Online Appendix: Strategic Mobilization
Why Proportional Representation Decreases Voter Mobilization

Carlisle Rainey[†]

Contents

A Proofs

A.1 Proof of Lemma 1 (Equal Effort)

In order for the strategy profile $S^* = (s_w^*, s_s^*)$ to be a Nash equilibrium, the first-order conditions, $\frac{\partial u_w}{\partial \epsilon_w} = 0$ and $\frac{\partial u_s}{\partial \epsilon_s} = 0$, and second order conditions, $\frac{\partial^2 u_w}{\partial^2 \epsilon_w} < 0$ and $\frac{\partial^2 u_s}{\partial^2 \epsilon_s} < 0$, must hold. Since the utility functions generate a discontinuous payoff space, I must also show that neither party has an incentive to defect from the proposed equilibrium to zero effort.

    Remark: The algebra below assumes that $\epsilon_i \neq 0$ in equilibrium. No Nash equilibrium exists in which any player chooses zero effort. If both players choose zero effort, then both have an

incentive to defect to an arbitrarily small amount of effort to win the whole prize. If only one player chooses to play zero effort, then the other has an incentive to play as close to zero as possible. Because no positive real number closest to zero exists, no Nash equilibrium exists where one player plays zero.

Given that $\epsilon_i \neq 0$ in equilibrium, Equation 1 gives the first order conditions for u_w.

$$\frac{\partial u_w}{\partial \epsilon_w} = \frac{\delta \beta^\delta \epsilon_w^{\delta-1} \epsilon_s^\delta}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^2} - 1 = 0$$

$$\frac{\delta \beta^\delta \epsilon_w^{\delta-1} \epsilon_s^\delta}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^2} = 1 \tag{1}$$

Equation 2 presents the similar first order condition for s.

$$\frac{\partial u_s}{\partial \epsilon_s} = \frac{\delta \beta^\delta \epsilon_w^\delta \epsilon_s^{\delta-1}}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^2} - 1 = 0$$

$$\frac{\delta \beta^\delta \epsilon_w^\delta \epsilon_s^{\delta-1}}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^2} = 1 \tag{2}$$

Since both the left hand side of Equation 1 and 2 equal one, setting one equal to the other shows that the first-order conditions hold if and only if the parties exert equal effort.

$$\frac{\delta \beta^\delta \epsilon_w^{\delta-1} \epsilon_s^\delta}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^2} = \frac{\delta \beta^\delta \epsilon_w^\delta \epsilon_s^{\delta-1}}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^2}$$

$$\epsilon_w = \epsilon_s$$

Given that parties must exert equal effort in any pure strategy Nash equilibrium, the second order conditions confirm that such a strategy maximizes (rather than minimizes) the local utility. Equation 3 gives the second-order condition for u_w.

$$\frac{\partial^2 u_w}{\partial^2 \epsilon_w} = \frac{\delta(\delta-1)\beta \epsilon_w^{\delta-2} \epsilon_s^\delta}{(\beta \epsilon_w^\delta + \epsilon_s^\delta)^2} - \frac{2\delta^2 \beta \epsilon_w^{2\delta-2} \epsilon_s^\delta}{(\beta \epsilon_w^\delta + \epsilon_s^\delta)^3} < 0 \tag{3}$$

After applying Lemma 1 and noting that δ and β are restricted to positive values, algebraic simplification shows that Equation 3 holds if and only if $\beta > \left(\frac{\delta-1}{\delta+1}\right)^{\frac{1}{\delta}}$. Next, I show that the second-order condition holds for s.

$$\frac{\partial^2 u_s}{\partial^2 \epsilon_s} = \frac{\delta(\delta-1)\beta^\delta \epsilon_w^\delta \epsilon_s^{\delta-2}}{[(\beta \epsilon_w^\delta) + \epsilon_s^\delta]^2} - \frac{2\delta^2 \beta^{2\delta} \epsilon_w^\delta \epsilon_s^{2\delta-2}}{[(\beta \epsilon_w)^\delta + \epsilon_s^\delta]^3} < 0 \tag{4}$$

After again applying Lemma 1 and noting that δ and β are restricted to positive values, algebraic simplification shows that Equation 4 holds if and only if $\beta > \left(\frac{\delta-1}{3-\delta}\right)^{\frac{1}{\delta}}$.

