat Share | $3\Delta_r < \Delta_l$ | $2\Delta_r < \Delta_l \leq 3\Delta_r$ | $\Delta_r < \Delta_l \leq 2\Delta_r$ | $\Delta_l = \Delta_r$ | $\Delta_l < \Delta_r \leq 2\Delta_l$ | $2\Delta_l < \Delta_r \leq 3\Delta_l$ | $3\Delta_l < \Delta_r$ |
|-----------------------------|------------------------|---------------------------------------|--------------------------------------|-----------------------|--------------------------------------|---------------------------------------|------------------------|
| $s_l > (D-1)/2$ | a_l | a_l | a_l | a_l | a_l | a_l | a_l |
| $(D+1)/2 > s_l > s_r > s_m$ | a_{lm} | a_{lm} | a_{lm} | a_{lm} | a_{lm} | $2a_m - a_{lr}$ | a_l |
| $(D+1)/2 > s_l > s_m > s_r$ | a_{lr} | a_{lr} | a_{lr} | a_m | a_m | a_m | a_m |
| $s_m > s_l, s_r$ | a_m | a_m | a_m | a_m | a_m | a_m | a_m |
| $(D+1)/2 > s_r > s_m > s_l$ | a_m | a_m | a_m | a_m | a_{lr} | a_{lr} | a_{lr} |
| $(D+1)/2 > s_r > s_l > s_m$ | a_r | $2a_m - a_{lr}$ | a_{mr} | a_{mr} | a_{mr} | a_{mr} | a_{mr} |
| $s_r > (D-1)/2$ | a_r | a_r | a_r | a_r | a_r | a_r | a_r |

Table 1: Policy outcomes under ASB bargaining for any seat distribution and distance between parties.

³⁴This is the assumption made in the literature and in the original Austen-Smith and Banks (1988) paper. If B is not large enough then a coalition could never form without m , so we would return to the benchmark model.

³⁵This three-period protocol is a departure from the infinite horizon of the other bargaining rules I use. Nonetheless, it is the standard ASB model and so is used for easy comparisons to the literature.

³⁶Unless of course one party has a majority of seats, in which case it implements a_c and keeps all of B .

The following proposition shows that, even when parties can bargain over perks as well as policy, a non-centrist party will only win a majority if the median voter in the median district prefers its policy to that of the centrist party.

Proposition 5. *In any equilibrium under ASB bargaining in which the moderate party is expected to win at least one seat:*

1. *If $3\Delta_l > \Delta_r$, then the expected policy can be a_l only if $\tilde{t}_{\frac{D+1}{2}} < a_{lm}$.*
2. *If $3\Delta_r > \Delta_l$, then the expected policy can be a_r only if $\tilde{t}_{\frac{D+1}{2}} > a_{mr}$.*

Proof. See Appendix B. □

The conditions on Δ_l and Δ_r above are simply to rule out cases where party platforms a_l and a_r are exceptionally asymmetrically positioned vis-a-vis a_m . Given that restriction, we reaffirm the result of Proposition 2, that moderate coalitions will be the norm in legislative elections unless the population is heavily biased in favour of one of the non-centrist parties. Moreover, bargaining over perks and policy can make it even more difficult for a non-centrist party to gain a majority than in the benchmark case. Using the same steps as in the proof, one can show that if $E(s_l) > \frac{D-1}{2}$, $E(s_m) > 1$ and $\frac{D-1}{4} < E(s_r) < \frac{D-1}{2}$, the most likely decisive event for each district must be $\lambda(a_l, a_{lm}, a_{lm})$. Therefore, a party l majority could only come about if $\tilde{t}_{\frac{D+1}{2}} < \frac{a_l + a_{lm}}{2}$ - an even stricter requirement than that of the benchmark case. This result is noteworthy as in U.K. and Canadian elections party seat shares have tended to be in line with this case: one of the non-centrist parties wins a majority, the other wins more than a quarter of the seats, while the centrist party wins much less than a quarter. I do not carry out the welfare comparisons of Proposition 3 for this protocol. Here there are many more potential policy outcomes and many more decisive events, making welfare comparison intractable.

On the other dimension of interest bargaining over policy and perks does not perform as well; the restrictions needed to completely rule out misaligned voting are more severe than in the benchmark model. However, as Proposition 7 will show, there are many equilibria in which a large subset of districts have no misaligned voting.

Proposition 6. *For any distribution of preferences, under ASB bargaining, there always exist equilibria that do not exhibit misaligned voting in any district.*

1. *When $\Delta_l = \Delta_r$, there is no misaligned voting if either l or r is a serious candidate in fewer than $\frac{D-1}{4}$ districts.*
2. *When $\Delta_l < \Delta_r$, there is no misaligned voting if r is a serious candidate in fewer than $\frac{D-1}{4}$ districts.*

3. When $\Delta_l > \Delta_r$, there is no misaligned voting if l is a serious candidate in fewer than $\frac{D-1}{4}$ districts.

Proof. See Appendix B. □

The intuition is the same as in Proposition 4: when a non-centrist party is not serious in enough districts, there is no hope of it influencing the outcome of legislative bargaining. The threshold for relevance is lower than in the benchmark case. This is because the order of parties matters for the policy outcome under ASB bargaining. Once it is possible for party r to win $\frac{D-1}{4}$ districts, two distinct decisive events exist: $\lambda(a_l, a_m, a_m)$ and $\lambda(a_l, a_m, a_{lm})$. No matter which of these two events a district focuses on, and which two candidates are serious, some voters in the district will be casting misaligned votes.

When party l or r have serious candidates in more than $\frac{D-1}{4}$ districts we cannot rule out misaligned voting. This should not be surprising in a model with three differentiated parties and multiple voter types on a policy line. We know that in single plurality elections there will always be misaligned voting. What *is* surprising is that misaligned voting can ever be ruled out in a district. The following proposition gives conditions for equilibria with no misaligned voting in a subset of districts. The result is general and not specific to the bargaining rules examined in this paper. That is, if either of the two conditions hold for a given bargaining rule, then there will be no misaligned voting in that district.

Proposition 7. *There will be no misaligned voting in a district if either*

1. *The most likely decisive event λ_d^1 is a $\lambda(2')$ event where candidates c and c' are serious and $z_{c'}^1$ is preferred by no voter.*
2. *The most likely decisive event λ_d^1 is a $\lambda(2)$ event where candidates c and c' are serious, $z_c^1 = z_{c'}^1$, and all those voting v_c must have $u_t(z_c^i) > u_t(z_{c'}^i)$ in the next most likely decisive event $\lambda^i \in \Lambda$ in which $z_c^i \neq z_{c'}^i$.*

Proof. See Appendix B. □

The proposition is best understood by way of example. To see the first part, take a $\lambda(2')$ event, for example, $S_{-d} = (\frac{D-3}{2}, 2, \frac{D-3}{2})$. Electing l will give $s_l > s_r > s_m$ resulting in $z = a_{lm}$, electing r instead will give $s_r > s_l > s_m$ and bring about $z = a_{mr}$, while electing m will lead to a tie for first place between l and r . A coin toss will decide which of the two policies comes about, but *ex ante* voters' expectation is $E(a_{lm}, a_{mr})$. As voters have concave utility functions, every voter strictly prefers either a_{lm} or a_{mr} to the lottery over the pair. If this decisive event is the most likely (i.e. infinitely more likely than all others) and the district focuses on a race between l and r , nobody in the district is casting a misaligned vote.

To see the second part of the proposition, suppose the most likely decisive event is $S_{-d} = (\frac{D-3}{2}, 3, \frac{D-5}{2})$. Here, electing l or m gives a_{lm} while electing r brings about a coin toss and an expected policy $E(a_{lm}, a_{mr})$. Suppose further that the second most likely decisive event is $S_{-d} = (\frac{D-5}{2}, 3, \frac{D-3}{2})$, where electing m or r gives policy a_{mr} while electing l gives $E(a_{lm}, a_{mr})$. In the most likely event, all voters below a certain threshold will be indifferent between electing l and electing m . However, in the second most likely decisive event, all of these voters would prefer to elect l than m . Given that each decisive event is infinitely more likely to occur than a less likely decisive event, these voters need only consider the top two decisive events. Any voter who is indifferent between l and m in the most likely decisive event strictly prefers l in the second most likely. So, if the district focuses on a race between l and r there will be no misaligned voting. A special case of this second condition is when λ^1 is a $\lambda(2)$ event and no λ^2 event exists. This is what rules out misaligned voting in Proposition 4 and Proposition 6.

Proposition 7 is quite useful, as it holds for any bargaining rule. It allows me to show in Appendix C that, even though we cannot get results such as Proposition 4 and Proposition 6, we do not return to the stand-alone election case of “always misaligned voting”. Instead, there are again many equilibria in which a subset of districts have no misaligned voting.

6 Discussion

In this paper, I introduced and analysed a model of three-party competition in legislative elections under plurality rule. I showed that two properties of plurality rule - extreme policy outcomes and misaligned voting - are significantly reduced when the rule is used to elect a legislature. The degree to which these phenomena are reduced depends on the institutional setup - specifically, on how legislative bargaining occurs - but overall the results show that a plurality rule legislative electoral system reflects voter preferences much more than a single district model would suggest.

In the benchmark model, parties are perfectly patient and bargain only over policy. Two clear results emerged from this model. First, while an extreme policy can always come about in standard plurality elections, in my setting a non-centrist policy needs broad support in the electorate in order to be implemented; specifically, a majority of voters in a majority of districts must prefer the extreme policy to the moderate party’s policy. Second, while standard plurality elections with three distinct choices always exhibit misaligned voting, in my benchmark model this is the case only if each non-centrist party is a serious candidate in more than half the districts - otherwise there is no misaligned voting in any district.

The results of the benchmark model largely hold up under the other bargaining rules

considered: the non-centrist parties cannot win for any voter preferences (unlike in standard plurality elections), and there are always equilibria in which there is no misaligned voting (at least in a subset of districts).

In a 2011 referendum UK voters were asked to choose between plurality rule and Instant Run-off Voting (IRV) as a means of electing parliament.^{37,38} Supporters of IRV campaigned by appealing to the undesirable properties of plurality rule; claiming plurality protected a two party system of the non-centrist Conservative and Labour parties while the votes of millions of centrist voters were wasted. This paper shows that applying what we know about single district elections to legislative elections can lead to wrong conclusions. If we really want to know whether plurality is outperformed by IRV (or proportional representation for that matter) as an electoral system then we need to analyse the full legislative election game. Such a comparison between systems is beyond the scope of this paper and is left for future work.

In the remainder of this section I discuss the robustness of my modelling assumptions. First, if utility functions are concave rather than specifically quadratic, the benchmark model is unchanged. When bargaining also involves perks, the same is true as long as there are enough perks to allow a coalition between the left and right parties to form.³⁹ Second, if parties bargain by making demands rather than offers, as in Morelli (1999), the results will be the same as in the benchmark model.⁴⁰ Third, if instead of a Poisson model I assumed a fixed population size drawn from a multinomial distribution, the results of my model would still go through.⁴¹

A key assumption is that voters only care about the policy implemented in the legislature. If they also have preferences over who wins their local district, the results of the model no longer hold: the probability of being pivotal locally would outweigh any possible utility gain at the national level so that voters would only consider the local dimension. However, in Westminster systems, a Member of Parliament has no power to implement policy at a local level; he merely serves as an agent of his constituents: bringing up local issues in parliament,

³⁷In the referendum, UK Voters decided to retain plurality rule.

³⁸IRV is used to elect the Australian Parliament. Under IRV voters rank the candidates in their district on the ballot paper. If no candidate receives more than 50% of first preferences, those with the least first preference votes are eliminated and their second preferences are distributed to the remaining candidates. This continues until one candidate passes the 50% threshold and is elected.

³⁹If the perks are not large enough or parties don't value perks enough, a coalition will always involve the moderate party and we return to the simpler bargaining over policy case.

⁴⁰Whenever there is no clear majority, the head of state selects party m as the first mover, so the coalition policy will be $z = a_m$.

⁴¹Myerson (2000) shows that the magnitude of an event with a multinomial distribution is a simple transformation of its magnitude with a Poisson distribution. This transformation preserves the ordering of events and so this means that the ordering of sets of pivotal and decisive events in my model would remain unchanged, and therefore so would the equilibria.

helping constituents with housing authority claims, etc. So, if voters *do* have preferences over their local winner, it should only be on a common-value, valence dimension. If this were indeed the case, party policies would be irrelevant for how voters cast their ballots.

Finally, the assumption of perfect information is unrealistic in a real world election. The asymptotic elements of the model mean voters can perfectly rank the probabilities of certain events. In real life we are never that confident: polls may not be accurate, or more often, polls may not exist at the district level. Myatt (2007) and Fisher and Myatt (2014) have analysed single plurality elections with aggregate uncertainty over voters' intentions. While this paper abstracts from aggregate uncertainty for the sake of comparison with standard models, including greater uncertainty in a multi-district model is an important path for future research.

Appendix A: Poisson Properties

The number of voters in a district is a Poisson random variable n_d with mean n . The probability of having exactly k voters is $Pr[n_d = k] = \frac{e^{-n}n^k}{k!}$. 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 n_d with mean n . By the *decomposition property*, the number of voters with type in the subset $T \subset [-1, 1]$ is Poisson distributed with mean $\int_{t \in T} n f_d(t) dt$, and is independent of the number of players with types in any other disjoint sets.

