Online Appendix to "Strategic Voting in Two-Party Legislative Elections"

Niall Hughes

Appendix B. Poisson Properties

Single-district Poisson Properties

The number of voters in a district is a Poisson random variable k_d with mean k. The probability of having exactly η voters is $\Pr[k_d = \eta] = (e^{-k}k^{\eta})/(\eta!)$. Poisson Voting games exhibit some useful properties. By *environmental equivalence*, from the perspective of a player in the game, the number of other players is also a Poisson random variable k_d with mean k. By the *decomposition property*, the number of voters with type $t \in \mathcal{T}$ is Poisson distributed with mean $kf_d^{\omega}(t)$, and is independent of the number of other types. In what follows, I allow a district's vote share to depend on the state of the world, though in the model this will only be the case in district $\underline{d}$.

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

$$\Pr[x_d|k\nu_d^{\omega}] = \frac{e^{-k\nu_d^{\omega}}(k\nu_d^{\omega})^{x_d(L)}}{x_d(L)!}\frac{e^{-k(1-\nu_d^{\omega})}(k(1-\nu_d^{\omega}))^{x_d(R)}}{x_d(R)!} \tag{B1}$$

Its associated magnitude is

$$mag[x_d|\nu_d^{\omega}] \equiv \lim_{k\to\infty}\frac{\log(\Pr[x_d|k\nu_d^{\omega}])}{k} = \lim_{k\to\infty}\nu_d^{\omega}\psi\left(\frac{x_d(L)}{k\nu_d^{\omega}}\right) + (1-\nu_d^{\omega})\psi\left(\frac{x_d(R)}{k(1-\nu_d^{\omega})}\right) \tag{B2}$$

where $\psi(\theta) = \theta(1-\log(\theta)) - 1$. In what follows, I keep the strategies of voters fixed, so probabilities and magnitudes depend only on the state ω.

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

$$mag[A_d|\omega] \equiv \lim_{k\to\infty}\frac{\log(\Pr[A_d|k\nu_d^{\omega}])}{k} = \lim_{k\to\infty}\max_{x_d\in A_d}\nu_d^{\omega}\psi\left(\frac{x_d(L)}{k\nu_d^{\omega}}\right) + (1-\nu_d^{\omega})\psi\left(\frac{x_d(R)}{k(1-\nu_d^{\omega})}\right) \tag{B3}$$

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

Corollary to the Magnitude Theorem If two events A_d and A'_d have $mag[A_d|\omega] > mag[A'_d|\omega']$ for $\omega, \omega' \in \{0, 1\}$, then their probability ratio converges to zero as $k \to \infty$.

$$mag[A'_d|\omega'] < mag[A_d|\omega] \Rightarrow \lim_{k\to\infty} \frac{\Pr[A'_d|\omega']}{\Pr[A_d|\omega]} = 0 \tag{B4}$$

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

Maximising Equation B3 subject to the appropriate constraints, we get

$$\begin{aligned} mag[Lwin|\omega] &= 0 \\ mag[Rwin|\omega] &= 2\sqrt{\nu_d^\omega(1-\nu_d^\omega)} - 1 \end{aligned} \tag{B5}$$

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

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

$$\Pr[x_d(L) = x_d(R)|\omega] \approx \frac{e^{k(2\sqrt{\nu_d^\omega(1-\nu_d^\omega)}-1)}}{\pi(k+\frac{1}{3})} \tag{B6}$$

Multi-District Poisson Properties

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

$$\Pr[\mathbf{x}|k\boldsymbol{\nu}^\omega] = \prod_{d\in\mathcal{D}} \frac{e^{-k\nu_d^\omega}(k\nu_d^\omega)^{x_d(L)}}{x_d(L)!} \frac{e^{-k(1-\nu_d^\omega)}(k(1-\nu_d^\omega))^{x_d(R)}}{x_d(R)!} \tag{B7}$$

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

$$mag[\mathbf{x}|\omega] \equiv \lim_{k\to\infty} \frac{\log(\Pr[\mathbf{x}|k\boldsymbol{\nu}^\omega])}{k} = \lim_{k\to\infty} \sum_{d\in\mathcal{D}} \nu_d^\omega \psi\left(\frac{x_d(L)}{k\nu_d^\omega}\right) + (1-\nu_d^\omega)\psi\left(\frac{x_d(R)}{k(1-\nu_d^\omega)}\right) = \sum_d mag[x_d|\omega] \tag{B8}$$

Multi-District Magnitude Theorem

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 $\bar{x_d} \in A_d = \underset{x_d}{\text{argmax}}\, \nu_d\psi\left(\frac{x_d(L)}{k\nu_d}\right) + (1-\nu_d)\psi\left(\frac{x_d(R)}{k(1-\nu_d)}\right)$, that is, $\bar{x_d}$ is the most likely district vote profile in A_d given some ν_d. Then, Multi-District Magnitude Theorem

Hughes (2016) states that:

$$mag[\mathbf{A}|\omega] = \sum_{d\in\mathcal{D}} mag[A_d|\omega] = \sum_{d\in\mathcal{D}} mag[\bar{x_d}|\omega] = mag[\bar{\mathbf{x}}|\omega] \tag{B9}$$

The first equality 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 B8. Together they show that the single-district magnitude theorem extends to a multi-district setting.

Following from this, and using Equation B4, we have that the Corollary to the Magnitude Theorem also extends to the multi-district case. If $mag[\mathbf{A}'|\omega'] < mag[\mathbf{A}|\omega]$ for $\omega, \omega' \in \{0, 1\}$, then

$$\lim_{k\to\infty} \frac{\Pr[\mathbf{A}'|k\boldsymbol{\nu}^{\omega'}]}{\Pr[\mathbf{A}|k\boldsymbol{\nu}^{\omega}]} = \lim_{k\to\infty} \frac{e^{k mag[\mathbf{A}'|\omega']}}{e^{k mag[\mathbf{A}|\omega]}} = \lim_{k\to\infty} e^{k(mag[\bar{\mathbf{x}}'|\omega'] - mag[\bar{\mathbf{x}}|\omega])} = 0 \tag{B10}$$

Appendix C. Ruling out knife-edge events

In this section, I show which features of an equilibrium are not robust for any population level k. This helps to characterise the set of equilibria when $k \to \infty$ in the main body of the paper.