2

Proof of Lemma 2 (Equilibrium Strategy)

Given Lemma 1 (Equal Effort), Equation 1 or 2 reduces to $\dfrac{\delta\beta^\delta\epsilon^{2\delta-1}}{(\beta^\delta+1)^2\epsilon^{2\delta}}=1$. Algebraic manipulation gives the equilibrium strategy, $s^*(\delta,\beta)=\epsilon=\delta\dfrac{\beta^\delta}{(1+\beta^\delta)^2}$. But since the utility functions produce a discontinuous payoff space, I must also show that neither party as an incentive to defect from the proposed equilibrium to the boundary of zero effort. First, note by Lemma 1 that $u_w < u_s$ for all $\beta \in (0,1)$ and $\delta > 1$. Therefore, if w has no incentive to change from the proposed equilibrium, then neither does s. To see when w has an incentive to change from the proposed equilibrium to zero effort, simply plug the proposed equilibrium effort into the utility function for w and set the payoff function greater than or equal to zero. This shows that neither party has an incentive to switch to zero for $\beta \geq (\delta-1)^{\frac{1}{\delta}}$. Any set of parameter values which satisfy this condition also satisfy the second order conditions. Therefore, the equilibrium exists if and only if neither party has an incentive to change from the proposed equilibrium strategy to zero effort.

Proof of Proposition 1 (Competitiveness)

Given the equilibrium strategy of $s^*(\delta,\beta)=\epsilon=\delta\dfrac{\beta^\delta}{(\beta^\delta+1)^2}$, I simply take the derivative of the equilibrium strategy function, s^*, with respect to the competitiveness parameter, β, to find how equilibrium mobilization efforts change as competitiveness changes.

$$\frac{\partial s^*}{\partial \beta} = \frac{\delta^2\beta^{\delta-1}}{(\beta^\delta+1)^2} - \frac{2\delta^2\beta^{2\delta-1}}{(\beta^\delta+1)^3} \tag{5}$$

Algebraic manipulation yields $\dfrac{\partial s^*}{\partial \beta} = \dfrac{\delta^2\beta^{\delta-1}(1-\beta^\delta)}{(\beta^\delta+1)^3}$. Since δ is defined to be greater than one and β is defined to be between zero and one, this derivative is clearly positive. Thus, equilibrium effort is increasing in competitiveness.

A.2   Proof of Proposition 2 (Disproportionality)

Given the equilibrium strategy of $s^*(\delta,\beta)=\epsilon=\delta\dfrac{\beta^\delta}{(\beta^\delta+1)^2}$, I simply take the derivative of the equilibrium strategy function, s^*, with respect to the disproportionality parameter, δ, to find how equilibrium mobilization efforts change as disproportionality changes. This derivative is more complicated, but differentiation shows that $\dfrac{\partial s^*}{\partial \delta} = \dfrac{\beta^\delta+\delta\beta^\delta\ln(\beta)}{(\beta^\delta+1)^2} - \dfrac{2\beta^{2\delta}\ln(\beta)}{(\beta^\delta+1)^3}$. Algebraic manipulation shows that this derivative is positive if and only if $\delta\ln\beta > \dfrac{\beta^\delta+1}{\beta^\delta-1}$. This derivative

is positive for much of the parameter space for which the proposed equilibrium exists and only negative when both β and δ are extremely small.

A.3   Proof of Proposition 3 (Marginal Effect of Competitiveness)

Given the equilibrium strategy of $s^*(\delta, \beta) = \epsilon = \delta\dfrac{\beta^\delta}{(\beta^\delta + 1)^2}$, I simply take the second derivative of the equilibrium strategy function, s^*, first differentiating with respect to the competitiveness parameter, β, and then with respect to the disproportionality parameter, δ, to find how the marginal effect of competitiveness on equilibrium effort changes as disproportionality varies. This differentiation is complicated, but algebraic manipulation after setting it greater than zero yields the condition $\dfrac{1}{3} < \beta^\delta(1 + \beta^\delta)$. This inequality holds for almost all of the parameter space for which the pure strategy equilibrium exists.

B   Extensions of the Formal Model

B.1   Expanding the Two-Party Model to the n-Party Case

I briefly examine a less general model of party competition, but one that allows the value of the seat (or seats) in the contest to vary and allows more than two political parties. Making somehwat restrictive, but reasonable assumptions, I show that the key results from the two party model do not depend on the value of the prize (the number of seats competed over) or the number of parties in the competition. In this more complicated model, I assume that n parties, where $P = \{1, 2, ..., n\}$, compete over a seat or set of seats with value normalized to one. Each party places the same value on the seat (or set of seats). Again, I assume for simplicity that the prize is infinitely divisible. Similar to the simpler two-party model, I assume that the first k parties are unable to mobilize voters as efficiently as the last $n - k$ parties. This gives two types of parties, $j \in \{w, s\}$, with the weak type, w, only able to mobilize a fraction of the voters that the strong type, s, is able to mobilize with the same amount of effort. Assuming risk neutrality gives the utility functions for each type.

$$u_w(\epsilon_w, \epsilon_s) = \frac{(\beta\epsilon_w)^\delta}{k(\beta\epsilon_w)^\delta + (n-k)\epsilon_s^\delta} - \epsilon_w \tag{6}$$

$$u_s(\epsilon_w, \epsilon_s) = \frac{\epsilon_s^\delta}{k(\beta\epsilon_w)^\delta + (n-k)\epsilon_s^\delta} - \epsilon_s \tag{7}$$

Definition (Total Effort) Define total effort, τ_ϵ, as the sum of the effort exerted by each party competing in the contest. Formally, $\tau_\epsilon = \sum_{i=1}^{n} \epsilon_i$, for $i = \{1, 2, ..., n\}$.