The probability of a vote profile $x_d = (x_d(l), x_d(m), x_d(r))$ given voter strategies is

$$Pr[x_d | n\tau_d] = \prod_{c \in \{l, m, r\}} \frac{e^{-n\tau_d(c)} (n\tau_d(c))^{x_d(c)}}{x_d(c)!} \quad (7)$$

Its associated magnitude is

$$\mu(x_d) \equiv \lim_{n \rightarrow \infty} \frac{\log(Pr[x_d | n\tau_d])}{n} = \lim_{n \rightarrow \infty} \sum_{c \in \{l, m, r\}} \tau_d(c) \psi \left(\frac{x_d(c)}{n\tau_d(c)} \right) \quad (8)$$

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 n , the

magnitude of an event, $\mu(A_d)$, is:

$$\mu(A_d) \equiv \lim_{n \rightarrow \infty} \frac{\log(Pr[A_d])}{n} = \lim_{n \rightarrow \infty} \max_{x_d \in A_d} \sum_{c \in \{l, m, r\}} \tau_d(c) \psi\left(\frac{x_d(c)}{n\tau_d(c)}\right) \quad (9)$$

where $\psi(\theta) = \theta(1 - \log(\theta)) - 1$. That is, as $n \rightarrow \infty$, the magnitude of an event A_d is simply the magnitude of the most likely vote profile $x_d \in A_d$. The magnitude $\mu(A_d) \in [-1, 0]$ represents the speed at which the probability of the event goes to zero as $n \rightarrow \infty$; 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 $\mu(A_d) > \mu(A'_d)$, then their probability ratio converges to zero as $n \rightarrow \infty$.

$$\mu(A'_d) < \mu(A_d) \Rightarrow \lim_{n \rightarrow \infty} \frac{Pr[A'_d]}{Pr[A_d]} = 0 \quad (10)$$

It is possible that two distinct events have the same magnitude. In this case, we must use the offset theorem to compare their relative probabilities.

Offset Theorem Take two distinct events, $A_d \neq A'_d$ with the same magnitude, then

$$\mu(A_d) = \mu(A'_d) \Rightarrow \lim_{n \rightarrow \infty} \frac{Pr[A_d]}{Pr[A'_d]} = \phi \quad 0 < \phi < \infty \quad (11)$$

Suppose we have $\tau_d(c_1) > \tau_d(c_2) > \tau_d(c_3)$, so that the subscript denotes a party's expected ranking in terms of vote share. Maximising Equation 9 subject to the appropriate constraints we get

$$\begin{aligned} \mu(piv(i, j)) &= \mu(piv(j, i)) \quad \forall i, j \in C & (12) \\ \mu(c_1\text{-win}) &= 0 \\ \mu(c_2\text{-win}) = \mu(piv(c_1, c_2)) &= 2\sqrt{\tau_d(c_1)\tau_d(c_2)} - 1 + \tau_d(c_3) \\ \mu(c_3\text{-win}) = \mu(piv(c_2, c_3)) = \mu(piv(c_1, c_3)) &= 3\sqrt[3]{\tau_d(c_1)\tau_d(c_2)\tau_d(c_3)} - 1 \quad \text{if } \tau_d(c_1)\tau_d(c_3) < \tau_d(c_2)^2 \\ \mu(c_3\text{-win}) = \mu(piv(c_1, c_3)) &= 2\sqrt{\tau_d(c_1)\tau_d(c_3)} - 1 + \tau_d(c_2) \quad \text{if } \tau_d(c_1)\tau_d(c_3) > \tau_d(c_2)^2 \end{aligned}$$

With a magnitude of zero, by the corollary, the probability of candidate c_1 winning goes to 1 as n gets large. Also, as the magnitude of a pivotal event between c_1 and c_2 is greater than all other pivotal events, a pivotal event between c_1 and c_2 is infinitely more likely than a pivotal event between any other pair as n gets large.

Appendix B: Main Proofs

Proof of Lemma 1

To prove the lemma, I show that the magnitude theorem and its corollary extend to a multi-district setting.

Step 1: $\mu(\mathbf{x}) = \sum_d \mu(x_d)$.

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

$$Pr[\mathbf{x}|n\boldsymbol{\tau}] = \prod_{\substack{d \in D \\ c \in \{l, m, r\}}} \left(\frac{e^{-n\tau_d(c)} (n\tau_d(c))^{x_d(c)}}{x_d(c)!} \right) \quad (13)$$

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

$$\mu(\mathbf{x}) \equiv \lim_{n \rightarrow \infty} \frac{\log(Pr[\mathbf{x}|n\boldsymbol{\tau}])}{n} = \lim_{n \rightarrow \infty} \sum_{c \in \{l, m, r\}} \sum_{d \in D} \tau_d(c) \psi\left(\frac{x_d(c)}{n\tau_d(c)}\right) \quad (14)$$

Notice that the magnitude of a particular profile of votes across districts is simply the sum of the magnitudes in each district. This is because each district's realised profile of votes is independent of all other districts. So while $\mu(x_d) \in (-1, 0)$ in a single district, we have $\mu(\mathbf{x}) \in (-D, 0)$ when considering the profile of votes in all districts.

Step 2: *If a particular profile of votes across districts, \mathbf{x} , has a larger magnitude than another, \mathbf{x}' , the former is infinitely more likely to occur as $n \rightarrow \infty$.*

The rate at which $Pr[\mathbf{x}|n\boldsymbol{\tau}]$ goes to zero is $e^{n\mu(\mathbf{x})}$. Take another profile of votes across districts \mathbf{x}' where $\mu(\mathbf{x}') < \mu(\mathbf{x})$, then

$$\lim_{n \rightarrow \infty} \frac{Pr[\mathbf{x}'|n\boldsymbol{\tau}]}{Pr[\mathbf{x}|n\boldsymbol{\tau}]} = \lim_{n \rightarrow \infty} \frac{e^{n\mu(\mathbf{x}')}}{e^{n\mu(\mathbf{x})}} = \lim_{n \rightarrow \infty} e^{n(\mu(\mathbf{x}') - \mu(\mathbf{x}))} = 0 \quad (15)$$

As their probability ratio converges to zero, \mathbf{x} is infinitely more likely than \mathbf{x}' to occur as $n \rightarrow \infty$.

Step 3: Multi-District Magnitude Theorem. The magnitude theorem (Equation 9) shows that the magnitude of an event occurring in a given district (such as a tie for first) equals the magnitude of the most likely district vote profile in the set of district vote profiles comprising that event. Here I extend this to the multi-district case. Let $\mathbf{A} = (A_1, \dots, A_d, \dots, A_D)$ be a multi-district event, where each A_d is a particular district

event. Let $\bar{x}_d \in A_d = \operatorname{argmax}_{x_d} \sum_{c \in \{l, m, r\}} \tau_d(c) \psi\left(\frac{x_d(c)}{n\tau_d(c)}\right)$, that is, \bar{x}_d is the most likely district vote profile in A_d given τ_d . Then, we have

$$\mu(\mathbf{A}) = \sum_{d=1}^D \mu(A_d) = \sum_{d=1}^D \mu(\bar{x}_d) = \mu(\bar{\mathbf{x}}) \quad (16)$$

The first inequality follows from the independence of districts, the second equality follows from the magnitude theorem and the independence of districts, the third equality follows from Equation 14. Together they show that the single district magnitude theorem extends to a multi-district setting.

Following from this, and using Equation 15, we have that the corollary to the magnitude theorem also extends to the multi-district case. If $\mu(\mathbf{A}') < \mu(\mathbf{A})$, then

$$\lim_{n \rightarrow \infty} \frac{Pr[\mathbf{A}' | n\boldsymbol{\tau}]}{Pr[\mathbf{A} | n\boldsymbol{\tau}]} = \lim_{n \rightarrow \infty} \frac{e^{n\mu(\mathbf{A}')}}{e^{n\mu(\mathbf{A})}} = \lim_{n \rightarrow \infty} e^{n(\mu(\bar{\mathbf{x}}') - \mu(\bar{\mathbf{x}}))} = 0 \quad (17)$$

Step 4: Voter's Decision Problem

Let $\mathcal{A}(z, z') = \{\mathbf{A}(z, z'), \mathbf{A}'(z, z'), \mathbf{A}''(z, z'), \dots\}$ be the set of multi-district events such that a single vote in district d is pivotal and decisive between policies z and z' . There are six different cases in which a vote may be pivotal and decisive here: $piv_d(l, m)\lambda(z, z', z'')$, $piv_d(l, m)\lambda(z', z, z'')$, $piv_d(l, r)\lambda(z, z'', z')$, $piv_d(l, r)\lambda(z', z'', z)$, $piv_d(m, r)\lambda(z'', z, z')$, $piv_d(m, r)\lambda(z'', z', z)$. Each case can occur from many different multi-district events e.g $\mathbf{A}(z, z')$ may have $x_{d'}(l) > x_{d'}(m) > x_{d'}(r)$ and $x_{d''}(m) > x_{d''}(l) > x_{d''}(r)$ while $\mathbf{A}'(z, z')$ has the winners in d' and d'' reversed.⁴² Using Equation 17 and Equation 16 we get

$$\mu(\mathcal{A}(z, z')) = \mu(\bar{\mathbf{A}}(z, z')) = \mu(\bar{\mathbf{x}}(z, z')) \quad (18)$$

Where $\bar{\mathbf{A}}(z, z')$ is the event in $\mathcal{A}(z, z')$ with largest magnitude, and $\bar{\mathbf{x}}(z, z')$ is the vote profile in $\bar{\mathbf{A}}(z, z')$ with largest magnitude.

Let $Piv\Lambda_d \equiv \{\mathcal{A}(z, z'), \mathcal{A}(\tilde{z}, \tilde{z}'), \dots\}$ be the set of all distinct pivotal and decisive cases. By Equation 18, comparing the elements in $Piv\Lambda_d$ requires simply comparing the vote profiles with largest magnitude in each \mathcal{A} . From Equation 15 we know that if one vote profile has a larger magnitude than another it is infinitely more likely to occur as n goes to infinity. Thus, if $\mu(\mathcal{A}(z, z')) > \mu(\mathcal{A}(\tilde{z}, \tilde{z}'))$ for all $\mathcal{A}(\tilde{z}, \tilde{z}') \in Piv\Lambda_d$ then the former event is infinitely more likely to occur. This being the case, a voter who is not indifferent between z and z' need only condition his vote choice on the unique most likely vote profile $\bar{\mathbf{x}}(z, z') \in \bar{\mathbf{A}}(z, z') \in$

⁴²From the perspective of a voter in district d the exact seats which are won in other districts don't matter, only their number.

$\mathcal{A}(z, z') \in \text{Piv}\Lambda_d$. □

Proof of Proposition 1

I first prove that in any asymptotic strictly perfect equilibrium, each district is duvergerian. That is, the set of voters voting for a third candidate is of measure zero.

Step 1: A district is Duvergerian if there is a single pivotal and decisive event with largest magnitude.

By Lemma 1, if there is a unique pivotal and decisive event with largest magnitude in district d , then all voters will condition their vote choice on this event. Call this event $\lambda_d^i \text{piv}_d(c_1, c_2)$. Each voter type compares his gain of voting for the three candidates.

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_1}|n\tau)}{Pr[\lambda_d^i]Pr[\text{piv}_d(c_1, c_2)]} &= u_t(z_{c_1}^i) - u_t(z_{c_2}^i) \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_2}|n\tau)}{Pr[\lambda_d^i]Pr[\text{piv}_d(c_1, c_2)]} &= u_t(z_{c_2}^i) - u_t(z_{c_1}^i) \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_3}|n\tau)}{Pr[\lambda_d^i]Pr[\text{piv}_d(c_1, c_2)]} &= 0 \end{aligned}$$

Those with $u_t(z_{c_1}^i) > u_t(z_{c_2}^i)$ vote for c_1 , those with $u_t(z_{c_1}^i) < u_t(z_{c_2}^i)$ vote for c_2 . Voters with $u_t(z_{c_1}^i) = u_t(z_{c_2}^i)$ are of measure zero. Therefore, in any district with a single pivotal and decisive event with largest magnitude, only two candidates will receive positive vote shares.

Step 2: These Duvergerian equilibria are asymptotically strictly perfect. Consider a small deviation from the profile of expected vote shares τ . Is σ^* still a best response for every player type in every district? Suppose not. Then there must exist a set of types with positive measure who adopt a different strategy in this perturbed game. Note, however, from Equation 12 that the magnitude of each event in a district is continuous in τ_d . It follows that the magnitude of a multi-district event $\mu[\lambda_d^i \text{piv}_d(c_1, c_2)]$ is also continuous in τ . By assumption $\mu[\lambda_d^i \text{piv}_d(c_1, c_2)]$ is larger than the magnitude of any other pivotal and decisive event. Therefore, by the continuity of $\mu[\lambda_d^i \text{piv}_d(c_1, c_2)]$ in τ , we can find an ϵ perturbation of τ small enough such that $\mu[\lambda_d^i \text{piv}_d(c_1, c_2)]$ always remains the largest magnitude of any pivotal and decisive event. Each voter will thus vote as in Step 1 above, that is, the set of players who adopt a different strategy when expected vote shares are perturbed is of measure zero.