In a given continuation game $(\mathbf{f}, \mathbf{a})$ for a fixed k, either of the following could occur in equilibrium: (i) a district's vote could be exactly tied in expectation, and (ii) two distinct pivotal events could have the same magnitude in the same state of the world. Lemma C1 below shows these knife-edge conditions are measure-zero events, implying no loss of predictive power from excluding them. First, I define some necessary additional notation. Let $\widetilde{a}_d^{\text{win}}(\nu_d, a_d)$ denote the platform with the largest probability of winning in district d given platforms a_d and expected vote shares ν_d. Let $\widetilde{\mathbf{a}}^{\text{win}}(\boldsymbol{\nu}^\omega, \mathbf{a})$ be the profile across all districts. For a pivotal event $i \in \mathcal{P}_d$, let $\widetilde{\mathbf{a}}_{-d}^{\text{win}}(i|\boldsymbol{\nu}^\omega, \mathbf{a})$ denote the profile of district outcomes in the remaining $D-1$ districts that maximises the probability of i given $\boldsymbol{\nu}^\omega$ and $\mathbf{a}$. Denote the corresponding profile excluding districts d and $\underline{d}$ as $\widetilde{\mathbf{a}}_{-\{d,\underline{d}\}}^{\text{win}}(i|\boldsymbol{\nu}, \mathbf{a})$. For two distinct pivotal events $i, j \in \mathcal{P}_d$, $\widetilde{\mathbf{a}}_{-\{d,\underline{d}\}}^{\text{win}}(i|\boldsymbol{\nu}, \mathbf{a}) \neq \widetilde{\mathbf{a}}_{-\{d,\underline{d}\}}^{\text{win}}(j|\boldsymbol{\nu}, \mathbf{a})$ means that i and j have distinct most likely profiles, irrespective of the state of the world. In other words, i and j come about by different outcomes in at least one district among the $D-2$ districts other than d and $\underline{d}$.

Lemma C1. *Let $\mathcal{Z}$ be the set of continuation histories in $\mathcal{F}^D \times \mathcal{A}^D$ for which at least one of the following knife–edge conditions holds:*

(i) Expected vote tie: $\nu_d = 0.5$ *for some d.*

(ii) Equal magnitudes with distinct most likely profiles:
$mag[i] = mag[j]$ *for some d where $i \neq j \in \mathcal{P}_d$ and* $\widetilde{\mathbf{a}}_{-\{d,\underline{d}\}}^{\text{win}}(i|\boldsymbol{\nu}, \mathbf{a}) \neq \widetilde{\mathbf{a}}_{-\{d,\underline{d}\}}^{\text{win}}(j|\boldsymbol{\nu}, \mathbf{a})$.

Then $\mathcal{Z}$ has Lebesgue measure 0*, hence probability* 0 *under the product prior. Its complement $E := \left(\mathcal{F}^D \times \mathcal{A}^D\right) \setminus \mathcal{Z}$ occurs with probability one.*

Proof. Fix a continuation game $(\mathbf{f}, \mathbf{a})$ and an equilibrium $\boldsymbol{\sigma}^*$ (possibly mixed). Let $\mathcal{M}$ be the set of districts in which at least one voter type mixes under $\boldsymbol{\sigma}^*$. For each mixing district $m \in \mathcal{M}$, let G_m denote the gain of the mixing type in m; at equilibrium $G_m = 0$.[1]

Step 1: *Ruling out the case of all districts mixing, $|\mathcal{M}| = D$.*

If $|\mathcal{M}| = D$, the D indifference equalities $G_m(\boldsymbol{\nu}) = 0$ pin down $\boldsymbol{\nu}^*(k, \mathbf{a})$; these equations depend on $\mathbf{a}$ and the cardinal utilities U, but not on $\mathbf{f}$, so $\boldsymbol{\nu}^*$ is locally independent of $\mathbf{f}$. Imposing either the tie $\nu_d^* = 0.5$ or, for distinct $i \neq j$ with different most–likely profiles, the equality $mag[i|\boldsymbol{\nu}^*] = mag[j|\boldsymbol{\nu}^*]$, adds one independent scalar restriction, yielding an overdetermined system with $D+1$ constraints in D unknowns. Generically in U (i.e., absent fine-tuned utility coincidences), no solution exists. Hence, the all–mixing case cannot satisfy the knife-edge conditions, and we may confine attention to $\mathcal{M} = \emptyset$ or $|\mathcal{M}| < D$.

Step 2: *(i) Expected vote tie.*

[1] If two conflicted types were to mix in the same district, their gain functions would have to vanish simultaneously at the same $\boldsymbol{\nu}$. Given the strict utility orderings, this is a non-generic coincidence in the cardinal utilities. In that case, we simply add one indifference equation per mixer; the dimension–counting arguments below remain unchanged.

- If $d \notin \mathcal{M}$, then $\nu_d = \sum_t \sigma^*_{t,d} f_d(t)$ is linear in f_d, so any tie is broken by an arbitrarily small perturbation of f_d. The set $\{\mathbf{f} : \nu_d(\mathbf{f}) = 0.5\}$ therefore has Lebesgue measure 0.

- If $d \in \mathcal{M}$ and $|\mathcal{M}| < D$, pick a pure district $d' \notin \mathcal{M}$ and perturb $f_{d'}$ so $\nu_{d'}$ changes; restoring $G_m = 0$ for all $m \in \mathcal{M}$ requires adjusting mixers' probabilities, and generically the required adjustment in d is nonzero, so ν_d moves off 0.5. Thus $\{\mathbf{f} : \nu_d(\mathbf{f}) = 0.5\}$ has measure 0.