Notice that for the special case of two parties discussed above, the total effort exerted by the political parties is simply two times the equilibrium effort given the competitiveness and

4

disproportionality of the contest. Formally, by Lemmas 1 and 2, $\tau_\epsilon(\delta,\beta) = 2\delta\dfrac{\beta^\delta}{(\beta^\delta+1)^2}$ for the special case of a contest between only two political parties. I now use this definition to briefly expand beyond the simple model laid out above and explore the effects of multiple parties and varying district magnitude on the equilibrium effort.

Lemma 1 *(n-party Case) The total effort, τ_ϵ, exerted in any contest does not depend on the number of parties in the contest. The number of parties in the contest only influences the amount of effort exerted by each individual party.*

Proving Lemma 3 requires first finding a strategy profile for each type, w and s that satisfies the first and second order conditions of their respective utility functions. Second, it requires showing the the total effort, τ_ϵ, as defined in Definition 1 does not change as the number of parties in the contest varies. While the equilibrium effort may depend on the ratio of types (i.e. $\dfrac{k}{(n-k)}$), which is an aspect of competitiveness, it may not depend on the total number of parties in the contest (i.e. increasing the number of parties while leaving the ratio of type the same).

I begin by examining the first-order condition for the weak type's utility function. The utility function of the weak type is given by $u_w(\epsilon_w,\epsilon_s) = \dfrac{(\beta\epsilon_w)^\delta}{k(\beta\epsilon_w)^\delta+(n-k)\epsilon_s^\delta} - \epsilon_w$. Taking the derivative with respect to ϵ_w, setting it equal to zero, and rearranging yields the first order condition for the weak type.

$$\frac{(n-k)\delta\beta^\delta\epsilon_w^{\delta-1}\epsilon_s^\delta}{[k(\beta\epsilon_w)^\delta+(n-k)\epsilon_s^\delta]^2} = 1 \tag{8}$$

I now examine the first-order conditions for the strong type's utility function. The utility function is given by $u_s(\epsilon_w,\epsilon_s) = \dfrac{\epsilon_s^\delta}{k(\beta\epsilon_w)^\delta+(n-k)\epsilon_s^\delta} - \epsilon_s$. Taking the derivative with respect to ϵ_s, setting it equal to zero, and rearranging yields the first order condition for the strong type.

$$\frac{k\delta\beta^\delta\epsilon_w^\delta\epsilon_s^{\delta-1}}{[k(\beta\epsilon_w)^\delta+(n-k)\epsilon_s^\delta]^2} = 1 \tag{9}$$

Setting the left hand sides of Equations 8 and 9 equal to each other and algebraic manipulating the result shows that in equilibrium the strategies of the weak and strong types are not always equal, but have the relationship $\dfrac{k}{n-k}\epsilon_w = \epsilon_s$.

Replacing ϵ_s in Equation 8 with $\dfrac{k}{n-k}\epsilon_w$ or replacing ϵ_w in Equation 9 with $\dfrac{n-k}{k}\epsilon_s$ and algebraically manipulating gives the identical equilibrium strategies for each type, s_w^* and s_s^*.

$$s_w^*(n,k,\delta,\beta) = \frac{k\delta\beta^\delta(\frac{k}{n-k})^\delta}{[k\beta^\delta + (n-k)(\frac{k}{n-k})^\delta]^2} \tag{10}$$

$$s_s^*(n,k,\delta,\beta) = \frac{k\delta\beta^\delta(\frac{k}{n-k})^{\delta+1}}{[k\beta^\delta + (n-k)(\frac{k}{n-k})^\delta]^2} \tag{11}$$

The above result gives the equilibrium effort for each individual party in the contest, depending on that party's type. All that remains is to show that these equilibrium strategies imply that the total effort, τ_ϵ does not vary with the number of parties. By the definition of total effort given in Definition 1, the total effort in the n-party contest is given by $\tau_\epsilon = k \times \dfrac{k\delta\beta^\delta(\frac{k}{n-k})^\delta}{[k\beta^\delta + (n-k)(\frac{k}{n-k})^\delta]^2} +$

$(n-k) \times \dfrac{k\delta\beta^\delta(\frac{k}{n-k})^{\delta+1}}{[k\beta^\delta + (n-k)(\frac{k}{n-k})^\delta]^2}$. Through algebraic manipulation, this reduces to $\tau_\epsilon =$

$2\delta\beta^\delta \dfrac{(\frac{k}{n-k})^{\delta-1}}{[\beta^\delta + (\frac{k}{n-k})^{\delta-1}]^2}$. Clearly, the total effort does not depend on the number of parties, only on the ratio of weak and strong types.