Step 3: If the pivotal and decisive event with largest magnitude in a district is not unique, non-Duvergerian equilibria may exist. Note that in any non-Duvergerian

| Largest Magnitude Decisive Events in d | Largest Magnitude Pivotal & Decisive Events in d | Case | Duv |
|--|---|------|-----|
| $\tau_d(c_1) > \tau_d(c_2) > \tau_d(c_3) > 0$ | | | |
| $\lambda_d^1, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_2)] = \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | i | Yes |
| $\lambda_d^1 \in \{\lambda(2 z_{c_1} = z_{c_2})\}, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)] > \mu[\lambda_d^1 piv_d(c_2, c_3)]$ | iv | No |
| | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)] = \mu[\lambda_d^1 piv_d(c_2, c_3)]$ | iv | No |
| | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_2, c_3)] > \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | ii | Yes |
| $\lambda_d^1, \lambda_d^2 \in \{\lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_2, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_2, c_3)]$ | iii | Yes |
| $\tau_d(c_1) = \tau_d(c_2) > \tau_d(c_3) > 0$ | | | |
| $\lambda_d^1, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_2)] = \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | i | Yes |
| $\lambda_d^1 \in \{\lambda(2 z_{c_1} = z_{c_2})\}, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)] = \mu[\lambda_d^1 piv_d(c_2, c_3)]$ | iv | No |
| | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_2, c_3)] > \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | ii | No |
| $\lambda_d^1, \lambda_d^2 \in \{\lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_2, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_2, c_3)]$ | iii | No |
| $\tau_d(c_1) > \tau_d(c_2) = \tau_d(c_3) > 0$ | | | |
| $\lambda_d^1, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_2)] = \mu[\lambda_d^1 piv_d(c_1, c_3)] > \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | iv | No |
| | $\mu[\lambda_d^1 piv_d(c_1, c_2)] = \mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)] = \mu[\lambda_d^2 piv_d(c_1, c_3)]$ | v | No |
| $\lambda_d^1 \in \{\lambda(2 z_{c_1} = z_{c_2})\}, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)] = \mu[\lambda_d^2 piv_d(c_1, c_3)]$ | iv | No |
| $\tau_d(c_1) = \tau_d(c_2) = \tau_d(c_3) > 0$ | | | |
| $\lambda_d^1, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_2, c_3)] > \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | ii | No |
| | $\mu[\lambda_d^1 piv_d(c, c')] = \mu[\lambda_d^2 piv_d(c, c')] \forall c, c'$ | v | No |
| $\lambda_d^1 \in \{\lambda(2 z_{c_1} = z_{c_2})\}, \lambda_d^2 \in \{\Lambda \setminus \lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_2, c_3)] > \mu[\lambda_d^2 piv_d(c_1, c_2)]$ | ii | No |
| | $\mu[\lambda_d^1 piv_d(c, c')] = \mu[\lambda_d^2 piv_d(c, c')] \forall c, c'$ | v | No |
| $\lambda_d^1, \lambda_d^2 \in \{\lambda(2 z_{c_1} = z_{c_2})\}$ | $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_2, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_2, c_3)]$ | iii | No |

Table 2: Cases where $\tau_1, \tau_2, \tau_3 > 0$ and pivotal & decisive event with largest magnitude is not unique

equilibrium we must have $\sigma_{t,d}(c_1), \sigma_{t',d}(c_2), \sigma_{t'',d}(c_3) > 0$ for some $t, t', t'' \in [-1, 1]$. By Equation 1 this means $\tau_d(c_1), \tau_d(c_2), \tau_d(c_3) > 0$ must also hold. Therefore I examine the various cases where both (i) $\tau_d(c_1), \tau_d(c_2), \tau_d(c_3) > 0$ and (ii) the largest magnitude pivotal and decisive event is not unique. These different events are listed in Table 2. The first column shows the type of decisive events which have largest magnitude. The set $\lambda(2|z_{c_1} = z_{c_2})$ consists of those $\lambda(2)$ decisive events where electing either of the front-runners would bring about the same policy. The third column splits the various events into five cases. The final shows - as I now prove - that case i is duvergerian, cases ii, iii may not be duvergerian, while iv and v are non-duvergerian.

Step3a: Case i is Duvergerian

i In each case $\mu[\lambda_d^1 \text{piv}_d(c_1, c_2)] = \mu[\lambda_d^2 \text{piv}_d(c_1, c_2)]$.

By the offset theorem (Equation 11) we have $\lim_{n \rightarrow \infty} \frac{\text{Pr}[\lambda_d^2 \text{piv}_d(c_1, c_2)]}{\text{Pr}[\lambda_d^1 \text{piv}_d(c_1, c_2)]} \equiv \phi \in (0, 1]$. Note $\phi > 1$ is not possible because, by assumption $\mu(\lambda_d^1) \geq \mu(\lambda_d^2)$. Therefore we have:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_1} | n\tau)}{\text{Pr}[\lambda_d^1] \text{Pr}[\text{piv}_d(c_1, c_2)]} &= u_t(z_{c_1}^1) - u_t(z_{c_2}^1) + \phi[u_t(z_{c_1}^2) - u_t(z_{c_2}^2)] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_2} | n\tau)}{\text{Pr}[\lambda_d^1] \text{Pr}[\text{piv}_d(c_1, c_2)]} &= u_t(z_{c_2}^1) - u_t(z_{c_1}^1) + \phi[u_t(z_{c_2}^2) - u_t(z_{c_1}^2)] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_3} | n\tau)}{\text{Pr}[\lambda_d^1] \text{Pr}[\text{piv}_d(c_1, c_2)]} &= 0 \end{aligned} \quad (19)$$

Those with $u_t(z_{c_1}^1) + \phi u_t(z_{c_1}^2) > u_t(z_{c_2}^1) + \phi u_t(z_{c_2}^2)$ vote for c_1 , those with $u_t(z_{c_1}^1) + \phi u_t(z_{c_1}^2) < u_t(z_{c_2}^1) + \phi u_t(z_{c_2}^2)$ vote for c_2 . Indifferent voters are of measure zero.

Step3b: Cases ii, iii, iv and v have non-Duvergerian equilibria

ii In each case $\mu[\lambda_d^1 \text{piv}_d(c_1, c_3)] = \mu[\lambda_d^1 \text{piv}_d(c_2, c_3)] > \mu[\lambda_d^2 \text{piv}_d(c_1, c_2)]$.

Here the event $\lambda_d^1 \text{piv}_d(c_1, c_2)$ is not relevant because $z_{c_1}^1 = z_{c_2}^1$. The magnitudes of $\text{piv}_d(c_1, c_3)$ and $\text{piv}_d(c_2, c_3)$ are the same because the most likely event in which they occur is when all three candidates get exactly the same number of votes. However, the events consist of more than just this event. In an event $\text{piv}_d(c_1, c_3)$, candidate c_2 must have the same or fewer votes than the others. Similarly, in an event $\text{piv}_d(c_2, c_3)$, candidate c_1 must have the same or fewer votes than the others. We can once again use the offset theorem, now letting $\phi \equiv \lim_{n \rightarrow \infty} \frac{\text{Pr}[\lambda_d^1 \text{piv}_d(c_2, c_3)]}{\text{Pr}[\lambda_d^1 \text{piv}_d(c_1, c_3)]}$. If $\tau_d(c_1) > \tau_d(c_2)$ then, by the decomposition property, for any given number of votes c_3 has, c_1 is always more likely to have more votes than c_2 . Therefore, it must be that $\phi < 1$. If instead $\tau_d(c_1) = \tau_d(c_2)$ then both c_1 and c_2 are equally likely to tie for first with c_3 , giving $\phi = 1$. The gain functions below show the decision facing each voter type.

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_1} | n\tau)}{\text{Pr}[\lambda_d^1] \text{Pr}[\text{piv}_d(c_1, c_3)]} &= u_t(z_{c_1}^1) - u_t(z_{c_3}^1) \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_2} | n\tau)}{\text{Pr}[\lambda_d^1] \text{Pr}[\text{piv}_d(c_1, c_3)]} &= \phi[u_t(z_{c_1}^1) - u_t(z_{c_3}^1)] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_3} | n\tau)}{\text{Pr}[\lambda_d^1] \text{Pr}[\text{piv}_d(c_1, c_3)]} &= u_t(z_{c_3}^1) - u_t(z_{c_1}^1) + \phi[u_t(z_{c_3}^1) - u_t(z_{c_1}^1)] \end{aligned} \quad (20)$$

If $\phi < 1$ then any voter with $u_t(z_{c_1}^1) > u_t(z_{c_3}^1)$ votes for c_1 while any voter with $u_t(z_{c_3}^1) < u_t(z_{c_1}^1)$ votes for c_3 . Voting for c_2 is dominated by voting for c_1 . With $\phi < 1$ there is no multi-candidate support. If $\phi = 1$ any voter with $u_t(z_{c_1}^1) > u_t(z_{c_3}^1)$ is indifferent between

voting for c_1 and c_2 while any voter with $u_t(z_{c_3}^1) < u_t(z_{c_1}^1)$ votes for c_3 . Thus, we have a non-duvergerian equilibrium when $\tau_d(c_1) = \tau_d(c_2) \geq \tau_d(c_3)$.

iii In each case $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_2, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_2, c_3)]$.

Letting $\lim_{n \rightarrow \infty} \frac{Pr[\lambda_d^2]}{Pr[\lambda_d^1]} \equiv \phi^\lambda \in (0, 1]$ and $\lim_{n \rightarrow \infty} \frac{Pr[piv_d(c_2, c_3)]}{Pr[piv_d(c_1, c_3)]} \equiv \phi^{piv} \in (0, 1]$ we have:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_1}|n\tau)}{Pr[\lambda_d^1]Pr[piv_d(c_1, c_3)]} &= u_t(z_{c_1}^1) - u_t(z_{c_3}^1) + \phi^\lambda[u_t(z_{c_1}^2) - u_t(z_{c_3}^2)] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_2}|n\tau)}{Pr[\lambda_d^1]Pr[piv_d(c_1, c_3)]} &= \phi^{piv}[u_t(z_{c_1}^1) - u_t(z_{c_3}^1) + \phi^\lambda[u_t(z_{c_1}^2) - u_t(z_{c_3}^2)]] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_3}|n\tau)}{Pr[\lambda_d^1]Pr[piv_d(c_1, c_3)]} &= (1 + \phi^{piv})[u_t(z_{c_3}^1) - u_t(z_{c_1}^1) + \phi^\lambda[u_t(z_{c_3}^2) - u_t(z_{c_1}^2)]] \end{aligned} \quad (21)$$

If $\phi^{piv} < 1$ any voter with $u_t(z_{c_1}^1) + \phi^\lambda u_t(z_{c_1}^2) > u_t(z_{c_3}^1) + \phi^\lambda u_t(z_{c_3}^2)$ votes for c_1 while any voter with $u_t(z_{c_1}^1) + \phi^\lambda u_t(z_{c_1}^2) < u_t(z_{c_3}^1) + \phi^\lambda u_t(z_{c_3}^2)$ votes for c_3 . Thus, with $\phi^{piv} < 1$ there is no multi-candidate support. If $\phi^{piv} = 1$ any voter with $u_t(z_{c_1}^1) + \phi^\lambda u_t(z_{c_1}^2) > u_t(z_{c_3}^1) + \phi^\lambda u_t(z_{c_3}^2)$ is indifferent between voting for c_1 and c_2 while any voter with $u_t(z_{c_1}^1) + \phi^\lambda u_t(z_{c_1}^2) < u_t(z_{c_3}^1) + \phi^\lambda u_t(z_{c_3}^2)$ votes for c_3 . Thus, we have a non-duvergerian equilibrium when $\tau_d(c_1) = \tau_d(c_2) \geq \tau_d(c_3)$.

iv In each case either $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)]$ or $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^1 piv_d(c_1, c_2)]$.

I show the case where $\mu[\lambda_d^1 piv_d(c_1, c_3)] = \mu[\lambda_d^2 piv_d(c_1, c_2)]$ is the largest magnitude event (the other case is almost identical). Note that in this case we must have $\lambda_d^1 \in \{\lambda(2|z_{c_1} = z_{c_2})\}$, therefore $z_{c_1}^1 = z_{c_2}^1$.

Letting $\lim_{n \rightarrow \infty} \frac{Pr[\lambda_d^2]}{Pr[\lambda_d^1]} \equiv \phi^\lambda$ and $\lim_{n \rightarrow \infty} \frac{Pr[piv_d(c_1, c_3)]}{Pr[piv_d(c_1, c_3)]} \equiv \phi^{piv}$ we have:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_1}|n\tau)}{Pr[\lambda_d^2]Pr[piv_d(c_1, c_2)]} &= \phi^\lambda[u_t(z_{c_1}^2) - u_t(z_{c_2}^2)] + \phi^{piv}[u_t(z_{c_1}^1) - u_t(z_{c_3}^1)] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_2}|n\tau)}{Pr[\lambda_d^2]Pr[piv_d(c_1, c_2)]} &= \phi^\lambda[u_t(z_{c_2}^2) - u_t(z_{c_1}^2)] \\ \lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_3}|n\tau)}{Pr[\lambda_d^2]Pr[piv_d(c_1, c_2)]} &= \phi^{piv}[u_t(z_{c_3}^1) - u_t(z_{c_1}^1)] \end{aligned} \quad (22)$$

Manipulating these expressions, one can show that there exist values of $z_{c_1}^1, z_{c_3}^1, z_{c_1}^2, z_{c_2}^2$ such that no action is dominated for all voter types. Thus, non-Duvergerian equilibria can exist.

v In each case we have (at least) $\mu[\lambda_d^1 \text{piv}_d(c_1, c_2)] = \mu[\lambda_d^2 \text{piv}_d(c_1, c_2)] = \mu[\lambda_d^1 \text{piv}_d(c_1, c_3)] = \mu[\lambda_d^2 \text{piv}_d(c_1, c_3)]$.

Here we always have $\Pr[\text{piv}_d(c_1, c_2)] = \Pr[\text{piv}_d(c_1, c_3)]$. Below I show the gain functions for $\mu[\lambda_d^1 \text{piv}_d(c, c')] = \mu[\lambda_d^2 \text{piv}_d(c, c')]$ for all c, c' . Letting $\lim_{n \rightarrow \infty} \frac{\Pr[\lambda_d^2]}{\Pr[\lambda_d^1]} \equiv \phi^\lambda$ and $\lim_{n \rightarrow \infty} \frac{\Pr[\text{piv}_d(c_1, c_3)]}{\Pr[\text{piv}_d(c_1, c_2)]} \equiv \phi^{\text{piv}(1,3)}$ and $\lim_{n \rightarrow \infty} \frac{\Pr[\text{piv}_d(c_2, c_3)]}{\Pr[\text{piv}_d(c_1, c_3)]} \equiv \phi^{\text{piv}(2,3)}$ we have:

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_1}|n\tau)}{\Pr[\lambda_d^1] \Pr[\text{piv}_d(c_1, c_2)]} &= u_t(z_{c_1}^1) - u_t(z_{c_2}^1) + \phi^{\text{piv}(1,3)}[u_t(z_{c_1}^1) - u_t(z_{c_3}^1)] \\
&\quad + \phi^\lambda[u_t(z_{c_1}^2) - u_t(z_{c_2}^2) + \phi^{\text{piv}(1,3)}[u_t(z_{c_1}^2) - u_t(z_{c_3}^2)]] \\
\lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_2}|n\tau)}{\Pr[\lambda_d^1] \Pr[\text{piv}_d(c_1, c_2)]} &= u_t(z_{c_2}^1) - u_t(z_{c_1}^1) + \phi^{\text{piv}(2,3)}[u_t(z_{c_2}^1) - u_t(z_{c_3}^1)] \\
&\quad + \phi^\lambda[u_t(z_{c_2}^2) - u_t(z_{c_1}^2)] + \phi^{\text{piv}(2,3)}[u_t(z_{c_2}^2) - u_t(z_{c_3}^2)] \\
\lim_{n \rightarrow \infty} \frac{G_{t,d}(v_{c_3}|n\tau)}{\Pr[\lambda_d^1] \Pr[\text{piv}_d(c_1, c_2)]} &= \phi^{\text{piv}(1,3)}[u_t(z_{c_3}^1) - u_t(z_{c_1}^1)] + \phi^{\text{piv}(2,3)}[u_t(z_{c_3}^1) - u_t(z_{c_2}^1)] \\
&\quad + \phi^\lambda[\phi^{\text{piv}(1,3)}[u_t(z_{c_3}^2) - u_t(z_{c_1}^2)] + \phi^{\text{piv}(2,3)}[u_t(z_{c_3}^2) - u_t(z_{c_2}^2)]]
\end{aligned} \tag{23}$$

Manipulating these expressions, one can show that there exist values of $z_{c_1}^1, z_{c_2}^1, z_{c_3}^1, z_{c_1}^2, z_{c_2}^2, z_{c_3}^2$ such that no action is dominated for all voter types. Thus, non-Duvergerian equilibria can exist.