Step 3: *(ii) Equal magnitudes for distinct events.*

Fix d and $i \neq j \in \mathcal{P}_d$ with $\widetilde{\mathbf{a}}^{\text{win}}_{-\{d,\underline{d}\}}(i) \neq \widetilde{\mathbf{a}}^{\text{win}}_{-\{d,\underline{d}\}}(j)$. By part (i), each district apart from $\underline{d}$ has a unique expected winner, so these most likely profiles are well defined. The condition $\widetilde{\mathbf{a}}^{\text{win}}_{-\{d,\underline{d}\}}(i) \neq \widetilde{\mathbf{a}}^{\text{win}}_{-\{d,\underline{d}\}}(j)$ implies there is at least one district in which the expected winner differs under events i and j. However, i and j must have different magnitudes if they differ in only one district - thus, imposing $mag[i] = mag[j]$, implies that there exist at least two distinct districts $d_1, d_2 \notin \{d, \underline{d}\}$ such that under event i the expected winner wins in d_1 but loses in d_2, whereas under event j the expected winner wins in d_2 but loses in d_1. Moreover, it cannot be that the same platform is expected to lose in both d_1 and d_2. If that were the case, it would imply $i = j$. I present the case where the only difference between $\widetilde{\mathbf{a}}^{\text{win}}_{-\{d,\underline{d}\}}(i)$ and $\widetilde{\mathbf{a}}^{\text{win}}_{-\{d,\underline{d}\}}(j)$ is via d_1, d_2; the same logic applies if they differ in more districts. Given that i and j differ only in d_1 and d_2, $mag[i] = mag[j]$ reduces to equality of the local underdog magnitudes, which is equivalent to $\nu_{d_1} = \nu_{d_2}$ if the expected winners in d_1, d_2 are from the same party (and to $\nu_{d_1} = 1 - \nu_{d_2}$ otherwise). Without loss of generality we assume $\nu_{d_1} = \nu_{d_2}$.

- Suppose $d_1, d_2 \notin \mathcal{M}$. Perturb f_{d_1} so ν_{d_1} changes to ν'_{d_1}, but remains above 0.5. This results in a change to $mag[j]$ but not $mag[i]$ because, by Equation B5, ν_{d_1} only enters the magnitude of the underdog winning. The set $\{\mathbf{f} : mag[i, \mathbf{f}] = mag[j, \mathbf{f}]\}$ therefore has Lebesgue measure 0.

- Suppose $d_1 \in \mathcal{M}$ but $d_2 \notin \mathcal{M}$. Perturb f_{d_2} so ν_{d_2} changes to ν'_{d_2}; restoring $G_m = 0$ for all $m \in \mathcal{M}$ requires adjusting mixers' probabilities, and generically the required adjustment in d_1 will be such that $\nu'_{d_1} \neq \nu'_{d_2}$. The set $\{\mathbf{f} : mag[i, \mathbf{f}] = mag[j, \mathbf{f}]\}$ therefore has Lebesgue measure 0.

- Suppose $d_1, d_2 \in \mathcal{M}$ and $|\mathcal{M}| < D$, so that there exists a district $d_3 \notin \mathcal{M}$. Perturb f_{d_3} so ν_{d_3} changes to ν'_{d_3}; restoring $G_m = 0$ for all $m \in \mathcal{M}$ requires adjusting mixers' probabilities. The required adjustment of the mixing types in d_1 and d_2 can only be equal if they have identical gain functions. This is the case only if the two districts are *structurally equivalent* - that is, the same voter type is mixing in d_1 and d_2 and platforms are identical in both districts. Such an equivalence implies $\nu_{d_1} = \nu_{d_2}$, which in turn implies the same policy is expected to lose in both d_1 and d_2. But this implies $i = j$, a contradiction. Hence, d_1 and d_2 cannot be structurally equivalent. Therefore, the required adjustment by the mixing types in d_1 and d_2 will be such that $\nu'_{d_1} \neq \nu'_{d_2}$. The set $\{\mathbf{f} : mag[i, \mathbf{f}] = mag[j, \mathbf{f}]\}$ therefore has Lebesgue measure 0.

Steps 2 and 3 thus show that each knife–edge condition defines a set of Lebesgue measure 0 in the $7D$ continuous coordinates $\mathcal{F}^D$. Because $\mathcal{A}^D$ is finite, the product $\mathcal{Z}$ of those null sets also has measure 0. The product prior $\bigotimes_{d=1}^{D} \Pi_d$ is absolutely continuous with respect to that measure, so $\Pr[\mathcal{Z}] = 0$ and $\Pr[E] = \Pr\big[(\mathcal{F}^D \times \mathcal{A}^D) \setminus \mathcal{Z}\big] = 1$.

$\square$

Ruling out (i) ensures that there is a unique expected winner in each district (apart from $\underline{d}$), while ruling out (ii) establishes that distinct pivotal events cannot generically have the same magnitude. The lemma yields the following corollary.

Corollary (to Lemma C1). *In any E we have the following:*

(i) For $\underline{d}$, the set of most likely pivotal events, $Piv^1_{\underline{d}}$, is a singleton.

(ii) For each $d \neq \underline{d}$, Piv^1_d contains at most two pivotal events. If it contains two, the two events occur in different states of the world.

Appendix D. Analysis of district $\underline{d}$

From the perspective of $\underline{d}$ there is no aggregate uncertainty in other districts. That is, the probability of any $\text{piv}_{\underline{d}}$ event does not depend on the state of the world. This simplifies the gain function of voters in $\underline{d}$ and leads to a modified version of Proposition 1:

Proposition D1. *In any equilibrium with* $k > \bar{k}_{\underline{d}}(\mathbf{f}, \mathbf{a})$*, voter best responses in* $\underline{d}$ *are represented in Table D1.*

<table>
<tr><th>District</th><th>Vote v_L</th><th>Vote v_R</th></tr>
<tr><td>(a_{LN}, a_{RN})</td><td rowspan="4">$t_{NL}, t_{LN}, t_{YL}, t_{LY}$</td><td rowspan="4">$t_{NR}, t_{RN}, t_{YR}, t_{RY}$</td></tr>
<tr><td>(a_{LY}, a_{RY})</td></tr>
<tr><td>(a_{LN}, a_{RY}) with $Piv^1_{\underline{d}} = \text{piv}(LR)$</td></tr>
<tr><td>(a_{LY}, a_{RN}) with $Piv^1_{\underline{d}} = \text{piv}(LR)$</td></tr>
<tr><td>(a_{LN}, a_{RY}) with $Piv^1_{\underline{d}} = \text{piv}(NY)$</td><td>$t_{NL}, t_{LN}, t_{NR}, t_{RN}$</td><td>$t_{YL}, t_{LY}, t_{YR}, t_{RY}$</td></tr>
<tr><td>(a_{LY}, a_{RN}) with $Piv^1_{\underline{d}} = \text{piv}(NY)$</td><td>$t_{YL}, t_{LY}, t_{YR}, t_{RY}$</td><td>$t_{NL}, t_{LN}, t_{NR}, t_{RN}$</td></tr>
<tr><td>(a_{LN}, a_{RY}) with $Piv^1_{\underline{d}} = \text{piv}(LR, NY)$</td><td>$t_{NL}, t_{LN}, t_{NR}, t_{LY}$</td><td>$t_{YL}, t_{RN}, t_{YR}, t_{RY}$</td></tr>
<tr><td>(a_{LY}, a_{RN}) with $Piv^1_{\underline{d}} = \text{piv}(LR, NY)$</td><td>$t_{YL}, t_{LY}, t_{YR}, t_{LN}$</td><td>$t_{NL}, t_{RY}, t_{NR}, t_{RN}$</td></tr>
</table>

Table D1: Voter best responses for $\underline{d}$.