B.2   Expanding the Model to Allow for Prizes with Variable Value

I now consider the consequences of changing the value of the prize over which the parties are competing. Again, I make restrictive but reasonable assumptions and show that the results from the simpler two-party model do not change. There are two methods of increasing the magnitudes of districts in electoral systems. First, policy-makers might lump districts together. This increases the district magnitude, but also makes the district larger, increasing the geographical size of the district and the population. This change makes mobilization more costly. I assume that doubling the size of the district doubles the cost of mobilization, that tripling the size of the district triples the cost of mobilization, and so on. Another method to increase district magnitude is to increase the size of the legislature. This seems to have little or no influence on the cost of mobilization, since the geographic size and population of the district remain the same. However, increasing the size of the legislature reduces the value of each seat. It is reasonable to assume that doubling the size of the legislature reduces the value of each seat by one-half, tripling

the size of the legislature reduces the value of each seat by one-third, and so on. Either of these changes is captured by the assumption that the cost of effort is a function, c, of district magnitude, m, where $c(m, \epsilon)i) = m\epsilon_i$.

Lemma 2 *(Size of the Prize) Assuming that the cost of effort is a function, c, of district magnitude, m, where $c(m, \epsilon_i) = m\epsilon_i$, the number of seats awarded in any contest does not influence the total effort exerted by the parties, τ_ϵ.*

Changing the value of the seat or seats that parties compete over to v rather than normalizing to one has no effect on the equilibrium efforts as long as reasonable adjustments are applied to the cost. As I argue in the main text, increasing the size of the district should increase the cost of mobilization. Assuming that a competition over a seat or seats v times as valuable is v times as costly, the equilibrium effort levels do not change. It is straightforward to see this in the general model or two-party. I show this for the strategy of the weak type, and the proof for the strong type is almost identical.

The utility function of the weak type with variable values and associated changes in costs is given by $u_w(\epsilon_w, \epsilon_s) = v\dfrac{(\beta\epsilon_w)^\delta}{k(\beta\epsilon_w)^\delta + (n-k)\epsilon_s^\delta} - v\epsilon_w$. Taking the derivative with respect to ϵ_w, setting it equal to zero, and rearranging yields the first order condition for the weak type.

$$v\frac{(n-k)\delta\beta^\delta\epsilon_w^{\delta-1}\epsilon_s^\delta}{[k(\beta\epsilon_w)^\delta + (n-k)\epsilon_s^\delta]^2} = v \tag{12}$$

It is easy to see that the v's on the left and right sides cancel, yielding a first-order condition that does not depend on v and exactly the same first-order condition as the model without variable values.

C   Robustness of Empirical Results

In the first section of this appendix, I expand on the comments I make about the robustness of the empirical results in Footnote **??** of the main text. I first show that the conclusions do not depend on whether I control for the number of political parties. Next, I show that the results are robust to complete pooling (single logit on the entire data set). Finally, I show that the results are robust to no pooling across countries (separate logits in each country). For computational ease, I (1) use listwise deletion for all the robustness checks, (2) assume that the random effects in the hierarchical models are independent, and (3) place Gamma(2, 0.01) priors on the standard deviation parameters of the hierarchical models (see **?**). Inferences do not depend on multiple imputation. Relaxing independence does not change any conclusions (and the estimated correlation in the main text is near zero). Placing priors on the standard deviation parameter gently pushes these parameters away from zero in a couple of instances. All explanatory variables in this section are scaled to have a minimum of zero and a maximum of one, with the exception of competitiveness, which has a theoretical range from zero to one but actually varies from 0.24 to 0.93. This scaling dramatically speeds the convergence of the MLE estimation.

C.1 Controlling for Number of Political Parties Does Not Affect the Results

One potential confound that I do not include in the main model is the number of political parties. Prior work shows that proportional electoral rules lead to more political parties, so it might be important to take this factor into account. Table 1 presents a series of hierarchical models that show the results do not change when I include the effective number of political parties (at the district level) in the model or by including its interaction with electoral rules. The results are remarkably consistent across the different specifications.