Step 4: These non-Duvergerian equilibria are not asymptotically strictly perfect.

As $n \rightarrow \infty$, if there is a unique most likely event in which his vote is both pivotal and decisive over two distinct policy outcomes, then a voter need only condition his choice on this single event. I now show that for some perturbation of τ in each of the duvergerian cases above, the best response of some player types changes, returning us to duvergerian equilibria. Thus non-duvergerian equilibria are not asymptotically strictly perfect.

- *Case ii:* Non-Duvergerian equilibria arise here only when $\tau_d(c_1) = \tau_d(c_2) \geq \tau_d(c_3)$. A perturbation of τ which yields $\tau_d(c_1) > \tau_d(c_2)$ will give $\phi < 1$ in Equation 20.
- *Case iii:* As above, non-Duvergerian equilibria arise here only when $\tau_d(c_1) = \tau_d(c_2) \geq \tau_d(c_3)$. A perturbation of τ which yields $\tau_d(c_1) > \tau_d(c_2)$ will give $\phi^{\text{piv}} < 1$ in Equation 21.
- *Case iv:* In each case here either $\lambda_d^1 \text{piv}_d(c_1, c_2)$ or $\lambda_d^2 \text{piv}_d(c_1, c_2)$ is one of the events with joint largest magnitude. A perturbation of τ which reduces $\tau_d(c_3)$ will give

$\mu[piv_d(c_1, c_2)] > \mu[piv_d(c_1, c_3)], \mu[piv_d(c_2, c_3)]$ and therefore $\phi^{piv} = 0$ in Equation 22. This means voting for c_3 is dominated for all players, returning us to a duvergerian equilibrium.

- *Case v:* Each case here has both $\lambda_d^1 piv_d(c_1, c_2)$ and $\lambda_d^2 piv_d(c_1, c_2)$ as two of the many events with joint largest magnitude. A perturbation of τ which reduces $\tau_d(c_3)$ will give $\mu[piv_d(c_1, c_2)] > \mu[piv_d(c_1, c_3)], \mu[piv_d(c_2, c_3)]$ and therefore $\phi^{piv(1,3)} = \phi^{piv(2,3)} = 0$ in Equation 23. Those with $u_t(z_{c_1}^1) + \phi^\lambda u_t(z_{c_1}^2) > u_t(z_{c_2}^1) + \phi^\lambda u_t(z_{c_2}^2)$ vote for c_1 , those with $u_t(z_{c_1}^1) + \phi^\lambda u_t(z_{c_1}^2) < u_t(z_{c_2}^1) + \phi^\lambda u_t(z_{c_2}^2)$ vote for c_2 , while no player type votes for the dominated choice c_3 . Indifferent voters are of measure zero.

Step 2: Multiple Equilibria always exist.

A simple example proves the existence of multiple pure strategy equilibria for any bargaining rule where a majority is needed to implement a policy z . For any \mathbf{f} , suppose the right party is never serious in any district. All races will be between l and m . We will have $E(S_{-d_l}) = (k - 1, D - k, 0)$ and $E(S_{-d_m}) = (k, D - k - 1, 0)$ for some $k \in (0, D)$. As party r receives no votes, every district must be conditioning on the same decisive event $(\frac{D-1}{2}, \frac{D-1}{2}, 0)$. When conditioning on this λ , voters will face a choice between a_l , a_m , and a third policy which would come about if r wins a seat. All the districts focusing on races between l and m is an equilibrium as no individual would deviate and vote for r . Similarly, it is possible that all districts ignore the left party, and so an equilibrium will have every district conditioning on $(0, \frac{D-1}{2}, \frac{D-1}{2})$, or that all districts ignore the moderate party and all condition on $(\frac{D-1}{2}, 0, \frac{D-1}{2})$. These three equilibria always exist for any majoritarian bargaining rule. \square

Proof of Corollary to Proposition 1

By Proposition 1, an equilibrium always exists in which l and m are the serious candidates in every district. Party m will be the expected winner in none of these districts only if the median voter in every district prefers a_l to a_m . Suppose this is the case. By Proposition 1, another equilibrium always exists in which m and r are the serious candidates in every district. Here voters face a choice between a_m and a_r . As the median voter in each district prefers a_l to a_m , they must each also prefer a_m to a_r . Thus, if this equilibrium is played, party m will win every seat. Thus by focusing on either one of these two equilibria we can ensure, for any distribution of voter preferences, that party m always wins at least 1 seat. \square

Lemma 2

A sufficient condition for $\mu(\lambda_d^i(z^i)) > \mu(\lambda_d^j(z^j))$ is that the set of pivotal events with largest magnitude required for $\lambda_d^i(z^i)$ to occur, $\mathbf{K}_d(z^i)$, is a subset of those required for $\lambda_d^j(z^j)$ to occur, $\mathbf{K}_d(z^j)$.

Proof of Lemma 2

From Equation 18 we have $\mu(\lambda_d^i(z^i)) = \sum_{d'=1}^{D-1} \mu(\bar{x}_{d'}) \in (-D+1, 0)$. That is the magnitude of a particular decisive event, from the point of view of district d is simply the sum of the magnitudes of the most likely vote profiles in each district $d' \neq d$ which bring about that decisive event. We can use $\mu(c_1 - \text{win}) = 0$ and $\mu(c_2 - \text{win}) = \mu(\text{piv}_{d'}(c_1, c_2))$ from Equation 12 and re-write

$$\mu(\lambda_d^i(z^i)) = \sum_{d'=1}^k \mu(\text{piv}_{d'}(c_1, c_2)) \in (-k, 0)$$

Where there are k districts in which the expected winner loses. Let $\mathbf{K}_d(z^i)$ be the set of these k pivotal events. Let another decisive event for district d be z^j with magnitude

$$\mu(\lambda_d^j(z^j)) = \sum_{d'=1}^{k'} \mu(\text{piv}_{d'}(c_1, c_2)) \in (-k', 0)$$

Where there are k' districts in which the expected winner loses. Let $\mathbf{K}_d(z^j)$ be the set of these k' pivotal events. Furthermore let $\mathbf{K}_d(z^i) \subset \mathbf{K}_d(z^j)$. As magnitudes are negative numbers, it must be that

$$\sum_{d'=1}^k \mu(\text{piv}_{d'}(c_1, c_2)) > \sum_{d'=1}^k \mu(\text{piv}_{d'}(c_1, c_2)) + \sum_{d'=k+1}^{k'} \mu(\text{piv}_{d'}(c_1, c_2))$$

or

$$\mu(\lambda_d^i(z^i)) > \mu(\lambda_d^j(z^j)) \tag{24}$$

Thus if $\mathbf{K}_d(z^i) \subset \mathbf{K}_d(z^j)$ then $\mu(\lambda_d^i(z^i)) > \mu(\lambda_d^j(z^j))$. \square

Corollary to Lemma 2

For a given $E(S_{-d})$ and Λ , we have $\lambda_d(z) = \lambda_d^1$ if $\forall \lambda(z') \neq \lambda(z) \in \Lambda$ we have $\mathbf{K}_d(z) \subset \mathbf{K}_d(z')$.

Proof of Proposition 2

From Proposition 0, for $E(z) = a_l$ we must have an expected majority for the left party, $E(s_l) > \frac{D-1}{2}$. By Proposition 1, each d_l district must have either l and m or l and r as the serious candidates. Therefore, also by Proposition 1, if a d_l district has $\lambda_{d_l}^1 = \lambda(a_l, a_m, a_m)$, it must be conditioning on this decisive event. If a district d is conditioning on $\lambda(a_l, a_m, a_m)$, the expected winner will be party l if the expected median voter prefers a_l to a_m , that is if $\tilde{t}_d < a_{lm}$. As party l is the expected winner in d_l districts, any d_l district conditioning on $\lambda(a_l, a_m, a_m)$ must have $\tilde{t}_d < a_{lm}$. If each d_l has $\lambda_{d_l}^1 = \lambda(a_l, a_m, a_m)$ and $E(s_l) > \frac{D-1}{2}$ then $\tilde{t}_{\frac{D+1}{2}} < a_{lm}$.

All that remains is to show that when $E(z) = a_l$ and $E(s_m) > 0$ each d_l has $\lambda_{d_l}^1 = \lambda(a_l, a_m, a_m)$. I show that case of $E(s_m) = 1$ as the case of $E(s_m) > 1$ is analogous.

Let $\mathbf{K}_{d_l}(\lambda(a_m, a_m, a_r))$ be the sequence of pivotal events with largest magnitude resulting in $\lambda_{d_l} = \lambda(a_m, a_m, a_r)$. Let $\mathbf{K}_{d_l}(\lambda(a_l, a_m, a_r))$ be the sequence of pivotal events with largest magnitude resulting in $\lambda_{d_l} = \lambda(a_l, a_m, a_r)$. Then given $E(S_{-d_l}) = (\frac{D-1}{2} + k, 1, \frac{D-3}{2} - k)$ it must be that $\mathbf{K}_{d_l}(\lambda(a_l, a_m, a_m)) \subset \mathbf{K}_{d_l}(\lambda(a_m, a_m, a_r))$ and $\mathbf{K}_{d_l}(\lambda(a_l, a_m, a_m)) \subset \mathbf{K}_{d_l}(\lambda(a_l, a_m, a_r))$. That is, the pivotal events needed to move from $E(S_{-d_l})$ to $\lambda(a_l, a_m, a_m)$ are a subset of those needed to move from $E(S_{-d_l})$ to $\lambda(a_m, a_m, a_r)$ or $\lambda(a_l, a_m, a_r)$. By the Corollary to Lemma 2 this means $\lambda_d^1 = \lambda(a_l, a_m, a_m)$.

It is analogous to show that a necessary condition for $E(z) = a_r$ is $\tilde{t}_{\frac{D+1}{2}} > a_{mr}$. \square

Proof of Proposition 3

Case 1: $\tilde{t}_{\frac{D+1}{2}}, \tilde{t}_{med} \in (a_{lm}, a_{mr})$.

In a legislative election with $E(s_m) > 0$ (for possibly exogenous reasons) the only two decisive events are $\lambda(a_l, a_m, a_m)$ and $\lambda(a_m, a_m, a_r)$. Given, $\tilde{t}_{\frac{D+1}{2}} \in (a_{lm}, a_{mr})$, in each case the winning policy will be a_m . Instead, in a single district election, there are three possible equilibria in which the serious candidates are (l, m) , (m, r) , and (l, r) respectively. Given, $\tilde{t}_{med} \in (a_{lm}, a_{mr})$, policy a_m will be implemented in the first two cases. When the race is between l and r , the winning policy will be either a_l or a_r - both of which \tilde{t}_{med} prefers less than a_m . Thus, \tilde{t}_{med} strictly prefers a legislative election to a single district election if there is a chance that (l, r) is focal, and weakly prefers a legislative election otherwise.

Case 2: $\tilde{t}_{med} < a_{lm} < \tilde{t}_{\frac{D+1}{2}}$ or if $\tilde{t}_{\frac{D+1}{2}} < a_{mr} < \tilde{t}_{med}$.

I show the case of $\tilde{t}_{\frac{D+1}{2}} < a_{mr} < \tilde{t}_{med}$ as the other is symmetric.

First, suppose $\tilde{t}_{\frac{D+1}{2}} \in (a_{lm}, a_{mr})$. As above, in a legislative election with $E(s_m) > 0$ the only two decisive events are $\lambda(a_l, a_m, a_m)$ and $\lambda(a_m, a_m, a_r)$. In each case the winning policy will be a_m . In a single district election, the winning policies in the races (l, m) , (m, r) , and

(l, r) are a_m , a_r and a_r respectively. As \tilde{t}_{med} has $a_r \succ a_m \succ a_l$ he will be weakly better off under a single district election, and will be strictly so if there is a chance that the focal race is not (l, m) . Next suppose $\tilde{t}_{\frac{D+1}{2}} < a_{lm}$. In the decisive legislative events $\lambda(a_l, a_m, a_m)$ and $\lambda(a_m, a_m, a_r)$ the winning policies are a_l and a_m respectively. As before, in a single district election, the winning policies in the races (l, m) , (m, r) , and (l, r) are a_m , a_r and a_r . As \tilde{t}_{med} will be weakly better off under a single district election if $\lambda(a_m, a_m, a_r)$ is focal in a legislative election while (l, m) is focal in a single district election. In any other case, he will be strictly better off under a single district election.

Case 3: In all other cases the welfare effect is ambiguous.

In the remainder of cases we either have $\tilde{t}_{\frac{D+1}{2}} < a_{lm}$ and $\tilde{t}_{med} < a_{mr}$; or else $\tilde{t}_{\frac{D+1}{2}} > a_{mr}$ and $\tilde{t}_{med} > a_{lm}$. I prove the former case as the latter is symmetric.

First, suppose $\tilde{t}_{med} \in (a_{lm}, a_{mr})$. In the decisive legislative events $\lambda(a_l, a_m, a_m)$ and $\lambda(a_m, a_m, a_r)$ the winning policies are a_l and a_m respectively. In a single district election, the winning policy is a_m in the races (l, m) and (m, r) , while the winning policy in (l, r) is either a_l or a_r depending on whether $\tilde{t}_{med} < a_{lr}$ or not. Under both legislative and single district elections, the preferred policy of \tilde{t}_{med} (a_m) may or may not be the policy outcome, depending on which equilibrium is played. Therefore, we can make no statement about when one system or the other is preferred.