Proof. By the Corollary to Lemma C1, $Piv^1_{\underline{d}}$ must be a singleton. Case 1 in the proof of Proposition 1 shows that there exists a threshold $\bar{k}_{\underline{d}}(\mathbf{f}, \mathbf{a})$ above which conflicted voters in $\underline{d}$ condition their vote choice only on the event in $Piv^1_{\underline{d}}$. Thus, for $k > \bar{k}_{\underline{d}}(\mathbf{f}, \mathbf{a})$, voter best responses are given by Table D1. □

The key difference is that $\underline{d}$ may condition on $\text{piv}(LR, NY)$ alone. Corollary 1 remains valid for $\underline{d}$. Corollary 2 does not if $Piv^1_{\underline{d}} = \text{piv}(LR, NY)$ and platforms are either (a_{LY}, a_{RN}) or (a_{LN}, a_{RY}). In those cases, t_{ij} and t_{ji} would not always vote the same way. In equilibrium, however, Corollary 2 *does* hold as $\underline{d}$ never faces (a_{LY}, a_{RN}) or (a_{LN}, a_{RY}) platforms.

Proposition D2. *For any* $k > \bar{k}_{\underline{d}}(\mathbf{f}, \mathbf{a})$*, equilibrium platforms in* $\underline{d}$ *are* (a_{LN}, a_{RN}) *if condition (i) in Assumption A holds;* (a_{LY}, a_{RY}) *if condition (ii) in Assumption A holds.*

Proof. I will prove the case of condition (i) as the proof for condition (ii) is symmetric. Condition (i) in Assumption A states $f^{\omega}_{\underline{d}}(t_N) > \max\{f^{\omega}_{\underline{d}}(t_L), 1 - f^{\omega}_{\underline{d}}(t_L)\}$ and $f^{\omega}_{\underline{d}}(\hat{t}_{LN}) > f^{\omega}_{\underline{d}}(t_L) > f^{\omega}_{\underline{d}}(\hat{t}_{LY})$ for $\omega \in \{0, 1\}$.

By Proposition D1 we have that $Piv^1_{\underline{d}}$ will be either $\text{piv}(NY)$, $\text{piv}(LR, NY)$ or $\text{piv}(LR)$ in any continuation game. If $Piv^1_{\underline{d}} = \text{piv}(NY)$, then step 2 in the proof of Proposition 2 already shows that each candidate has unique best response under condition (i), giving equilibrium platforms (a_{LN}, a_{RN}). If $Piv^1_{\underline{d}} = \text{piv}(LR, NY)$, the problem candidates face is equivalent to that in the single district model of Appendix E. Equilibrium platforms are shown in Table E2. Under condition (i) each candidate has a unique best response, giving equilibrium platforms (a_{LN}, a_{RN}). If $Piv^1_{\underline{d}} = \text{piv}(LR)$ any candidate strategy is a best response, as the reform dimension is irrelevant. Therefore, (a_{LN}, a_{RN}) is the unique outcome under condition (i) as long as $\Pr[Piv^1 = \text{piv}(LR)] < 1$. This inequality is guaranteed by Lemma A1.

□

Appendix E. Single-District Model

In this section, I characterise the equilibria of a single-district model first developed by Krasa and Polborn (2010). Two candidates compete in a single district. As in the legislative model, each candidate is constrained on the left-right dimension but free to choose a pro- or anti-reform policy. Voters vote by majority rule, and the winning candidate implements his platform as policy. Most of the results in this section follow directly from Krasa and Polborn (2010). However, they do not analyse the equilibrium properties as the size of the electorate increases.[2] By modelling the setup as a Poisson game and looking for asymptotic equilibria, I can show what happens to chosen platforms and implemented policies as the number of voters increase. In the legislative model, there was aggregate uncertainty over the distribution of voters in district $\underline{d}$. For simplicity, in the single district model here and in Appendix F, I do not explicitly model aggregate uncertainty. One can think of the uncertainty as so small that it has no effect on the relative frequency of various voter types when all districts are combined into a single district.[3]

The voting stage is straightforward in a single-district election. There are three differences between voting here and in legislative elections. First, voting can no longer be a strategic choice because voters can only be pivotal between the two policy platforms offered by candidates. For each platform pair a voter faces, her preferred option is pinned down by her type, so her vote choice is deterministic. Table E1 shows the voting behaviour of each type in each of the four scenarios.

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

Table E1: Best responses of each voter type

Second, from Table E1, we see that preference intensity does matter here. Types t_{ij} and t_{ji} vote the same way in only three out of the four cases. This creates a conflict of interest between voter groups that want the same policy implemented but disagree about which is the second-best policy. Third, we see that t_{LN} and t_{LY} types always vote v_L while t_{RN} and t_{RY} types always vote v_R. The fact that these voters are never swing voters means candidates can safely ignore them when making their platform choices.

At the candidate competition stage, knowing how each voter type will vote, the left candidate chooses μ_L to maximise the expression below, while the right candidate chooses μ_R to minimise it.

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

[3] That is, despite any aggregate uncertainty, candidates know which case in Table E2 they face.

$$\begin{aligned}\Pr[Lwin \mid \mu_L, \mu_R, \sigma_d, f_d, k] = \big(\mu_L\mu_R + (1-\mu_L)(1-\mu_R)\big) &\Pr[Lwin \mid (a_{LN}, a_{RN}), \sigma_d, f_d, k] \\ + \mu_L(1-\mu_R) &\Pr[Lwin \mid (a_{LY}, a_{RN}), \sigma_d, f_d, k] \qquad \text{(E1)} \\ + (1-\mu_L)\mu_R &\Pr[Lwin \mid (a_{LN}, a_{RY}), \sigma_d, f_d, k].\end{aligned}$$

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

Proposition E1. *For any given f_d, the candidate equilibrium in a single-district election is unique. The equilibrium strategies are represented in Table E2.*

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

Table E2: Equilibrium platforms in the single-district candidate competition game. As $k \to \infty$, mixed strategy equilibria converge as shown.

