

Online Appendix for  
*Going It Alone? A Structural Analysis of Coalition  
Formation in Elections*\*

Sergio Montero  
University of Rochester

## Contents

A Additional Figures and Tables	i
B Campaign-Stage Details	ix
C Estimation Details	xi
D Alternative Specifications	xviii
References	xxiii

---

\*This version: May 3, 2024.

## A Additional Figures and Tables

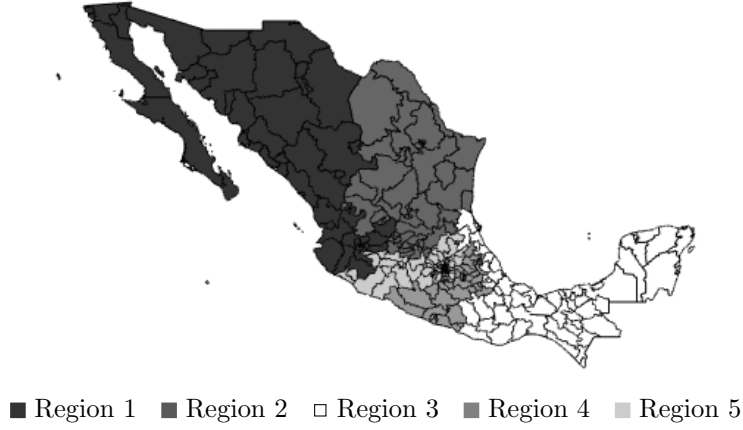


Figure A1: Mexican Electoral Regions and Districts (delimited)

Table A1: Composition of the Chamber of Deputies Following the 2012 Election

Party	FPTP Seats	PR Seats	Total
MP	71	64	135
NA		10	10
<b>PVEM</b>	19	15	34
<b>PRI</b>	158	49	207
PAN	52	62	114
Total	300	200	500

Table A2: Election Outcomes by PRI-PVEM Coalition Configuration

Party	Districts with Distinct PRI, PVEM Candidates		Districts with Joint PRI Candidate		Districts with Joint PVEM Candidate	
	Victory Rate (%)	Avg. Vote Share (%)	Victory Rate (%)	Avg. Vote Share (%)	Victory Rate (%)	Avg. Vote Share (%)
MP	22.8	25.5	21.2	29.4	34.9	31.4
NA	0	5.3	0	3.9	0	3.8
<b>PVEM</b>	3.0	4.9	-	7.0	37.2	7.7
<b>PRI</b>	51.5	36.7	67.9	33.2	-	28.7
PAN	22.8	27.6	10.9	26.4	27.9	28.4

Table A3: Votes in Support of PRI-PVEM Coalition Candidates

Type of PR Vote	Districts with Joint PRI Candidate	Districts with Joint PVEM Candidate
	Avg. Vote Share (%)	Avg. Vote Share (%)
PVEM	3.8	4.6
PRI	30.0	25.7
50-50 Split	6.4	6.1

*Notes.* The first two rows correspond to voters who gave 100% of their PR vote to the corresponding party. Thus, adding half of the third row to the other two yields the vote shares in Table A2.

Table A4: District Characteristics

	Districts with Distinct PRI, PVEM Candidates		Districts with Joint PRI Candidate		Districts with Joint PVEM Candidate	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Female Head of Household (%)	23.8	3.1	24.7	4.3	26.8	5.1
Pop. over 60 (% Voting-Age Pop.)	15.0	3.1	13.8	3.5	14.7	3.1
Rural Neighborhoods (%)	36.4	25.9	23.7	25.3	18.3	25.3

Table A5: Campaign Expenditures (Thousands of USD)

Party	Districts with Distinct PRI, PVEM Candidates		Districts with Joint PRI Candidate		Districts with Joint PVEM Candidate	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
MP	56.4	19.7	55.1	11.7	56.6	14.3
NA	19.7	8.5	16.7	4.4	19.1	8.5
<b>PVEM</b>	18.3	7.6	80.6	27.3	94.3	40.9
<b>PRI</b>	54.9	11.0				
PAN	38.0	10.4	41.4	12.7	44.6	14.2