Next suppose $\tilde{t}_{med} < a_{lm}$. As $\tilde{t}_{\frac{D+1}{2}}$ is unchanged, the winning policies in the decisive legislative events $\lambda(a_l, a_m, a_m)$ and $\lambda(a_m, a_m, a_r)$ are again a_l and a_m . In a single district election, the winning policies in the races (l, m) , (m, r) , and (l, r) are now a_l , a_m and a_l . Under both legislative and single district elections, the preferred policy of \tilde{t}_{med} (a_l) may or may not be the policy outcome, depending on which equilibrium is played. Therefore, we can make no statement about when one system or the other is preferred.

Proof of Proposition 4

By Proposition 1, only two candidates will receive votes in each district. With D districts there will be $2D$ serious candidates. If party r 's candidates are serious in fewer than $\frac{D-1}{2}$ districts, the decisive event in which an extra seat for party r gives them a majority can never come about. Therefore, in any equilibrium where party r is serious in fewer than $\frac{D-1}{2}$ districts, the only decisive event voters can condition on is $\lambda(a_l, a_m, a_m)$. As this is the only decisive event, in each district, voters with $t < a_{lm}$ will vote v_l while those with $t > a_{lm}$ will vote for whichever of m or r is a serious candidate. An analogous result holds when party l is serious in fewer than $\frac{D-1}{2}$ districts. \square

Proof of Proposition 5

From Table 1 we see that when $3\Delta_l > \Delta_r$, $z = a_l$ is only possible if $s_l > \frac{D-1}{2}$. As such $E(z) = a_l$ requires $E(s_l) > \frac{D-1}{2}$. By Lemma 2, if $E(s_l) > \frac{D-1}{2}$ then $\lambda_{d_l}^1$ must be in the set of decisive events $\Lambda_{s_l = \frac{D-1}{2}} \subset \Lambda$. In addition if $E(s_m) > 1$ then $\lambda_{d_l}^1$ must be in the set of decisive events $\Lambda_{s_m > 1, s_l = \frac{D-1}{2}} \subset \Lambda_{s_l = \frac{D-1}{2}}$.

From Table 1 we see that when $3\Delta_l > \Delta_r$, any decisive event λ^i in the set $\Lambda_{s_m > 1, s_l = \frac{D-1}{2}}$ has $z_l^i = a_l \neq z_m^i, z_r^i$. From Proposition 1, each d_l must have either l and r or else l and m as serious candidates. Therefore, with $E(s_l) > \frac{D-1}{2}$, all d_l districts must be conditioning on decisive events in $\Lambda_{s_m > 1, s_l = \frac{D-1}{2}}$.

By the definition of a d_l district, the median voter in d_l must prefer $z_l^i = a_l$ to the some alternative: either z_m^i or z_r^i . The set of policies in $\Lambda_{s_m > 1, s_l = \frac{D-1}{2}}$ are $\{a_l, a_{lm}, a_{lr}, 2a_m - a_{lr}, a_m\}$. Suppose a district d_l conditions on a race where $z_l^i = a_l$ and the serious alternative is a_m . It is immediate to see that $\tilde{t}_{d_l} < a_{lm}$. It is also immediate to see that if the serious alternative was some other policy z' then $\tilde{t}_{d_l} < \frac{a_l + z'}{2}$. As a_m is the rightmost policy in the set, a necessary condition for a d_l district when $E(s_l) > \frac{D-1}{2}$ is $\tilde{t} < a_{lm}$. Thus, it must be that $\tilde{t}_{\frac{D+1}{2}} < a_{lm}$ in order for $z = a_l$ to come about.

Following the same steps as above, it is analogous to show that when $3\Delta_r > \Delta_l$ and $E(s_m) > 1$, then $E(z) = a_r$ requires $\tilde{t}_{\frac{D+1}{2}} > a_{mr}$.

Proof of Proposition 6

Case 1: $\Delta_l = \Delta_r$. Suppose party r is a serious candidate in fewer than $\frac{D-1}{4}$ districts; then it can win at most that many seats. The possible election outcomes are either a party l majority giving $z = a_l$, a party m majority giving $z = a_m$, or no majority but where party r has the least seats. From Table 1 we see that when $\Delta_l = \Delta_r$, the policy outcome will be a_m if no party has a majority and r is the smallest party. Conditional on r winning fewer than $\frac{D-1}{4}$ districts, the only decisive event is therefore $\lambda(a_l, a_m, a_m)$, when l wins $\frac{D-1}{2}$ seats and m is the second largest party. Given that only this distinct decisive event exists, all voters must be conditioning on it. As electing m or r here brings about the same policy, voters are indifferent between the two. In each district, those with $t < a_{lm}$ will vote v_l while those with $t > a_{lm}$ will coordinate on either v_m or v_r . With a choice over 2 policies in each district, no voter will be casting a misaligned vote. The case of l being a serious candidate in fewer than $\frac{D-1}{4}$ districts is analogous.

Case 2: $\Delta_l < \Delta_r$. The only difference from Case 1 we need to consider is when $s_r < \frac{D-1}{4}$ and no party has a majority. From Table 1 we see that when $\Delta_l < \Delta_r$, then with $\frac{D+1}{2} > s_l, s_m > s_r$ the policy outcome will be a_m . As shown above, when r is serious in fewer than

$\frac{D-1}{4}$ districts, the only distinct decisive case is $\lambda(a_l, a_m, a_m)$. All districts will condition on this event and, as before, there is no misaligned voting.

Case 3: $\Delta_l < \Delta_r$. The possible election outcomes are either a party r majority giving $z = a_r$, a party m majority giving $z = a_m$, or no majority but where party l has the least seats. From Table 1 we see that when $\Delta_r < \Delta_l$, if $\frac{D+1}{2} > s_r, s_m > s_l$ the policy outcome will be a_m . Therefore, when l is serious in fewer than $\frac{D-1}{4}$ districts, the only distinct decisive event is $\lambda(a_m, a_m, a_r)$. All districts will condition on this event and there is no misaligned voting. \square

Proof of Proposition 7

Case 1: Recall from Proposition 1 that if $\lambda_d^1 = \lambda(2')$ then voters in d must be conditioning on it. Furthermore, by Lemma 1 and Proposition 1 voters condition on the unique most likely pivotal and decisive event in which they are not indifferent. Whichever option a voter prefers in this case will also be his preferred over all possible decisive events. By the definition of a $\lambda(2')$ event, one of the 3 policy outcomes, $z_{c''}^1$, is dominated for each voter by one of the two other policies, z_c^1 and $z_{c'}^1$. Therefore as long as $z_{c''}^1$ is not serious, there will be no misaligned voting - no voter would wish to change their vote if it could unilaterally decide the district.

Case 2: Let the most likely decisive event λ_d^1 be a $\lambda(2)$ event where candidates c and c' are serious. If voters are conditioning on λ_d^1 it must be that $z_c^1 = z_{c''}^1$ or $z_{c'}^1 = z_{c''}^1$; Here, I take it to be the former. Without loss of generality let $z_c^1 < z_{c'}^1$. Any voter type with $t > \frac{z_c^1 + z_{c'}^1}{2}$ will vote $v_{c'}$, while any other type will vote v_c . The former group cannot be casting misaligned votes as they have $u_t(z_{c'}^1) > u_t(z_c^1)$, and the decisive event λ_d^1 is infinitely more likely than all others. Next, we need to consider whether any of the voters choosing v_c might be misaligned. All of these voters have $u_t(z_c^1) = u_t(z_{c''}^1) > u_t(z_{c'}^1)$, so that they want to beat c' but are indifferent between c and c'' in this most likely decisive event. If one of these voters could unilaterally decide which candidate coordination takes place on, he would decide by looking at the most likely pivotal event in which $z_c \neq z_{c''}$, call this event λ^i . If $u_t(z_c^i) > u_t(z_{c''}^i)$ then voter type t would prefer coordination to take place on candidate c , while if $u_t(z_c^i) < u_t(z_{c''}^i)$ she'd want coordination on c'' . Therefore, there is no misaligned voting in the district if there exists no type such that $u_t(z_c^i) < u_t(z_{c''}^i)$ and $u_t(z_c^1) = u_t(z_{c''}^1) > u_t(z_{c'}^1)$ when c and c' are the serious candidates. \square

Appendix C: Impatient Parties

In this section, I examine how the results of the benchmark model change when $\delta < 1$, so that parties are no longer perfectly patient. It is likely that the discount rates of politicians vary across countries depending on things such as constitutional constraints of bargaining, the status quo, and the propensity of politicians to be reelected.⁴³

In the benchmark model it didn't matter whether the bargaining protocol was random or had a fixed order; a coalition would always implement $z = a_m$. Once parties discount the future, we get different policy outcomes depending on which of the two is used. Also, once discounting is introduced into a bargaining model, one needs to decide whether players receive payoffs at each stage of bargaining or only once an agreement is reached. In the former case, the location of a status quo policy, Q , will be important for final policy outcomes. The literature is far from united in the treatment of stage utilities in government formation. With no stage payoffs, Jackson and Moselle (2002) show that there exists $\delta^* < 1$ such that if $\delta \geq \delta^*$ then the coalition policy will be within ϵ of a_m . However, their model doesn't lend itself easily to the current setup as mixing over proposals means precise policy outcomes are not pinned down. Therefore, I follow Austen-Smith and Banks (2005) and Banks and Duggan (2006) in having players receive stage payoffs. I assume the status quo is neither too extreme, $Q \in (-\Delta_c, \Delta_c)$ where $\Delta_c = \min\{\Delta_l, \Delta_r\}$, nor too central $Q \neq a_m$.⁴⁴

In each period where no agreement is reached, the status quo policy remains and enters party's payoff functions. All parties discount the future at the same rate of $\delta \in (0, 1)$. Therefore, if a proposal y is passed in period t , the payoff of party c is

$$W_c = -(1 - \delta^{t-1})(Q - a_c)^2 - \delta^{t-1}(y - a_c)^2 \quad (25)$$

For ease of analysis I assume, without loss of generality, that $a_m = 0$.⁴⁵ Banks and Duggan (2000) show that all stationary equilibria are no-delay equilibria and are in pure strategies when the policy space is one-dimensional and $\delta < 1$.

Fixed Order Bargaining

The order of recognition is fixed and follows the ranking of parties' seat shares. In Appendix D, I derive the policy outcomes for any ordering of parties; these are presented in

⁴³After the 2010 Belgian elections, legislative bargaining lasted for a record-breaking 541 days, suggesting high values of δ . Conversely, after the 2010 U.K. elections, a coalition government was formed within five days.

⁴⁴If $Q = a_m$ the result is the same as the benchmark case of $\delta = 1$.

⁴⁵Taking any original positions $(a_l, a_m \neq 0, a_r)$, we can always alter \mathbf{f} so that the preferences of all voter types are the same when $(a'_l, a'_m = 0, a'_r)$.

Table 3 below along with policies for specific values of δ and Q . From the table we see that the further party m moves down the ranking of seat shares, the further the coalition policy moves away from a_m .

| Seat Share | Policy | $\delta = 0.95$ $ Q = 0.5$ | $\delta = 0.95$ $ Q = 0.25$ | $\delta = 0.9$ $ Q = 0.5$ |
|-----------------------------|---------------------------|--------------------------------|---------------------------------|-------------------------------|
| $s_l > (D-1)/2$ | a_l | < -0.5 | < -0.25 | < -0.5 |
| $(D+1)/2 > s_l > s_r > s_m$ | $-\sqrt{(1-\delta^2)Q^2}$ | -0.156 | -0.078 | -0.218 |
| $(D+1)/2 > s_l > s_m > s_r$ | $-\sqrt{(1-\delta)Q^2}$ | -0.112 | -0.056 | -0.158 |
| $s_m > s_l, s_r$ | $a_m = 0$ | 0 | 0 | 0 |
| $(D+1)/2 > s_r > s_m > s_l$ | $\sqrt{(1-\delta)Q^2}$ | 0.112 | 0.056 | 0.158 |
| $(D+1)/2 > s_r > s_l > s_m$ | $\sqrt{(1-\delta^2)Q^2}$ | 0.156 | 0.078 | 0.218 |
| $s_r > (D-1)/2$ | a_r | > 0.5 | > 0.25 | > 0.5 |

Table 3: Policy outcomes with fixed order bargaining over policy and $\delta < 1$.

The proposition below shows that when the bargaining protocol is fixed, parties discount the future, and the status quo is not exactly a_m , it is even more difficult to have extreme outcomes than is the case in the benchmark model.

Proposition 8. *In any equilibrium with a fixed order of bargaining over policy and $\delta < 1$ in which the moderate party is expected to win at least one seat:*

1. *If $E(s_r) > 1$, then the expected policy can be a_l only if $\tilde{t}_{\frac{D+1}{2}} < \frac{a_l - \sqrt{(1-\delta)Q^2}}{2} < a_{lm}$.*
2. *If $E(s_l) > 1$, then the expected policy can be a_r only if $\tilde{t}_{\frac{D+1}{2}} > \frac{a_r + \sqrt{(1-\delta)Q^2}}{2} > a_{mr}$.*

Proof. See Appendix D. □

Here, a majority government will only come about if the electorate is even more biased in favour of policy extreme policies than in the benchmark case.⁴⁶ The reason is that in the benchmark case every coalition implements $z = a_m$, while with discounting and a fixed order protocol, the largest party has an advantage in coalition negotiations and can use this to get an alternative policy passed. Voters anticipate this power in coalition formation, so will only vote to bring about an l majority if they prefer it to an l -led coalition.

For example, if $a_l = -1$, $\delta = 0.95$ and $|Q| = 0.25$ then an l coalition would implement a policy -0.078 or -0.056 . Realising the advantage of party l in bargaining, voters will only approve a_l if $\tilde{t}_{\frac{D+1}{2}} < -0.528$, which is slightly to the left of the indifferent type a_{lm} in the benchmark case. What the proposition also shows is that the further the status quo is from

⁴⁶It is worth mentioning that without the restriction to $E(s_r) > 1$ and $E(s_l) > 1$ in the proposition, the thresholds become $\tilde{t}_{\frac{D+1}{2}} < a_{lm}$ and $\tilde{t}_{\frac{D+1}{2}} > a_{mr}$ as in the benchmark case.

a_m , the more likely we are to have coalition governments, all else equal. This is because a more distant status quo gives the formateur even more bargaining power over the moderate party. For example, if $a_l = -1$, $\delta = 0.95$ and $|Q| = 0.5$ then an l majority implementing a_l can only come about if $\tilde{t}_{\frac{D+1}{2}} < -0.556$. Similarly, reducing the discount factor will strengthen the bargaining hand of the formateur: setting $\delta = 0.9$ means $z = a_l$ requires $\tilde{t}_{\frac{D+1}{2}} < -0.579$.