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

Step 1: *Equilibrium strategies.*

The left candidate will choose μ_L to maximise Equation E1, while the right candidate will choose μ_R to minimise it. This yields the following best response correspondences:

$$BR_L(\mu_R) = \begin{cases} 1 & \text{if } \mu_R > \bar{\mu}_R \\ [0,1] & \text{if } \mu_R = \bar{\mu}_R \\ 0 & \text{if } \mu_R < \bar{\mu}_R \end{cases} \qquad \text{(E2)}$$

$$BR_R(\mu_L) = \begin{cases} 1 & \text{if } \mu_L < \bar{\mu}_L \\ [0,1] & \text{if } \mu_L = \bar{\mu}_L \\ 0 & \text{if } \mu_L > \bar{\mu}_L \end{cases} \tag{E3}$$

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

$$\bar{\mu}_L \equiv \frac{\Pr[Lwin|(a_{LN}, a_{RN})] - \Pr[Lwin|(a_{LN}, a_{RY})]}{2\Pr[Lwin|(a_{LN}, a_{RN})] - \Pr[Lwin|(a_{LY}, a_{RN})] - \Pr[Lwin|(a_{LN}, a_{RY})]} \tag{E4}$$

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

Where for ease of comprehension, I suppress explicit dependence of the probability of winning on σ_d, f_d and k. We can also write Equation E4 as

$$\bar{\mu}_L \equiv \frac{\Pr[Rwin|(a_{LN}, a_{RN})] - \Pr[Rwin|(a_{LN}, a_{RY})]}{2\Pr[Rwin|(a_{LN}, a_{RN})] - \Pr[Rwin|(a_{LY}, a_{RN})] - \Pr[Rwin|(a_{LN}, a_{RY})]} \tag{E5}$$

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

- **Case 1:** If $f_d(\hat{t}_{LN}) > f_d(t_L) > f_d(\hat{t}_{LY})$, voter strategies given in Table E1 imply that $\Pr[Lwin|(a_{LN}, a_{RY})] > \Pr[Lwin|(a_{LN}, a_{RN})] > \Pr[Lwin|(a_{LY}, a_{RN})]$. Examining Equation E1, we have a corner solution - both are playing a best response at $(\mu_L, \mu_R) = (0, 0)$.

- **Case 2:** If $f_d(\hat{t}_{LY}) > f_d(t_L) > f_d(\hat{t}_{LN})$, voter strategies given in Table E1 imply that $\Pr[Lwin|(a_{LY}, a_{RN})] > \Pr[Lwin|(a_{LN}, a_{RN})] > \Pr[Lwin|(a_{LN}, a_{RY})]$. Examining Equation E1, we have a corner solution - both are playing a best response at $(\mu_L, \mu_R) = (1, 1)$.

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

  - **Case 3a:** With $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L), 0.5$ and voter strategies given in Table E1 it must be that $\Pr[Lwin|(a_{LY}, a_{RN})], \Pr[Lwin|(a_{LN}, a_{RY})] > \Pr[Lwin|(a_{LN}, a_{RN})], 0.5$. This in turn means that $\bar{\mu}_L \in (0, 1)$ and $\bar{\mu}_R \in (0, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 3b:** With $0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$ and voter strategies given in Table E1 it must be that $\Pr[Lwin|(a_{LN}, a_{RY})], 0.5 > \Pr[Lwin|(a_{LY}, a_{RN})] > \Pr[Lwin|(a_{LN}, a_{RN})]$. This in turn means that $\bar{\mu}_L \in (0.5, 1)$ and $\bar{\mu}_R \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 3c:** With $0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$ and voter strategies given in Table E1 it must be that $\Pr[Lwin|(a_{LY}, a_{RN})], 0.5 > \Pr[Lwin|(a_{LN}, a_{RY})] > \Pr[Lwin|(a_{LN}, a_{RN})]$. This in turn means that $\bar{\mu}_L \in (0, 0.5)$ and $\bar{\mu}_R \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.

- **Case 4:** If $f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$, then $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$. There are 3 sub-cases to consider here:
  - **Case 4a:** With $0.5, f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$ and voter strategies given in Table E1 it must be that $\Pr[Lwin|(a_{LN}, a_{RN})], 0.5 > \Pr[Lwin|(a_{LY}, a_{RN})], \Pr[Lwin|(a_{LN}, a_{RY})]$. This in turn means that $\bar{\mu}_L \in (0, 1)$ and $\bar{\mu}_R \in (0, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 4b:** With $f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$ and voter strategies given in Table E1 it must be that $\Pr[Lwin|(a_{LN}, a_{RN})] > \Pr[Lwin|(a_{LN}, a_{RY})] > \Pr[Lwin|(a_{LY}, a_{RN})], 0.5$. This in turn means that $\bar{\mu}_L \in (0, 0.5)$ and $\bar{\mu}_R \in (0.5, 1)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.
  - **Case 4c:** With $f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$ and voter strategies given in Table E1 it must be that $\Pr[Lwin|(a_{LN}, a_{RN})] > \Pr[Lwin|(a_{LY}, a_{RN})] > \Pr[Lwin|(a_{LN}, a_{RY})], 0.5$. This in turn means that $\bar{\mu}_L \in (0.5, 1)$ and $\bar{\mu}_R \in (0, 0.5)$ and that both are playing a best response at $(\mu_L, \mu_R) = (\bar{\mu}_L, \bar{\mu}_R)$.

Step 2: *Equilibrium strategies as $k \to \infty$.*
We can now examine what happens to $(\bar{\mu}_L, \bar{\mu}_R)$ as $k \to \infty$ in cases 3 and 4.