Notice first that, as I expect, district competitiveness has a positive effect in SMDP systems (shown by the large, positive coefficient for competitiveness) and PR rules eliminate most of this positive effect (shown the large, negative coefficient for the product of competitiveness and PR rules). These findings hold regardless of the model specification. Next, notice that whether I include or exclude individual-level covariates in the hierarchical models, the the model specification that I rely on in the main text (excluding the number of parties) minimizes the BIC, a criterion commonly used to choose model specifications. However, Table 2 shows that the effect of the number of parties marginally statistically significant when individual-level controls are excluded from the model. But because the BIC points toward the model excluding the number of political parties, I present that model in the main text. However, the conclusions do not depend on this choice.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	−2.63*	−2.63*	−2.63*	−4.43*	−4.43*	−4.44*
	(0.42)	(0.43)	(0.43)	(0.46)	(0.46)	(0.46)
Competitiveness	3.62*	3.36*	3.23*	4.15*	4.15*	4.01*
	(0.52)	(0.54)	(0.55)	(0.66)	(0.68)	(0.70)
PR	0.14	−0.07	0.12	0.98	0.98	1.12
	(0.76)	(0.77)	(0.79)	(0.74)	(0.75)	(0.76)
Competitiveness × PR	−2.34*	−2.18*	−1.96*	−3.10*	−3.10*	−2.91*
	(0.98)	(0.98)	(0.99)	(1.06)	(1.07)	(1.08)
ENEP		1.11	1.68*		0.00	0.66
		(0.66)	(0.83)		(0.74)	(1.04)
ENEP × PR			−1.56			−1.30
			(1.38)			(1.47)
Age				0.38*	0.38*	0.38*
				(0.17)	(0.17)	(0.17)
Male				0.07	0.07	0.07
				(0.07)	(0.07)	(0.07)
Education				1.11*	1.11*	1.12*
				(0.17)	(0.17)	(0.17)
Married				0.02	0.02	0.02
				(0.08)	(0.08)	(0.08)
Union Member				0.24*	0.24*	0.24*
				(0.08)	(0.08)	(0.08)
Household Income				0.41*	0.41*	0.42*
				(0.14)	(0.14)	(0.14)
Close to a Party				0.43*	0.43*	0.43*
				(0.07)	(0.07)	(0.07)
BIC	8352.17	8358.18	8365.89	5615.87	5624.42	5632.20
Num. obs.	7651	7651	7651	5218	5218	5218
Num. groups: District	613	613	613	568	568	568
Num. groups: Country	5	5	5	5	5	5

$^{*}p < 0.05$

TABLE 1: This table shows that the findings presented in the main text are robust to the inclusion of ENEP. Notice also that excluding ENEP leads to the smallest BIC, although some models including ENEP fit the data statistically significantly better (see Table 2 for the details).

Comparison	F-test	Direction
Model 1 to Model 2	$p = 0.09$	+
Model 1 to Model 3	$p = 0.12$	+
Model 4 to Model 5	$p = 1.00$	0
Model 4 to Model 6	$p = 0.68$	+/−

TABLE 2: This table compares models that include and exclude ENEP. The Comparison column reports which two models from Table 1 are being compared. The F-test column reports the p-value that that larger model (including ENEP) fits the data statistically significantly better than than the smaller model (excluding ENEP). The Direction column reports the direction of the estimated effect of ENEP.

C.2 Estimating a Model with Complete Pooling Does Not Affect Results

Table 3 presents estimates using a non-hierarchical logit model that completely pools the data across countries and districts. These results show that the conclusions in the main text and the conclusions about the limited impact of the number of parties on these results are robust to a complete pooling estimation strategy. Notice that across all model specifications, competitiveness has a large, positive effect in SMDP systems (shown by the coefficient for competitiveness) while PR rules consistently undermine this effect (shown by the large, negative coefficient for the product of competitiveness and PR rules).

By design, the empirical model I present in the main text partially pools information across districts. Because districts in Canada and Great Britain have fewer respondents and thus less information per district, there is more partial pooling in these districts. One might wonder whether this has an unintended effect on the results. However, the results from Table 3 using complete pooling closely mirror those from Table 1 using partial pooling. The substantive conclusions do not change.

However, Table 4 shows that all specifications in Table 3 that include the number of political parties fit the data statistically significantly better than models excluding the number of political parties. However, notice that signs of the estimated effects depend on whether I include individual-level covariates. Also again notice that the model specification that I rely on in the main text minimizes the BIC.