Under this bargaining protocol, we cannot rule out misaligned voting completely. The decisive events where one party has no seats are all $\lambda(3)$ events, so for any $E(S)$ at least one of these $\lambda(3)$ events can always be conditioned on. As shown earlier, if a district conditions on a $\lambda(3)$ event there must be misaligned voting in that district. Nonetheless, Proposition 7 still holds here: there are equilibria with no misaligned voting in a subset of districts - an improvement on a single plurality election. The following proposition summarises the state of misaligned voting under this bargaining rule.

Proposition 9. *In any equilibrium with a fixed order of bargaining over policy and $\delta < 1$, there is always misaligned voting in at least one district. However, equilibria exist in which there is no misaligned voting in a subset of districts.*

Random Recognition Bargaining

In each period one party is randomly selected as formateur, where the probability of each party being chosen is equal to its seat share in the legislature, $\frac{s_i}{D}$. Party payoffs are again given by Equation 25. As usual if a party has a majority of seats it will implement its preferred policy. Following Banks and Duggan (2006), when no party has a majority I look for an equilibrium of the form $y_l = a_m - \Omega$, $y_m = a_m$, $y_r = a_m + \Omega$. Cho and Duggan (2003) show that this stationary equilibrium is unique. As bargaining is only over policy, any minimum winning coalition will include party m . When there was no discounting this meant party m could always achieve $z = a_m$. Now, however, the presence of discounting and $Q \neq 0$ allows parties l and r to offer policies further away from a_m , which party m will nonetheless support. The moderate party will be indifferent between accepting and rejecting an offer y when

$$W_m(y) = -(\Omega)^2 = -(1 - \delta)(Q)^2 - \frac{\delta(D - s_m)}{D}(\Omega)^2 \quad (26)$$

which, when rearranged gives

$$\Omega = \sqrt{\frac{(1 - \delta)Q^2}{1 - \delta \frac{D - s_m}{D}}} \quad (27)$$

Table 4 shows the equilibrium offer each party will make when chosen as formateur. Notice that the policies offered by l and r depend on the seat share of party m . The more seats

party m has, the closer these offers get to zero. The other thing to notice is that the policies lie inside $(-Q, Q)$.

| Formateur | Policy |
|-----------|--|
| y_l | $-\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}}Q^2 > - Q $ |
| y_m | $a_m = 0$ |
| y_r | $\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}}Q^2 < Q $ |

Table 4: Policy proposals with random order bargaining over policy, $\delta < 1$.

For a seat distribution such that no party has a majority, the expected policy outcome from bargaining is

$$E(z) = -\frac{s_l}{D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}}Q^2 \right) + \frac{s_m}{D}(0) + \frac{s_r}{D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}}Q^2 \right) \quad (28)$$

An extra seat for any of the three parties will increase their respective probabilities of being the formateur and so affect the expected policy outcome. Thus, every district always conditions on a choice between three distinct (expected) policies. We also see that as s_m increases, the expected policy moves closer and closer to zero. This occurs for two reasons; firstly because there is a higher probability of party m being the formateur, and secondly because s_m enters the policy offers of l and r ; as s_m increases the absolute value of these policies shrink.

The proposition below shows that when the bargaining protocol is random, parties discount the future, and the status quo is not exactly a_m , it is easier for a non-centrist party to win a majority of seats and implement its preferred policy than is the case in the benchmark model (though still more difficult than in a single district election).

Proposition 10. *In any equilibrium with a random order of bargaining over policy and $\delta < 1$ in which the moderate party is expected to win at least one seat:*

1. *The expected policy can be a_l only if $\tilde{t}_{\frac{D+1}{2}} < z_l^*$, where $a_{lm} < z_l^* < a_m, a_{lr}$.*
2. *The expected policy can be a_r only if $\tilde{t}_{\frac{D+1}{2}} > z_r^*$, where $a_m, a_{lr} < z_r^* < a_{mr}$.*

Proof. See Appendix D. □

The proposition implies that we should witness more majority governments than coalition governments when the bargaining protocol is random.⁴⁷ The reason is that with a random

⁴⁷Testing this empirically is difficult. We would need to look at the set of multi-party systems which use

recognition rule voters face uncertainty if they choose to elect a coalition. The implemented policy will vary greatly depending on which party is randomly chosen as formateur. As voters are risk averse, they find the certainty of policy provided by a majority government appealing. The median voter in the median district need not prefer the policy of a non-centrist party to that of party m in order for the former to win a majority of seats.

For example, if $a_l = -1$, $D = 101$, $\delta = 0.95$ and $|Q| = 0.25$ then $S = (50, 1, 50)$ would bring about a lottery over policies $(-0.229, 0, 0.229)$. The uncertainty generated means a voter with $t < -0.474$ would prefer to elect an l majority government, a type slightly to the right of a_{lm} , the indifferent type in the benchmark case. If instead we had $|Q| = 0.5$ and kept the other parameters unchanged voters would face a lottery over policies $(-0.459, 0, 0.459)$; a type $t < -0.396$ would now prefer an l majority to a coalition government.

The proposition stacks the deck against a coalition government. It gives us the right-most type who might ever want an l majority, that is when $S_{-d} = (\frac{D-1}{2}, 0, \frac{D-3}{2})$. This almost-equal split of seats between the l and r party ensures huge variance in the coalition policy, making a single party government very appealing. It is worth considering when an l majority would be preferred to a more balanced coalition. Keep the parameters as in the example above and let a district condition on $S_{-d} = (50, 25, 25)$ with l and r being serious in the district. If $|Q| = 0.25$ then a victory for r will bring about a lottery $(-0.105, 0, 0.105)$. A voter in this district will prefer a_l to this lottery if $t < -0.508$. If instead $|Q| = 0.5$, a victory for r will bring about a lottery $(-0.209, 0, 0.209)$. A voter in this district will prefer a_l to this lottery if $t < -0.509$. In contrast to the proposition, these less extreme cases have $\tilde{t}_{\frac{D+1}{2}} < z_l^* < a_{lm}$ meaning it is more difficult for a_l to come about than in the benchmark model.

Along with the previous propositions on extreme policy outcomes, Proposition 10 shows that no matter which of the bargaining rules is used, there is less scope for extreme outcomes in legislative elections using plurality rule than there is in stand-alone plurality elections such as mayoral or presidential elections. This remains true even when the coalition formation process leads to large uncertainty over policy outcomes.

Almost all decisive events under this bargaining rule are $\lambda(3)$ events and, as we know, if a district is conditioning on such an event it must have misaligned voting. There are, however, a selection of $\lambda(2')$ events which occur when $s_m = \frac{D-1}{2}$.⁴⁸ From Proposition 7, if a district is conditioning on such an event and the dominated candidate is not serious, there will be

plurality to elect their legislature. Furthermore, we would need to know which of the two explicit proposer rules (if any) they use in coalition formation. Thirdly, even if we could divide this set of legislatures into fixed and random protocol systems, we would still face the problem that the distribution of voters in one policy is different from those in another. Thus any evidence of more or less coalition governments between systems could simply be the result of different underlying voter preferences.

⁴⁸A figure showing the various decisive events in this case is available upon request.

no misaligned voting.

Proposition 11. *For any distribution of voter preferences, in a legislative election with a random order of bargaining over policy, $\delta < 1$, and where $\zeta \equiv \frac{D+1}{D-3} \sqrt{\frac{1-\delta}{1-\delta\frac{D+1}{2D}}}$:*

1. *If $|Q| > |\frac{a_l}{\zeta}|$ or $|Q| > |\frac{a_r}{\zeta}|$ there exist equilibria with no misaligned voting in any district.*
2. *If $|Q| < |\frac{a_l}{\zeta}|, |\frac{a_r}{\zeta}|$ in every equilibrium there will be misaligned voting in at least one district.*

Proof. See Appendix D. □

If $|Q| > |\frac{a_l}{\zeta}|$ and each district has l and m as serious candidates then no voter type would like to see party r win a seat. Electing r in a district would give $S = (\frac{D-1}{2}, \frac{D-1}{2}, 1)$. Coalition bargaining would lead to a lottery over policies very close to policies a_l and a_m , so voters would prefer the certainty of either a l or m majority government. This result contrasts with the case of fixed order bargaining in Proposition 9. There the cases where a party is expected to win no seats are what drives misalignment; Here, it is exactly these cases where misaligned voting can be excluded.

For large D and δ we will have $|Q| < |\frac{a_l}{\zeta}|, |\frac{a_r}{\zeta}|$. However, we can still find equilibria with no misaligned voting in as many as $D - 1$ districts. As I show in the proof, when $s_m = \frac{D-1}{2}$ the expected policy which comes about by electing the smallest party in the legislature is actually not preferred by any voter type. In this case, if a district focuses on the two national frontrunners, there will be no misaligned voting. We can even find cases where there will be misalignment in only a single district. If party m has a majority of seats, party r has one seat, and l has the rest, then as long as all d_m and d_l districts have m and l as serious candidates, only that single d_r district will have misaligned voting. This gives a fresh insight into the idea of “wasted votes”: if party l or r is expected to be the smallest of the parties in the legislature, and party m is expected to have a majority, then the least popular non-centrist party should optimally be abandoned by voters. Any district which actually elects the weakest national party does so due to a coordination failure; a majority there would instead prefer to elect one of the other two parties. Notice, however, that for this to be the case, the moderate party must be expected to win an overall majority. So, while the idea of a wasted vote does carry some weight, it clearly does not apply to the case of the Liberal Democrats.

Appendix D: Proofs for Appendix C

Bargaining Equilibrium for Fixed Order Protocol and $\delta < 1$

As equilibria are stationary we need only consider two orderings: $l > r > m > l > r > \dots$ and $r > l > m > r > l > \dots$. I will derive the equilibrium offers for the case of $l > r > m$, the other is almost identical. I solve the game by backward induction. At stage 3, party m will make an offer y_m which maximises its payoff subject to the proposal being accepted by either party l or r .

At stage 2, party r will either make an offer $y_r(m)$ to attract party m , or an offer $y_r(l)$ to attract party l . For these proposals to be accepted by m and l respectively requires

$$\begin{aligned} -y_r(m)^2 &\geq -(1-\delta)Q^2 - \delta y_m^2 \\ -(a_l - y_r(l))^2 &\geq -(1-\delta)(a_l - Q)^2 - \delta(a_l - y_m)^2 \end{aligned}$$

If $y_r(m)$ is chosen then the first inequality will bind and we have $y_r(m) = \sqrt{(1-\delta)Q^2 + \delta y_m^2}$. We can now compare the payoff of party l when $y_r(m)$ and $y_r(l)$ are implemented.

$$\begin{aligned} -(a_l - \sqrt{(1-\delta)Q^2 + \delta y_m^2})^2 &= -a_l^2 - (1-\delta)Q^2 - \delta y_m^2 + 2a_l\sqrt{(1-\delta)Q^2 + \delta y_m^2} \\ -(a_l - y_r(l))^2 &= -a_l^2 - (1-\delta)Q^2 - \delta y_m^2 + (1-\delta)2a_lQ + \delta 2a_ly_m \end{aligned}$$

Party l prefers policy $y_r(l)$ when

$$\begin{aligned} (1-\delta)2a_lQ + \delta 2a_ly_m &> 2a_l\sqrt{(1-\delta)Q^2 + \delta y_m^2} \\ 2a_l((1-\delta)Q + \delta y_m) &> 2a_l\sqrt{(1-\delta)Q^2 + \delta y_m^2} \\ (1-\delta)Q + \delta y_m &< \sqrt{(1-\delta)Q^2 + \delta y_m^2} \end{aligned}$$

the final inequality always holds. As party l gets a higher payoff from $y_r(l)$ than $y_r(m)$, the former must be closer to a_l on the policy line, and therefore further away from a_r . Clearly then, party r maximises its utility by choosing $y_r = \sqrt{(1-\delta)Q^2 + \delta y_m^2}$.

At stage 1, party l will either make an offer $y_l(m)$ to attract party m , or an offer $y_l(r)$ to

attract party r . For these proposals to be accepted by m and r respectively requires

$$\begin{aligned} -y_l(m)^2 &\geq -(1-\delta)Q^2 - \delta(-\sqrt{(1-\delta)Q^2 + \delta y_m^2})^2 \\ -(a_r - y_l(r))^2 &\geq -(1-\delta)(a_r - Q)^2 - \delta(a_r - \sqrt{(1-\delta)Q^2 + \delta y_m^2})^2 \end{aligned}$$

If $y_l(m)$ is chosen then the first inequality will bind and we have $y_l(m) = -\sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2}$. We can now compare the payoff of party r when $y_l(m)$ and $y_l(r)$ are implemented.

$$\begin{aligned} -(a_r + \sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2})^2 &= -a_r^2 - (1-\delta^2)Q^2 - \delta^2 y_m^2 - 2a_r \sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2} \\ -(a_r - y_l(r))^2 &= -(1-\delta)(a_r^2 + Q^2 - 2a_r Q) - \delta a_r^2 - \delta(1-\delta)Q^2 - \delta^2 y_m^2 + \delta 2a_r \sqrt{(1-\delta)Q^2 + \delta y_m^2} \end{aligned}$$

Party r prefers policy $y_l(r)$ when

$$\begin{aligned} (1-\delta)2a_r Q + \delta 2a_r \sqrt{(1-\delta)Q^2 + \delta y_m^2} &> -2a_r \sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2} \\ (1-\delta)Q + \delta \sqrt{(1-\delta)Q^2 + \delta y_m^2} &> -\sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2} \end{aligned}$$

the final inequality always holds. As party r gets a higher payoff from $y_l(r)$ than $y_l(m)$, the former must be closer to a_r on the policy line, and therefore further away from a_l . Clearly then, party l maximises its utility by choosing $y_l = -\sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2}$.