- **Case 3a:** With $f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN}) > f_d(t_L), 0.5$ and voter strategies given in Table E1 it must be that $\nu_d[(a_{LY}, a_{RN})], \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LN}, a_{RN})], 0.5$.
  This implies $mag[Lwin|(a_{LY}, a_{RN})] = mag[Lwin|(a_{LN}, a_{RY})] = 0$
  and $mag[Rwin|(a_{LN}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RY})], mag[Rwin|(a_{LY}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation E5 by $\Pr[Rwin|(a_{LN}, a_{RN})]$ to get

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

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

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

- **Case 3b:** With $0.5, f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}) > f_d(t_L)$ and voter strategies given in Table E1 it must be that $0.5, \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RN})]$. This implies $mag[Lwin|(a_{LN}, a_{RY})] > mag[Lwin|(a_{LY}, a_{RN})] > mag[Lwin|(a_{LN}, a_{RN})]$.

  We can divide the top and bottom of Equation E4 by $\Pr[Lwin|(a_{LN}, a_{RY})]$ to get

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

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

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

- **Case 3c:** With $0.5, f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}) > f_d(t_L)$ and voter strategies given in Table E1 it must be that $0.5, \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LN}, a_{RN})]$. This implies $mag[Lwin|(a_{LY}, a_{RN})] > mag[Lwin|(a_{LN}, a_{RY})] > mag[Lwin|(a_{LN}, a_{RN})]$.

  We can divide the top and bottom of Equation E4 by $\Pr[Lwin|(a_{LY}, a_{RN})]$ to get

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

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

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

- **Case 4a:** With $0.5, f_d(t_L) > f_d(\hat{t}_{LY}), f_d(\hat{t}_{LN})$ and voter strategies given in Table E1 it must be that $0.5, \nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN})], \nu_d[(a_{LN}, a_{RY})]$.
  This implies $mag[Lwin|(a_{LN}, a_{RN})] > mag[Lwin|(a_{LN}, a_{RY})], mag[Lwin|(a_{LY}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation E4 by $\Pr[Lwin|(a_{LN}, a_{RN})]$ to get

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

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

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

- **Case 4b:** With $f_d(t_L) > f_d(\hat{t}_{LN}) > f_d(\hat{t}_{LY}), 0.5$ and voter strategies given in Table E1 it must be that $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})] > \nu_d[(a_{LY}, a_{RN})], 0.5$. This implies $mag[Rwin|(a_{LY}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RY})] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation E5 by $\Pr[Rwin|(a_{LY}, a_{RN})]$ to get

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

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

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

- **Case 4c:** With $f_d(t_L) > f_d(\hat{t}_{LY}) > f_d(\hat{t}_{LN}), 0.5$ and voter strategies given in Table E1 it must be that $\nu_d[(a_{LN}, a_{RN})] > \nu_d[(a_{LY}, a_{RN})] > \nu_d[(a_{LN}, a_{RY})], 0.5$. This implies $mag[Rwin|(a_{LN}, a_{RY})] > mag[Rwin|(a_{LY}, a_{RN})] > mag[Rwin|(a_{LN}, a_{RN})] > -1$.

  We can divide the top and bottom of Equation E5 by $\Pr[Rwin|(a_{LN}, a_{RY})]$ to get

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

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

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

□

Examining Table E2, a number of differences stand out from the legislative model in Table 2.

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

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

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

Finally, from Table E2 we can see that a Condorcet winner (if it exists) may not win a single-district election. This can happen in two different ways. First, it may be that there is a Condorcet winner but preferences correspond to case 3a or 4a so that candidates randomise equally over platforms. Here, with probability 0.5, the Condorcet winner policy will not be on the ballot. Second, in cases 1 and 2, a Condorcet winner policy may exist but not be chosen by a candidate. For example, this occurs in case 1 if $f_d(t_Y) > f_d(t_N)$ so that a_{LY} or a_{RY} is a Condorcet winner, but the equilibrium platforms are (a_{LN}, a_{RN}).

Appendix F. Applications

Parliamentary v Presidential Systems

One of the key choices any new democracy faces is whether to become a presidential or parliamentary system. A vast literature in economics and political science has studied the effect of having either system on the size and composition of government spending, growth, responses to crises, tax rates, corruption, and electoral campaign spending.[4] Any analysis must first define how exactly these two systems differ. Persson and Tabellini (2005) consider the difference to be along two dimensions: (i) presidential systems have greater separation of power between the executive and legislature, and (ii) presidents do not require the confidence of the legislature to remain in office. In reality, however, there is a large variance in the degree of separation of power both within and across these two systems. In particular, in many strong presidential systems, such as Brazil, Chile and the Philippines, legislative power is highly concentrated in the hands of the president. Equally, in a strong parliamentary system, power rests with legislators rather than the executive (prime-minister).[5] Here, I focus on a stark comparison between a strong presidential system and a strong parliamentary system. In the former, a president is directly elected by a national (single-district) election and he/she alone determines policy. In the latter, voters elect the legislature and a legislative majority on each dimension determines policy. These represent two extremes.[6] One can think of regimes where policy outcomes are determined jointly by the legislature and executive to varying degrees as existing between these extremes.[7] We have already seen that the incentives of voters are different under these two types of elections and that this results in different candidate platforms and implemented policies. But which system is best for voters?

If there was only one dimension of policy, I could compare under which system the expected national median voter $\widetilde{m}$ fares better. However, with multiple dimensions of policy there is no single median voter. The natural equivalent is to examine whether the median voter on each dimension fares better under a presidential or parliamentary system. A further complication in comparing welfare between these two systems is that the respective (expected) national median voters $\widetilde{m}^{LR} \in \{t_L, t_R\}$ and $\widetilde{m}^{NY} \in \{t_N, t_Y\}$ need not coincide with the median voters in the median district on each dimension $\widetilde{m}^{LR}_{d^*} \in \{t_L, t_R\}$ and $\widetilde{m}^{NY}_{d^*} \in \{t_N, t_Y\}$.[8] Whether they coincide or not will depend on how voters are distributed across districts. For example, in a US context it could be that the national median voter on the left-right dimension is a Democrat but that the median voter in the median district (Illinois's 13th District) is a Republican.[9] Similarly, the national median on the social issues

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

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

[6]They closely resemble the pure presidential model in Persson, Roland and Tabellini (1997) and the simple legislature in Persson, Roland and Tabellini (2000).

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

[8]$\widetilde{m}^{LR}$ and $\widetilde{m}^{NY}$ are respectively at the 50th percentile of the national voter population on that dimension, but $\widetilde{m}^{LR}_{d^*}$ and $\widetilde{m}^{NY}_{d^*}$ can be anywhere between the 25th and 75th percentile.