	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Intercept	−1.89*	−1.88*	−1.88*	−4.12*	−4.17*	−4.18*
	(0.24)	(0.24)	(0.24)	(0.39)	(0.39)	(0.39)
Competitiveness	2.77*	2.60*	2.61*	3.50*	3.68*	3.55*
	(0.37)	(0.38)	(0.41)	(0.54)	(0.54)	(0.58)
PR	0.71*	0.61	0.61	1.91*	2.04*	2.06*
	(0.32)	(0.32)	(0.32)	(0.43)	(0.44)	(0.44)
Competitiveness × PR	−3.07*	−2.99*	−3.00*	−4.12*	−4.18*	−4.04*
	(0.47)	(0.47)	(0.51)	(0.64)	(0.64)	(0.67)
ENEP		0.65*	0.63		−0.78*	−0.20
		(0.26)	(0.68)		(0.32)	(0.92)
ENEP × PR			0.03			−0.65
			(0.73)			(0.98)
Age				0.29	0.34*	0.34*
				(0.17)	(0.17)	(0.17)
Male				0.05	0.05	0.05
				(0.07)	(0.07)	(0.07)
Education				1.38*	1.41*	1.41*
				(0.16)	(0.16)	(0.16)
Married				0.04	0.03	0.03
				(0.07)	(0.07)	(0.07)
Union Member				0.22*	0.27*	0.27*
				(0.07)	(0.08)	(0.08)
Household Income				0.34*	0.36*	0.36*
				(0.13)	(0.13)	(0.13)
Close to a Party				0.43*	0.42*	0.42*
				(0.07)	(0.07)	(0.07)
BIC	8597.55	8600.44	8609.38	5692.73	5695.14	5703.26
Num. obs.	7651	7651	7651	5218	5218	5218

$^*p < 0.05$

TABLE 3: This table shows that the findings presented in the main text are robust to the inclusion of ENEP and complete pooling. Notice again that excluding ENEP leads to the smallest BIC, although some models including ENEP fit the data statistically significantly better (see Table 4 for the details).

Comparison	F-test	Direction
Model 7 to Model 8	$p = 0.01$	+
Model 7 to Model 9	$p = 0.05$	+
Model 10 to Model 11	$p = 0.01$	−
Model 10 to Model 12	$p = 0.04$	−

TABLE 4: This table compares models that include and exclude ENEP. The Comparison column reports which two models from Table 1 are being compared. The F-test column reports the p-value that that larger model (including ENEP) fits the data statistically significantly better than than the smaller model (excluding ENEP). The Direction column reports the direction of the estimated effect of ENEP.

C.3 Estimating a Model with No Pooling Across Countries Does Not Affect the Results

The model I present in the main text assumes complete pooling across districts and elections. In this section, I also show that the results are robust to a no pooling assumption by estimating separate logistic regressions for each election in the data. This approach uses complete pooling across districts and no pooling across elections. The results are consistent with the main model (i.e. notice that competitiveness has a much larger effect in Canada and Great Britain). Table 5 presents results without individual-level controls, while Table 6 includes them.

	Great Britain (SMDP)	Canada (SMDP)	Finland (PR)	Portugal 2002 (PR)	Portugal 2005 (PR)
Intercept	−3.17*	−2.01*	−1.58*	−3.00*	−0.49*
	(0.53)	(0.29)	(0.61)	(0.47)	(0.25)
Competitiveness	3.61*	3.54*	0.33	1.90*	−1.15*
	(0.77)	(0.45)	(0.82)	(0.65)	(0.35)
Num. obs.	858	1650	1192	1254	2697

*$p < 0.05$

TABLE 5: This table shows that the findings presented in the main text are robust to estimating separate models in each country with no individual-level covariates. Notice that the estimated effect of competitiveness is much larger in SMDP countries than in PR countries.

	Great Britain (SMDP)	Canada (SMDP)	Finland (PR)	Portugal 2002 (PR)	Portugal 2005 (PR)
Intercept	−3.42*	−4.54*	−2.24*	−4.00*	−2.34*
	(0.63)	(0.67)	(0.72)	(0.75)	(0.37)
Competitiveness	3.73*	4.67*	−0.57	1.63*	−1.01*
	(0.83)	(0.76)	(0.90)	(0.83)	(0.42)
Age	0.00	1.53*	0.20	1.20*	0.30
	(0.37)	(0.60)	(0.36)	(0.54)	(0.29)
Male	−0.34*	−0.09	0.07	0.01	0.30*
	(0.16)	(0.16)	(0.15)	(0.20)	(0.11)
Education	0.29	1.08*	0.56	0.50	1.65*
	(0.42)	(0.47)	(0.44)	(0.54)	(0.26)
Married	0.02	0.18	−0.07	−0.14	0.02
	(0.18)	(0.18)	(0.18)	(0.23)	(0.13)
Union Member	0.15	0.13	0.36*	0.41	0.21
	(0.19)	(0.17)	(0.16)	(0.27)	(0.16)
Household Income	−0.11	0.10	0.57	1.24*	1.01*
	(0.29)	(0.28)	(0.32)	(0.52)	(0.26)
Close to a Party	0.44*	0.48*	0.79*	0.32	0.23*
	(0.17)	(0.17)	(0.16)	(0.20)	(0.12)
Num. obs.	757	682	1086	759	1934

*$p < 0.05$

TABLE 6: This table shows that the findings presented in the main text are robust to estimating separate models in each country with individual-level covariates. Notice that the estimated effect of competitiveness is much larger in SMDP countries than in PR countries.