Now, we can return to stage 3 to show that $y_m = 0$. By stationarity, if y_m is rejected at stage 3, then in stage 4 $y_l = -\sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2}$ will be proposed and accepted. Parties l and r will accept proposal y_m if

$$\begin{aligned} -(a_l - y_m)^2 &\geq -(1-\delta)(a_l - Q)^2 - \delta(a_l + \sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2})^2 \\ -(a_r - y_m)^2 &\geq -(1-\delta)(a_r - Q)^2 - \delta(a_r + \sqrt{(1-\delta^2)Q^2 + \delta^2 y_m^2})^2 \end{aligned}$$

Party m 's payoff is maximised when $y_m = 0$ (because $a_m = 0$), so we want to check whether this is an implementable proposal. Letting $y_m = 0$ and rearranging, the two inequalities

above become

$$\begin{aligned} 0 &\leq (1 - \delta^3)Q^2 + 2a_l[\delta\sqrt{(1 - \delta^2)Q} - (1 - \delta)Q] \\ 0 &\leq (1 - \delta^3)Q^2 + 2a_r[\delta\sqrt{(1 - \delta^2)Q} - (1 - \delta)Q] \end{aligned}$$

The term in square brackets may be positive or negative. If it is positive then, party r will accept $y_m = 0$, if the term is negative then party l will accept $y_m = 0$.⁴⁹

Given $y_m = 0$, we can now characterise the accepted policy proposals (and therefore policy outcomes) for the fixed order protocol when $l > r > m > l > r > \dots$

$$\begin{aligned} y_l &= -\sqrt{(1 - \delta^2)Q^2} \\ y_r &= \sqrt{(1 - \delta)Q^2} \\ y_m &= 0 \end{aligned}$$

Instead when $r > l > m > r > l > \dots$, the same process gives:

$$\begin{aligned} y_r &= \sqrt{(1 - \delta^2)Q^2} \\ y_l &= -\sqrt{(1 - \delta)Q^2} \\ y_m &= 0 \end{aligned}$$

Proof of Proposition 8

For $z = a_l$ to be the expected outcome it must be that $E(s_l) > \frac{D-1}{2}$. Given the restriction that $E(s_m), E(s_r) > 1$, the set of distinct decisive events which d_l districts can be conditioning on is reduced to

$$\begin{aligned} \Lambda_{s_l = \frac{D-1}{2}, s_m > 1, s_r > 1} &= \{ \lambda(a_l, -\sqrt{(1 - \delta^2)Q^2}, -\sqrt{(1 - \delta^2)Q^2}), \\ &\lambda(a_l, -\sqrt{(1 - \delta^2)Q^2}, -\sqrt{(1 - \delta)Q^2}), \\ &\lambda(a_l, -\sqrt{(1 - \delta)Q^2}, -\sqrt{(1 - \delta)Q^2}) \} \end{aligned}$$

⁴⁹Whenever $\delta > 0.543689$ then the term is positive. Given that we mostly care about values of δ close to one, we can say that it is generally party r who accepts m 's offer.

Any race between a_l and $-\sqrt{(1-\delta)Q^2}$, where the former is the expected winner, must have $\tilde{t} < \frac{a_l - \sqrt{(1-\delta)Q^2}}{2}$. Any race between a_l and $-\sqrt{(1-\delta^2)Q^2}$, where the former is the expected winner, must have $\tilde{t} < \frac{a_l - \sqrt{(1-\delta^2)Q^2}}{2}$, a stricter condition. Therefore in order to for a party l to win a majority in expectation when $E(s_m), E(s_r) > 1$ it must be at least that $\tilde{t}_{\frac{D+1}{2}} < \frac{a_l - \sqrt{(1-\delta)Q^2}}{2}$. Notice that since $-\sqrt{(1-\delta)Q^2} < a_m$, then $\frac{a_l - \sqrt{(1-\delta)Q^2}}{2} < a_{lm}$. Similarly, for party r to win a majority in expectation when $E(s_m), E(s_l) > 1$ it must be that $\tilde{t}_{\frac{D+1}{2}} > \frac{a_r + \sqrt{(1-\delta)Q^2}}{2} > a_{mr}$. \square

Proof of Proposition 10

Under this bargaining rule, we can only have $E(z) = a_l$ if $E(s_l) > \frac{D-1}{2}$. If this is the case, any d_l district must be conditioning on a decisive event in $\Lambda_{s_l = \frac{D-1}{2}}$. Each of these decisive events are distinct: by increasing a party's seat share by one, it alters the expected policy outcome $E(z)$. In order to find the weakest condition for $E(z) = a_l$ to occur, I proceed in the following steps. Steps 1 and 2 show that electing party l is always the worst option for a player with $t < 0$. Step 3 shows that $S = (\frac{D-1}{2}, 1, \frac{D-1}{2})$ gives the lowest utility for any $t < 0$ type conditional on $s_m > 0$. Step 4 identifies the type who is indifferent between $S = (\frac{D-1}{2}, 1, \frac{D-1}{2})$ and a party l majority. Let $\Omega(r) \equiv \sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}Q^2}$ and $\Omega(m) \equiv \sqrt{\frac{1-\delta}{1-\delta\frac{D-(s_m+1)}{D}}Q^2}$. Note that $\Omega(m) < \Omega(r)$.

Step 1: If $u_t(a_l) > u_t(E(z_m^i))$ or $u_t(a_l) > u_t(E(z_r^i))$, it must be that $t < 0$. I show the case of $u_t(a_l) > u_t(E(z_m^i))$ as the other is analogous. Comparing the expected utility of a voter voting for l or m where $s_l = \frac{D-1}{2}$, $s_m > 0$ and $s_r > 0$ we have:

$$\begin{aligned} u_t(l) &= -(a_l - t)^2 \\ u_t(m) &= -\frac{D-1}{2D}(-\Omega(m) - t)^2 - \frac{s_m+1}{D}(t)^2 - \frac{D-1-2s_m}{2D}(\Omega(m) - t)^2 \end{aligned}$$

Some algebra shows that a player has $u_t(l) > u_t(m)$ if

$$t < \frac{Da_l^2 - (D-s_m)\Omega(m)^2}{2Da_l - (1-s_m)2\Omega(m)}$$

The RHS must always be negative. It is easy to see that this would also be the case comparing $u_t(l)$ to $u_t(r)$. Thus any player type who prefers a_l to a lottery over coalition policies must have $t < 0$.

Step 2: At any decisive event $\lambda^i \in \Lambda_{s_l = \frac{D-1}{2}, s_r} > 0$ every $t \leq 0$ prefers $E(z_m^i)$ to $E(z_r^i)$. Comparing the expected utility of a voter voting for m or r for any case where $s_l = \frac{D-1}{2}$,

$s_m > 0$ and $s_r > 0$ we have:

$$\begin{aligned} u_t(m) &= -\frac{D-1}{2D}(-\Omega(m)-t)^2 - \frac{s_m+1}{D}(t)^2 - \frac{D-1-2s_m}{2D}(\Omega(m)-t)^2 \\ u_t(r) &= -\frac{D-1}{2D}(-\Omega(r)-t)^2 - \frac{s_m}{D}(t)^2 - \frac{D+1-2s_m}{2D}(\Omega(r)-t)^2 \end{aligned}$$

Some algebra shows that a player has $u_t(m) > u_t(r)$ if

$$t < \frac{(D-s_m)(\Omega(r)^2 - \Omega(m)^2) + \Omega(r)^2}{2(s_m-1)(\Omega(r) - \Omega(m)) + 2\Omega(m)}$$

The RHS must always be positive. This means that any type $t \leq 0$ always prefers $E(z_m^i)$ to $E(z_r^i)$.

Step 3: $S = (\frac{D-1}{2}, 1, \frac{D-1}{2})$ gives the lowest utility for any $t < 0$ type conditional on $s_m > 0$. Using the result from Step 2 and noting that $E(z|S_{-d_r} = (\frac{D-1}{2}, k, \frac{D-1-2k}{2})) = E(z|S_{-d_m} = (\frac{D-1}{2}, k-1, \frac{D+1-2k}{2}))$ we can see that for a $t \leq 0$ the least preferred expected policy is $E(z|S = (\frac{D-1}{2}, 1, \frac{D-1}{2}))$.

Step 4: Identifying the type who prefers a_l to $E(z|S = (\frac{D-1}{2}, 1, \frac{D-1}{2}))$. To find the rightmost type who prefers a_l to a coalition we thus examine $S = (\frac{D-1}{2}, 1, \frac{D-3}{2})$. At this point, electing party l gives them a majority and brings about $z = a_l$, while electing party r leads to a coalition with an ex ante expected policy of

$$E\left(z|S = \left(\frac{D-1}{2}, 1, \frac{D-1}{2}\right)\right) = -\frac{D-1}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-1}{D}}} Q^2 \right) + \frac{2}{2D}(0) + \frac{D-1}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-1}{D}}} Q^2 \right)$$

A voter will prefer the former if

$$-(a_l - t)^2 > -\frac{D-1}{2D} \left(-\sqrt{\frac{1-\delta}{1-\delta\frac{D-1}{D}}} Q^2 - t \right)^2 - \frac{2}{2D}(-t)^2 - \frac{D-1}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-1}{D}}} Q^2 - t \right)^2$$

rearranging this we get that a voter prefers a_l if

$$t < \frac{a_l}{2} - \frac{D-1}{2Da_l} \left(\frac{1-\delta}{1-\delta(\frac{D-1}{D})} Q^2 \right)$$

As $a_l < 0$, the right hand side is greater than $\frac{a_l}{2}$, which is the cutoff point in the benchmark case (recalling that $a_{lm} = \frac{a_l}{2}$ when $a_m = 0$). The cutoff for a party l majority is thus given

by

$$\tilde{t}_{\frac{D+1}{2}} < z_l^* \equiv \frac{a_l}{2} - \frac{D-1}{2Da_l} \left(\frac{1-\delta}{1-\delta(\frac{D-1}{D})} Q^2 \right) > a_{lm} \quad (29)$$

As $z_l^* > a_{lm}$, the above result must also hold when $s_r = 0$ as well as $s_r > 0$.

Next we show that $z_l^* < a_m, a_{lr}$. First, I show $z_l^* < a_m$. Some manipulation shows

$$\frac{D-1}{2Da_l} \left(\frac{1-\delta}{1-\delta(\frac{D-1}{D})} Q^2 \right) = \frac{Q^2}{2a_l} \frac{\frac{D-1}{D} - \delta \frac{D-1}{D}}{\frac{D-1}{D} - \delta \frac{D-1}{D} + \frac{1}{D}} > \frac{Q^2}{2a_l} \quad (30)$$

Let $P \equiv \frac{\frac{D-1}{D} - \delta \frac{D-1}{D}}{\frac{D-1}{D} - \delta \frac{D-1}{D} + \frac{1}{D}}$ where the inequality in Equation 30 comes from $P < 1$. In order to show $z_l^* < a_m$ we need to show $\frac{a_l}{2} - \frac{P(Q)^2}{2a_l} < 0$. Given $P < 1$ and $|Q| \leq |a_l|$, this immediately follows.

Second, I show $z_l^* < a_{lr}$. If $\Delta_l < \Delta_r$ then necessarily $a_{lr} > a_m$, so the proof above applies. If $\Delta_l > \Delta_r$ we have $z_l^* < a_{lr}$ if

$$\begin{aligned} \frac{a_l}{2} - \frac{P(Q)^2}{2a_l} &< \frac{a_l}{2} + \frac{a_r}{2} \\ \frac{P(Q)^2}{2|a_l|} &< \frac{a_r}{2} \end{aligned}$$

As $Q^2 \leq a_r^2$ and $|a_l| > |a_r|$, the inequality above must hold.

Therefore, we have that a majority for party l , which brings about a_l can only occur if $\tilde{t}_{\frac{D+1}{2}} < z_l^* \equiv \frac{a_l}{2} - \frac{D-1}{2Da_l} \left(\frac{1-\delta}{1-\delta(\frac{D-1}{D})} Q^2 \right)$ where $z_l^* < a_m, a_{lr}$.

Analogously, a majority for party r , which brings about a_r can only occur if $\tilde{t}_{\frac{D+1}{2}} > z_r^* \equiv \frac{a_r}{2} + \frac{D-1}{2Da_r} \left(\frac{1-\delta}{1-\delta(\frac{D-1}{D})} Q^2 \right)$ where $z_r^* > a_m, a_{lr}$. \square

Proof of Proposition 11

Case 1: If $-|Q| < \frac{a_l}{\zeta}$ or $\frac{a_r}{\zeta} < |Q|$ there exist equilibria with no misaligned voting in any district. When $E(s_l) > \frac{D-1}{2}$ and $E(s_r) = 0$, voters will be conditioning on the point $(\frac{D-1}{2}, \frac{D-1}{2}, 0)$. Electing l or m brings about $z = a_l$ or $z = a_m = 0$ respectively, each of which are the preferred policies of $t = -1$ and $t = 0$. Electing r leads to an expected coalition policy given by

$$u_t(r) = -\frac{D-1}{2D} \left(-\sqrt{\frac{1-\delta}{1-\delta\frac{D+1}{2D}}} Q^2 - t \right)^2 - \frac{D-1}{2D} (t)^2 - \frac{2}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D+1}{2D}}} Q^2 - t \right)^2 \quad (31)$$

Some algebra shows that there exists a type for which Equation 31 is greater than both $u_t(l)$ and $u_t(m)$ if

$$-|Q| > \frac{a_l}{\zeta} \quad (32)$$

where $\zeta \equiv \frac{D+1}{D-3} \sqrt{\frac{1-\delta}{1-\delta \frac{D+1}{2D}}}$. Otherwise, if $-|Q| < \frac{a_l}{\zeta}$, then $u_t(r)$ is dominated for all voter types, making $(\frac{D-1}{2}, \frac{D-1}{2}, 0)$ a $\lambda(2')$ event. If each district has l and m as serious candidates then there will be no misaligned voting. It is analogous to show that if $\frac{a_r}{\zeta} < |Q| < a_r$ and all districts have m and r as serious candidates, there is no misaligned voting.

Case 2: If $|Q| > |\frac{a_l}{\zeta}|, |\frac{a_r}{\zeta}|$ there will be misaligned voting in at least one district.

Case 2A: When $s_l > \frac{D-1}{2}$ or $s_r > \frac{D-1}{2}$ there will always be misaligned voting in some districts. I examine the case where l is expected to win a majority; the other case is identical.