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

dimension may be pro-choice but the median in the median district on that dimension may be pro-life. Such cases occur if the distribution of voter preferences across districts is highly asymmetric. In many cases, however, the identities of the national median and the median voter in the median district will coincide. If this is so, the next proposition says that the median voter on each dimension is better off under a parliamentary system than a presidential one.

Proposition F1. *For $k > \bar{k}$, consider any distributions of voter preferences across districts such that $\widetilde{m}^h = \widetilde{m}^h_{d^*} \forall h \in \{LR, NY\}$. Then, as $k \to \infty$, with probability $\to 1$, the utility of the median voter $\widetilde{m}^h$ under a parliamentary system is weakly greater than under a presidential system.*

Proof. By applying $\widetilde{m}^h = \widetilde{m}^h_{d^*} \forall h \in \{LR, NY\}$ to Proposition 4, we know that as $k \to \infty$ then with probability going to one, the implemented policy in a parliamentary system is that which is preferred by $\widetilde{m}^h$ on dimension $h \in \{LR, NY\}$. All that remains to show is that the utility of $\widetilde{m}^h$ under a presidential system is the same for some voter distributions, but strictly lower for the rest. As a presidential election is simply a national single-district election, the mapping from preferences to policy outcomes is given by Table E2. There are four broad preference categories to consider:

(i) $f_d(t_L), f_d(t_N) > 0.5$:
Here, $\widetilde{m}^{LR} \in t_L$ and $\widetilde{m}^{NY} \in t_N$. From Table E2, we see z_{LN} is implemented only in case 1 and case 4b. For all other feasible preferences (cases 2, 3a, 4a, 4c), the expected implemented policy does not maximise the utility of $\widetilde{m}^h$.

(ii) $f_d(t_L) > 0.5 > f_d(t_N)$:
Here, $\widetilde{m}^{LR} \in t_L$ and $\widetilde{m}^{NY} \in t_Y$. From Table E2, we see z_{LY} is implemented only in case 2 and case 4c. For all other feasible preferences (cases 1, 3a, 4a, 4b), the expected implemented policy does not maximise the utility of $\widetilde{m}^h$.

(iii) $f_d(t_N) > 0.5 > f_d(t_L)$:
Here, $\widetilde{m}^{LR} \in t_R$ and $\widetilde{m}^{NY} \in t_N$. From Table E2, we see z_{RN} is implemented only in case 1 and case 3b. For all other feasible preferences (cases 2, 3a, 3c, 4a), the expected implemented policy does not maximise the utility of $\widetilde{m}^h$.

(iv) $f_d(t_L), f_d(t_N) < 0.5$:
Here, $\widetilde{m}^{LR} \in t_R$ and $\widetilde{m}^{NY} \in t_Y$. The policy z_{RY} is implemented in case 2 and 3c. For all other feasible preferences (cases 1, 3a, 3b, 4a), the expected implemented policy does not maximise the utility of $\widetilde{m}^h$.

Thus, while the utility of the median voter $\widetilde{m}^h$ is the same under a presidential system for some voter preference distributions, it is strictly lower in others.

□

Voting Index.

That is, for some distributions of voter preferences, the median voter on each dimension gets the same utility under both systems, while under all other distributions they get strictly greater utility under a parliamentary system. From Proposition 4, we know that the implemented policy in a legislative election is that preferred by the median voter in the median district on each dimension. If these types coincide with national median voters, then a parliamentary election always implements the policies favoured by the national medians on each dimension. Instead, presidential systems suffer from the problems of single-district elections laid out in Appendix E - candidates choose policies to attract swing voters, resulting in final policies which may not be the preferred choice of the median voter on each dimension. The result is somewhat counter-intuitive: presidential systems have a direct election to choose their policy-maker yet it is parliamentary systems with their indirect election of policy-makers (the median legislator on each dimension) that are more representative of voters' preferences. This difference occurs because there is more than one dimension of policy. In a single-district (presidential) election, voters must choose between two bundles of policies - it is this bundling of issues that can lead to sub-optimal policies. In a legislative (parliamentary) election, voters can act strategically to unbundle the issues and vote on a dimension-by-dimension basis. It is this strategic play of voters that brings about more representative policies under parliamentary systems.

Polarisation

There has been much debate about the causes and consequences of political polarisation in the US and around the world.[10] In this section, I ask whether increased polarisation can affect candidate platforms and implemented policies; and whether the impact differs between single-district and legislative elections.

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

Notice that because polarisation simply shifts voters' preference intensity, any increase in polarisation will leave the respective median voters $\widetilde{m}^h$ and $\widetilde{m}^h_{d^*}$ for $h \in \{LR, NY\}$ unchanged. As such, polarisation does not alter the *optimal* policy, but as the next proposition shows, it may change the *implemented* policy.

Proposition F2. *In a legislative election, polarisation will have no effect on voter behaviour, candidate platforms or implemented policies for $k > \bar{k}$ so long as Assumption A holds. In a single-district election, for any f_d^1 such that $f_d^1(t) > 0 \forall t \in \mathcal{T}$, there always exists a distribution with increased polarisation f_d^2 such that candidate platforms and implemented policies differ from those under f_d^1.*

Proof. The fact that polarisation has no effect on voter behaviour, candidate platforms, or implemented policies in legislative elections (as long as Assumption A holds) follows directly from Corollary 2 to Proposition 1.

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

Now, consider single-district elections. Take a distribution f_d^1 such that $f_d^1(t) > 0$ for all $t \in \mathcal{T}$, and assume without loss of generality that $f_d(t_L), f_d(t_N) > 0.5$, so that $\widetilde{m}^{LR} \in t_L$ and $\widetilde{m}^{NY} \in t_N$. Without further restrictions, the equilibrium can be any of the cases {1, 2, 3a, 4a, 4b, 4c} in Table E2, depending on the relative sizes of $f_d(\hat{t}_{LN})$, $f_d(t_L)$, and $f_d(\hat{t}_{LY})$.