I also replicate Figure **??** from the main text using separate models in each country (i.e. the estimates presented in Table 6). Figure 1 shows these results.
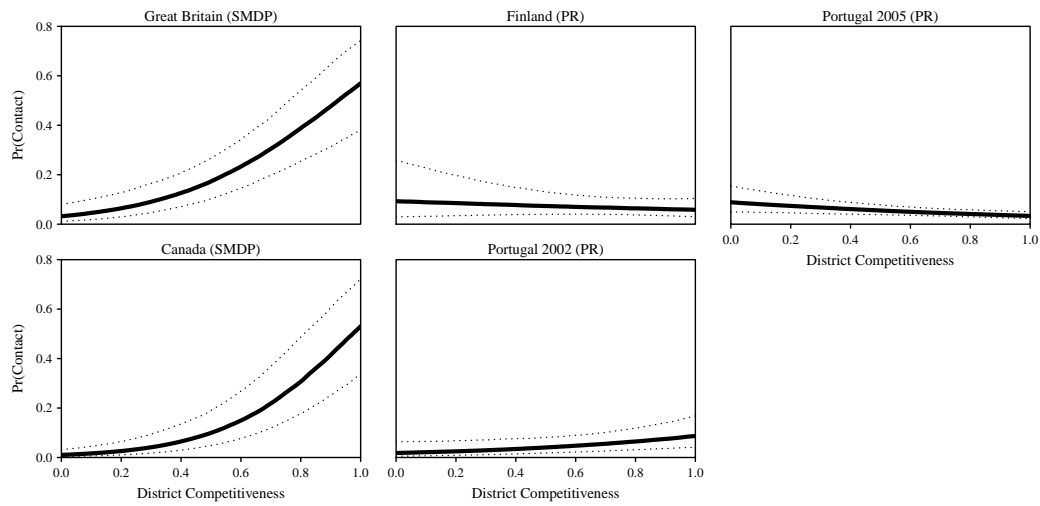
FIGURE 1: This figure replicates Figure **??** in the main text using separate logistic regression models in each election. That is, it exhibits no pooling across elections and complete pooling across districts. The results are very similar to Figure **??** in the main text and the substantive conclusions are the same.

D Variance Priors

The data I use to estimate the model parameters contain sufficient information to estimate all parameters reasonably well with the exception of the covariance matrix for the election-level effects. In this situation, the choice of prior can impact the results and one must choose a reasonable prior. A default, "uninformative" prior uniformly distributed from 0 to 100 yields results substantively similar to the results I discuss in the main text, but overestimates the variance parameters (**?**).

To alleviate concerns of overestimating variance parameters, I construct a weakly informative prior distribution for the diagonal entries of the covariance matrix for the election-level effects. The goal is to place a weakly-informative prior on the variance parameters (election-level variances, in this case) that contains much less prior information than is actually available, allowing the likelihood to drive the inferences. To do this, I use a half-Cauchy(5) distribution to model the standard deviations. Figure 2 shows that the half-Cauchy(5) is nearly uniform for reasonable values of the standard deviation, which for my purposes is roughly the interval [0,3]. Given the scale of the variables, it would be very strange if the true coefficients (conditional on electoral rules) across elections had a standard deviation of greater than three. One would probably expect a standard deviation of less than one. If the standard deviation were three, and the mean, for example, 2, then roughly half of the elections would have logit coefficients larger than four or less than zero. Given the scaling of the variables of interest, this is not a realistic scenario.

Another important property of the half-Cauchy(5) prior is that it goes to zero very slowly compared to the half-normal prior. So while it is roughly uniform from zero to three, it is quite flat out to ten. This property rules out very large standard deviations, such as 30, while allowing the likelihood to drive the estimates. Compare the half-Cauchy(5) prior to the uniform(0,100) and the half-normal(2). Notice that the uniform prior places an unrealistic, flat prior on the standard deviation, suggesting that large values near 50 are just as likely as small values near one. **?** notes that the uniform(0,100) prior is problematic because it leads to an overestimation of the variance parameters, especially when the number of groups is small. The half-normal(2) prior, on the other hand, acts as a highly informative prior, ruling out large standard deviations and placing much higher prior weight on parameter values less than three. Although I use the half-Cauchy(5) prior in the main analysis, the results are substantively similar for both the uniform and the half-normal prior.