1) *When $E(s_l) > \frac{D-1}{2}$, $E(s_r) > 0$ and $E(s_m) > 0$ there will be misaligned voting.* Voters must be conditioning on a decisive event where $s_l = \frac{D-1}{2}$; the expected utility of electing the three different candidates is

$$\begin{aligned} u_t(l) &= -(a_l - t)^2 \\ u_t(m) &= -\frac{D-1}{2D} \left(-\sqrt{\frac{1-\delta}{1-\delta \frac{D-(s_m+1)}{D}}} Q^2 - t \right)^2 - \frac{s_m+1}{D} (t)^2 - \frac{D-1-2s_m}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta \frac{D-(s_m+1)}{D}}} Q^2 - t \right)^2 \\ u_t(r) &= -\frac{D-1}{2D} \left(-\sqrt{\frac{1-\delta}{1-\delta \frac{D-s_m}{D}}} Q^2 - t \right)^2 - \frac{s_m}{D} (t)^2 - \frac{D+1-2s_m}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta \frac{D-s_m}{D}}} Q^2 - t \right)^2 \end{aligned}$$

It suffices to consider the most extreme types $t = -1$, $t = 0$ and $t = 1$. Subbing these values in we see that a type $t = -1$ will always want to elect l and a type $t = 0$ will always want to elect m . For any case $E(s_l) > \frac{D-1}{2}$, $E(s_r) > 0$, $E(s_m) > 0$, there must therefore be misaligned voting as in the districts where r is expected to win, the other voters coordinate on either l or m . The supporters of that candidate which is not serious must be casting misaligned votes.

2) *When $E(s_l) > \frac{D-1}{2}$ and $E(s_m) = 0$ there will be misaligned voting in every district.* Voters will be conditioning on the point $(\frac{D-1}{2}, 0, \frac{D-1}{2})$ where electing l or r brings about $z = a_l$ or $z = a_r$ respectively, each of which are the preferred policies of $t = -1$ and $t = 1$. Electing m leads to an expected coalition policy given by

$$u_t(m) = -\frac{D-1}{2D} \left(-\sqrt{\frac{1-\delta}{1-\delta \frac{D-1}{D}}} Q^2 - t \right)^2 - \frac{2}{2D} (t)^2 - \frac{D-1}{2D} \left(\sqrt{\frac{1-\delta}{1-\delta \frac{D-1}{D}}} Q^2 - t \right)^2$$

Subbing in for $t = 0$, it is clear that $u_0(m) > u_0(l), u_0(r)$. As only 2 candidates receive votes,

one of these player types must be casting a misaligned vote.

3) When $E(s_l) > \frac{D-1}{2}$ and $E(s_r) = 0$ there will be misaligned voting when $-\zeta|Q| > a_l$. This follows directly from Case 1.

Case 2B: When no party has an expected majority there will be misaligned voting in some districts. For any expected seat distribution where no party has a majority, the expected utility of electing the three different candidates is

$$\begin{aligned} u_t(l) &= -\frac{s_l+1}{D} \left(-\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}} Q^2 - t \right)^2 - \frac{s_m}{D} (t)^2 - \frac{s_r}{D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}} Q^2 - t \right)^2 \\ u_t(m) &= -\frac{s_l}{D} \left(-\sqrt{\frac{1-\delta}{1-\delta\frac{D-(s_m+1)}{D}}} Q^2 - t \right)^2 - \frac{s_m+1}{D} (t)^2 - \frac{s_r}{D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-(s_m+1)}{D}}} Q^2 - t \right)^2 \\ u_t(r) &= -\frac{s_l}{D} \left(-\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}} Q^2 - t \right)^2 - \frac{s_m}{D} (t)^2 - \frac{s_r+1}{D} \left(\sqrt{\frac{1-\delta}{1-\delta\frac{D-s_m}{D}}} Q^2 - t \right)^2 \end{aligned}$$

Where I abuse notation slightly to let s_c to be the expected number of seats of party c before district d votes, so that $s_l + s_m + s_r = D - 1$. By subbing in $t = 0$, we see that this type will always want m elected. For there to be no misaligned voting it must therefore be the case that m is a serious candidate in every district. Suppose this is the case so that in a d_r district r and m are serious candidates and in a d_l district l and m are the serious candidates. In a d_r district a type $t = 1$ must have $u_1(r) > u_1(m)$. Note that a d_r district conditions on r having one less seat and l having one more seat than a d_l district conditions on. Using the equations above one can show that if $u_1(r|(s_l, s_m, s_r)) > u_1(m|(s_l, s_m, s_r))$ then $u_1(r|(s_l - 1, s_m, s_r + 1)) > u_1(m|(s_l - 1, s_m, s_r + 1))$. That is, in a given equilibrium, if a type $t = 1$ in a d_r district prefers r to m , then a type $t = 1$ in a d_l district also prefers r to m . As this type also prefers r to l and the focal candidates in the d_l district are l and m , he must be casting a misaligned vote.

Case 2C: When $s_m > \frac{D-1}{2}$ there will be misaligned voting in at least one district.

If $s_m > \frac{D-1}{2}$, all districts must be conditioning on decisive events where $s_m = \frac{D-1}{2}$. In

such cases the expected utility of a type t voter in voting for each of the parties is

$$\begin{aligned}
u_t(l) &= -\frac{s_l + 1}{D} \left(-\sqrt{\frac{1 - \delta}{1 - \delta^{\frac{D+1}{2D}}}} Q^2 - t \right)^2 - \frac{D - 1}{2D} (t)^2 - \frac{D - 1 - 2s_l}{2D} \left(\sqrt{\frac{1 - \delta}{1 - \delta^{\frac{D+1}{2D}}}} Q^2 - t \right)^2 \\
u_t(m) &= -(t)^2 \\
u_t(r) &= -\frac{s_l}{D} \left(-\sqrt{\frac{1 - \delta}{1 - \delta^{\frac{D+1}{2D}}}} Q^2 - t \right)^2 - \frac{D - 1}{2D} (t)^2 - \frac{D + 1 - 2s_l}{2D} \left(\sqrt{\frac{1 - \delta}{1 - \delta^{\frac{D+1}{2D}}}} Q^2 - t \right)^2
\end{aligned}$$

Note that for $t = 0$ we have $u_t(m) > u_t(l) = u_t(r)$. Any voter type with $t < 0$ has $u_t(l) > u_t(r)$, while any voter with $t > 0$ has $u_t(l) < u_t(r)$. However, it could be that some of these types prefer $u_t(m)$ to either of the other two. In order to check this I calculate the derivative of each of the expected utilities with respect to t .

$$\begin{aligned}
\frac{d[u(l)]}{dt} &= -2t + \frac{D - 3 - 4s_l}{D} \left(\sqrt{\frac{1 - \delta}{1 - \delta^{\frac{D+1}{2D}}}} Q^2 \right) \\
\frac{d[u(m)]}{dt} &= -2t \\
\frac{d[u(r)]}{dt} &= -2t + \frac{D + 1 - 4s_l}{D} \left(\sqrt{\frac{1 - \delta}{1 - \delta^{\frac{D+1}{2D}}}} Q^2 \right)
\end{aligned}$$

When $s_l < \frac{D-3}{4}$ then for any $t < 0$ we have $\frac{d[u(m)]}{dt} < \frac{d[u(l)]}{dt} < \frac{d[u(r)]}{dt}$. Utility is increasing for all three cases as we move towards 0. Given $\frac{d[u(m)]}{dt} < \frac{d[u(l)]}{dt}$ for any $t < 0$ and $u_t(m) > u_t(l) = u_t(r)$ for $t = 0$, it must be that for $s_l < \frac{D-3}{4}$ there is no type with $u_t(l) > u_t(m), u_t(r)$. When $s_l > \frac{D+1}{4}$ then for any $t > 0$ we have $\frac{d[u(m)]}{dt} < \frac{d[u(r)]}{dt} < \frac{d[u(l)]}{dt}$. Combined, with the fact that we have $u_t(m) > u_t(l) = u_t(r)$ for $t = 0$, this means that for $s_l > \frac{D+1}{4}$ there is no type with $u_t(r) > u_t(m), u_t(l)$.

What this means is that, conditional on $s_m = \frac{D-1}{2}$, if $s_l < \frac{D-3}{4}$, a district in which m and r are the serious candidates will have no misaligned voting; and if $s_l > \frac{D+1}{4}$ then a district in which m and l are the serious candidates will have no misaligned voting.

However, each equilibrium has misaligned voting in at least one district. Recall that $E(S_{-d_l}) = (s_l - 1, s_m, s_r)$, $E(S_{-d_m}) = (s_l, s_m - 1, s_r)$ and $E(S_{-d_r}) = (s_l, s_m, s_r - 1)$. As all the relevant decisive events occur at $s_m = \frac{D-1}{2}$, d_l and d_r districts will have the same “route” to being decisive. That is, in any equilibrium if d_l districts are conditioning on $(k, \frac{D-1}{2}, \frac{D-1}{2} - k)$, then d_r districts must be conditioning on $(k + 1, \frac{D-1}{2}, \frac{D-1}{2} - (k + 1))$.

When $0 < s_l < \frac{D-3}{4}$, all d_m and d_r districts are conditioning on $\lambda(2')$ events. In any of

these districts if the serious candidates are m and r , there is no misaligned voting. However, we know that each d_l district must either be conditioning on a $\lambda(2')$ event or else a $\lambda(3)$ event (if it conditions on $S_{-d} = (0, \frac{D-1}{2}, \frac{D-1}{2})$). Whichever one of these is the case, there will always be misaligned voting in these d_l districts. Indeed, if a single d_l district conditions on $S_{-d_l} = (0, \frac{D-1}{2}, \frac{D-1}{2})$, and the other districts are all races between m and r , it must be that there is only misaligned voting in this single d_l district. Examining the $\frac{D+1}{4} < s_l < \frac{D-1}{2}$ case gives the same insight for the mirror case; there'll be no misaligned voting in d_l or d_m districts if they focus on races between l and m , but there will always be misaligned voting in the d_r districts. \square

References

- AUSTEN-SMITH, D., AND J. S. BANKS (1988): “Elections, Coalitions, and Legislative Outcomes,” *American Political Science Review*, 82(2), 405–422.
- (2005): *Positive Political Theory II*. University of Michigan Press.
- BANDYOPADHYAY, S., K. CHATTERJEE, AND T. SJÖSTRÖM (2011): “Pre-electoral Coalitions and Post-election Bargaining,” *Quarterly Journal of Political Science*, 6(1), 1–53.
- BANDYOPADHYAY, S., AND M. P. OAK (2008): “Coalition governments in a model of parliamentary democracy,” *European Journal of Political Economy*, 24(3), 554–561.
- BANKS, J. S., AND J. DUGGAN (2000): “A Bargaining Model of Collective Choice,” *American Political Science Review*, 94(1), 73–88.
- (2006): “A General Bargaining Model of Legislative Policy-making,” *Quarterly Journal of Political Science*, 1(1), 49–85.
- BARON, D. P. (1991): “A Spatial Bargaining Theory of Government Formation in Parliamentary Systems,” *American Political Science Review*, 85(1), 137–164.
- BARON, D. P., AND D. DIERMEIER (2001): “Elections, Governments, and Parliaments in Proportional Representation Systems,” *Quarterly Journal of Economics*, 116(3), 933–967.
- BARON, D. P., AND J. FEREJOHN (1989): “Bargaining in Legislatures,” *American Political Science Review*, 83(4), 1181–1206.
- 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., AND G. GRATTON (2015): “Majority Runoff Elections: Strategic Voting and Duverger’s Hypothesis,” *Theoretical Economics*, 10(2), 283–314.
- CHO, S.-J., AND J. DUGGAN (2003): “Uniqueness of Stationary Equilibria in a One-dimensional Model of Bargaining,” *Journal of Economic Theory*, 113(1), 118–130.
- COX, G. W. (1997): *Making Votes Count*. Cambridge University Press.

- DIERMEIER, D., AND A. MERLO (2000): “Government Turnover in Parliamentary Democracies,” *Journal of Economic Theory*, 94(1), 46–79.
- ERASLAN, H., D. DIERMEIER, AND A. MERLO (2003): “A Structural Model of Government Formation,” *Econometrica*, 71(1), 27–70.
- FEY, M. (1997): “Stability and Coordination in Duverger’s Law: A Formal Model of Pre-election Polls and Strategic Voting,” *American Political Science Review*, 91(1), 135–147.
- FISHER, S. D., AND D. P. MYATT (2014): “Strategic Voting in Plurality Rule Elections,” *mimeo*.
- JACKSON, M. O., AND B. MOSELLE (2002): “Coalition and Party Formation in a Legislative Voting Game,” *Journal of Economic Theory*, 103(1), 49–87.
- KALANDRAKIS, A. (2004): “A Three-Player Dynamic Majoritarian Bargaining Game,” *Journal of Economic Theory*, 116(2), 294–322.
- KAWAI, K., AND Y. WATANABE (2013): “Inferring Strategic Voting,” *American Economic Review*, 103(4), 624 – 662.
- KRISHNA, V., AND J. MORGAN (2011): “Overcoming Ideological Bias in Elections,” *Journal of Political Economy*, 119(2), 183–211.
- LASLIER, J.-F. (2012): “And the Loser Is Plurality Voting,” in *Electoral Systems*, ed. by D. S. Felsenthal, and M. Machover, Studies in Choice and Welfare, pp. 327–351. Springer Berlin Heidelberg.
- MCCARTY, N., K. POOLE, AND H. ROSENTHAL (2008): *Polarized America: The Dance of Ideology and Unequal Riches*. MIT press.
- MILGROM, P. R., AND R. J. WEBER (1985): “Distributional Strategies for Games with Incomplete Information,” *Mathematics of Operations Research*, 10(4), 619–632.
- MORELLI, M. (1999): “Demand Competition and Policy Compromise in Legislative Bargaining,” *American Political Science Review*, 93(4), 809–820.
- (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.

- 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.
- PALFREY, T. R. (1989): “A Mathematical Proof of Duverger’s Law,” in *Models of Strategic Choice in Politics*, ed. by P. C. Ordeshook, pp. 69–91. University of Michigan Press.
- POLBORN, M., AND S. KRASA (2015): “Political Competition in Legislative Elections,” *mimeo*.