For the more polarised distribution f_d^2, polarisation does not change the total share of t_L voters, so $f_d^2(t_L) = f_d^1(t_L) > 0.5$. I will now show that a polarised distribution f^2 can move the equilibrium from any of the initial cases to any other. To show this, I consider four distinct cases based on the relative frequencies of key voter types:

(i) <u>If $f_d^1(\hat{t}_{LN}) > f_d^1(t_L)$, there exists f_d^2 such that $f_d^2(\hat{t}_{LN}) < f_d^2(t_L)$.</u>
Since $f_d^1(\hat{t}_{LN}) > f_d^1(t_L)$, it must be that $f_d^1(t_{NR}) > f_d^1(t_{YL}) > 0$. There exists a f_d^2 where $f_d^2(t_{NR}) \in (0, f_d^1(t_{NR}))$. Hence, $f_d^2(t_{NR}) < f_d^1(t_{YL}) = f_d^2(t_{YL})$, leading to $f_d^2(\hat{t}_{LN}) < f_d^2(t_L)$.

(ii) <u>If $f_d^1(\hat{t}_{LN}) < f_d^1(t_L)$, there exists f_d^2 such that $f_d^2(\hat{t}_{LN}) > f_d^2(t_L)$.</u>
Here, $f_d^1(t_{YL}) > f_d^1(t_{NR}) > 0$. There exists a f_d^2 where $f_d^2(t_{YL}) \in (0, f_d^1(t_{YL}))$. Thus, $f_d^2(t_{YL}) < f_d^1(t_{NR}) = f_d^2(t_{NR})$, leading to $f_d^2(\hat{t}_{LN}) > f_d^2(t_L)$.

(iii) <u>If $f_d^1(\hat{t}_{LY}) > f_d^1(t_L)$, there exists f_d^2 such that $f_d^2(\hat{t}_{LY}) < f_d^2(t_L)$.</u>
Since $f_d^1(\hat{t}_{LY}) > f_d^1(t_L)$, it must be that $f_d^1(t_{YR}) > f_d^1(t_{NL}) > 0$. There exists a f_d^2 such that $f_d^2(t_{YR}) \in (0, f_d^1(t_{YR}))$, so $f_d^2(t_{YR}) < f_d^1(t_{NL}) = f_d^2(t_{NL})$, giving $f_d^2(\hat{t}_{LY}) < f_d^2(t_L)$.

(iv) <u>If $f_d^1(\hat{t}_{LY}) < f_d^1(t_L)$, there exists f_d^2 such that $f_d^2(\hat{t}_{LY}) > f_d^2(t_L)$.</u>
Here, $f_d^1(t_{NL}) > f_d^1(t_{YR}) > 0$. There exists a f_d^2 where $f_d^2(t_{NL}) \in (0, f_d^1(t_{NL}))$. Thus, $f_d^2(t_{NL}) < f_d^1(t_{YR}) = f_d^2(t_{YR})$, leading to $f_d^2(\hat{t}_{LY}) > f_d^2(t_L)$.

Notice that in each of the four scenarios, we change the distribution of a different voter type. Any change to $f_d(t_{YL})$ or $f_d(t_{NR})$ does not affect the ordering of $f_d(\hat{t}_{LY})$ and $f_d(t_L)$. Similarly, changes to $f_d(t_{YR})$ or $f_d(t_{NL})$ do not affect the ordering of $f_d(\hat{t}_{LN})$ and $f_d(t_L)$. Therefore, regardless of the initial ordering of $f_d^1(\hat{t}_{LY})$, $f_d^1(\hat{t}_{LN})$, and $f_d^1(t_L)$, any reordering can be achieved by selecting an appropriate f_d^2. Hence, polarisation can shift the equilibrium from any one in {1, 2, 3a, 4a, 4b, 4c} to any other. A similar argument holds for the case $f_d(t_L) > 0.5$, $f_d(t_Y) > 0.5$, while for $f_d(t_R) > 0.5$, the relevant equilibria in Table E2 are {1, 2, 3a, 3b, 3c, 4a}. In all cases, increased polarisation can shift a single-district election from one equilibrium to another. □

The fact that polarisation has no impact in a legislative election stems directly from Corollary 2 to Proposition 1 - types t_{ij} and t_{ji} always vote the same way. The requirement that Assumption A holds is because it is possible to increase polarisation in $\underline{d}$ in such a way that the assumption no longer holds.

In a single-district election, an increase in polarisation will change the composition of swing voters, which in turn changes the platforms candidates campaign on and the policies that are implemented. To take an extreme example, if polarisation is so pronounced that only core party supporters $t_{LN}, t_{LY}, t_{RN}, t_{RY}$ exist - we get multiple equilibria. As there are no more swing voters, choosing pro- or anti-reform policies are both best responses for each candidate. More generally, changes in polarisation in a single-district election can lead to any of the four policies being implemented despite no change in $\widetilde{m}^{LR}$, $\widetilde{m}^{NY}$. As the implemented policy in single-district elections may already have been sub-optimal, an increase in polarisation may either increase or decrease the utility of the median voter. In a legislative election, polarisation has no effect on outcomes, so the implemented policy remains that preferred by $\widetilde{m}^{LR}_{d^*}$ on the left-right dimension and $\widetilde{m}^{NY}_{d^*}$ on the reform dimension.

Though Proposition F2 considers the specific form of polarisation defined above, the result is more general. For any f^1_d, f^2_d such that $f^1_d(t_L) = f^2_d(t_L)$ and $f^1_d(t_N) = f^2_d(t_N)$, voter behaviour, candidate platforms and implemented policies in a legislative election will not vary when moving from f^1_d to f^2_d. In a single-district election, as before, one can always find a f^2_d which satisfies the constraints but leads to different equilibrium outcomes than f^1_d. This more general result encompasses the case where voter preferences become highly correlated across dimensions, something widely documented among US voters (Barber and McCarty, 2015).

References

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

Gentzkow, Matthew. 2016. "Polarization in 2016." *Toulouse Network of Information Technology whitepaper*, 1–23.

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

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

Myerson, Roger B. 2000. "Large Poisson Games." *Journal of Economic Theory*, 94(1): 7–45.

Persson, Torsten, and Guido Tabellini. 2005. *The economic effects of constitutions.* Vol. 54, The MIT Press.

Persson, Torsten, Gérard Roland, and Guido Tabellini. 1997. "Separation of powers and political accountability." *Quarterly Journal of Economics*, 112(November): 1163–1202.

Persson, Torsten, Gérard Roland, and Guido Tabellini. 2000. "Comparative Politics and Public Finance." *Journal of Political Economy*, 108(6): 1121–1161.

Shugart, Matthew Soberg, John M Carey, et al. 1992. *Presidents and assemblies: Constitutional design and electoral dynamics.* Cambridge University Press.