To illustrate that the weakly informative half-Cauchy prior is not driving the inferences, Figure 3 shows the MCMC samples of the standard deviation parameter in the models using half-Cauchy(5), uniform(0,100), and half-normal(2) priors. In each case, plotting the prior is not useful because it appears as a barely visible flat line across the bottom of the histogram, suggesting the the likelihood is overwhelming the prior, as intended. However, notice that the posteriors based on the weakly informative half-Cauchy priors and the informative half-normal priors are more peaked and focused in a realistic region of the parameter space than the posteriors based on the uniform priors.
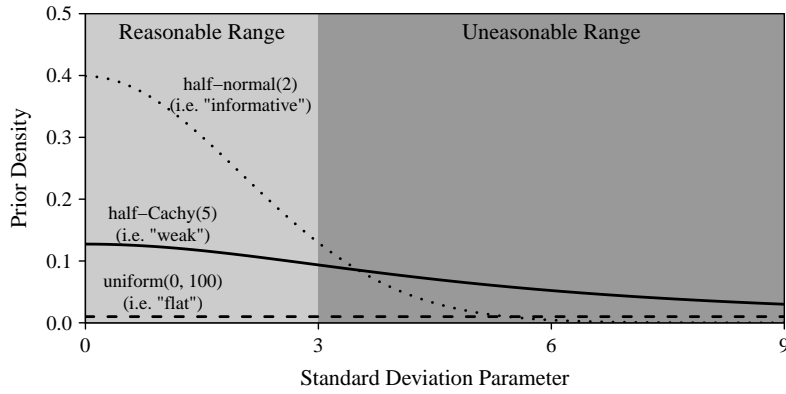
FIGURE 2: This figure shows that the half-Cauchy(5) prior is a reasonable, weakly informative prior. It is roughly flat for reasonable standard deviations and somewhat flat for larger, but unrealistic values. **?** points out that this prior structure will yield less upwardly biased estimates of standard deviation parameters. Compare the half-Cauchy(5) to the uniform(0,100), which places substantial prior weight on truly unrealistic standard deviations above ten, and the half-normal(2) prior, which places a much higher prior probability on parameter values less than three.
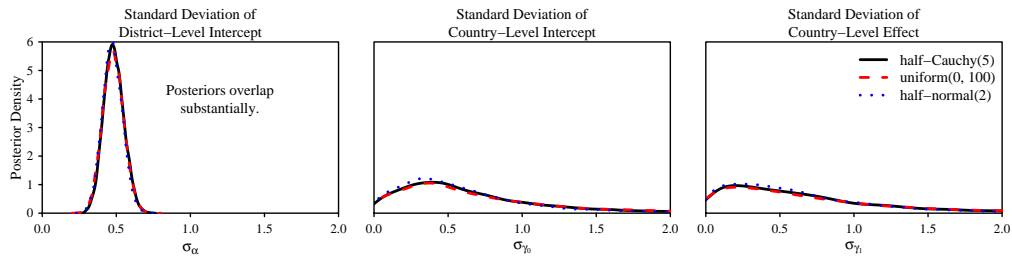

FIGURE 3: This figure shows that the half-Cauchy(5) prior does not overwhelm the likelihood. If the prior density were plotted, it would be a barely visible, flat line across the bottom of the plot. Further, all the posteriors are concentrated well below one, while both the half-Cauchy and uniform priors place a substantial prior probability above one.

15

E   Case Selection

Table 7 lists each country in Module 2 of the CSES and explains if and why I exclude it from the sample.

Country	Reason Excluded
Albania	second-tier adjustments
Australia	concurrent upper house election
Belgium	concurrent upper house election
Brazil	concurrent presidential election, concurrent upper house election
Bulgaria	overlapping single and multimember districts (though not strictly a second-tier correction)
Canada	Included
Chile	concurrent presidential election, concurrent upper house election
Czech Republic	second-tier adjustment
Denmark	second-tier adjustment
Finland	Included
France	concurrent presidential election
Germany	second-tier adjustment
Great Britain	Included
Hong Kong	overlapping geographic and functional constituencies
Hungary	second-tier adjustment
Iceland	second-tier adjustment
Ireland	Droop Quota
Israel	single national district
Italy	concurrent upper house election
Japan	concurrent upper house election
Korea	second-tier correction
Kyrgyzstan	concurrent presidential election
Mexico	second-tier adjustment
Netherlands	single national district
New Zealand	second-tier adjustment
Norway	second-tier adjustment
Peru	concurrent presidential election
Philippines	concurrent presidential election
Poland	concurrent upper house election
Portugal	Included
Romania	concurrent presidential election, concurrent upper house election, second-tier adjustment
Russia	concurrent presidential election
Slovenia	data unavailability
Spain	concurrent upper house elections
Sweden	second-tier adjustment
Switzerland	concurrent upper house elections
Taiwan	concurrent presidential election, second-tier adjustment
United States	concurrent presidential election, concurrent upper house election

TABLE 7: A table listing the countries in Module 2 of the CSES and explaining why each country was excluded.