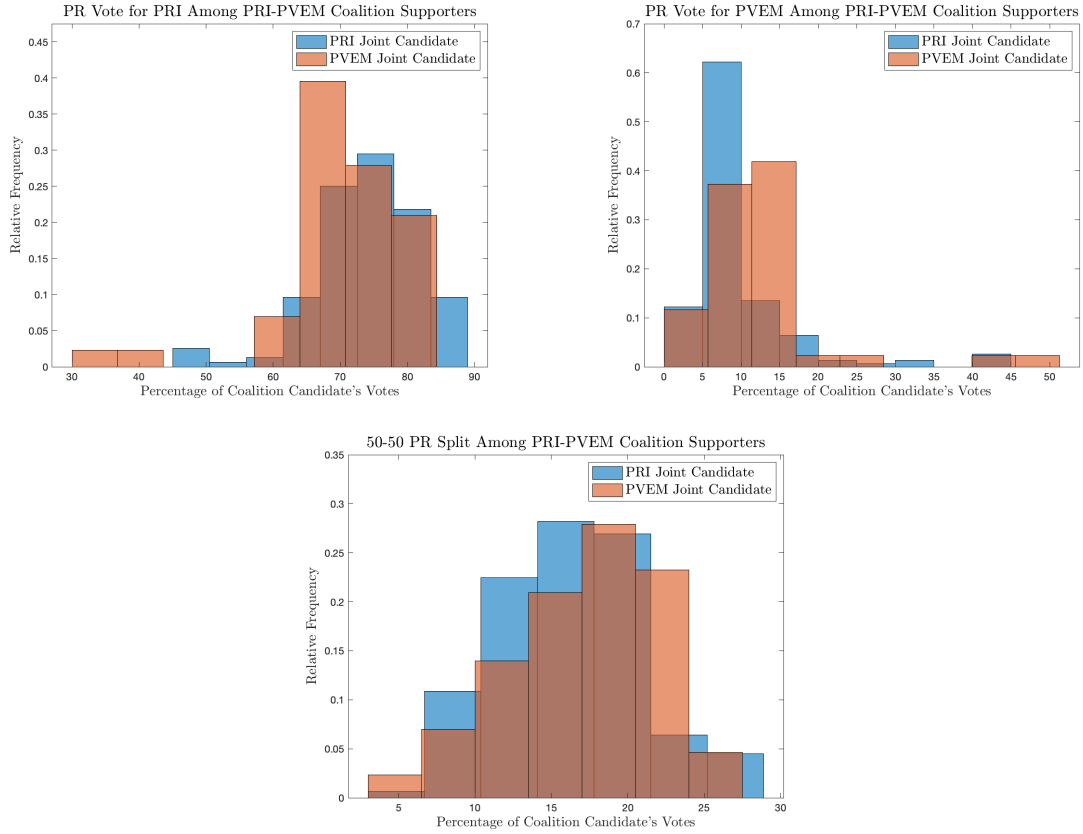


Figure A2: Composition of Votes in Support of PRI-PVEM Coalition Candidates

*Notes.* The top panels show the distribution across districts—by party affiliation of the coalition candidate—of the percentage of coalition supporters who gave their PR vote entirely to PRI (left) or PVEM (right). The bottom panel corresponds to a 50-50 split of the PR vote between the two partners.

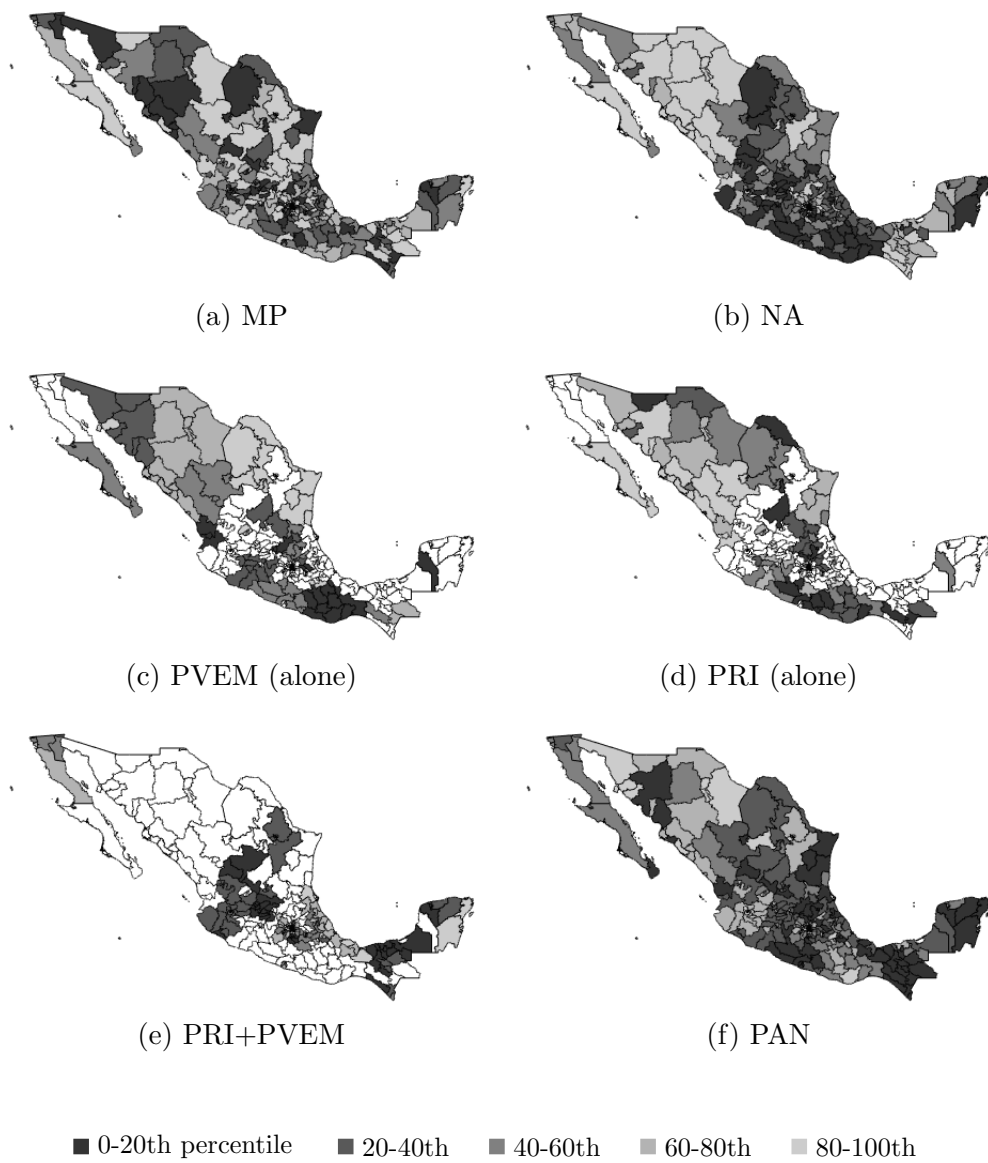


Figure A3: Geographic Distribution of Campaign Spending by Party

Table A6: Prior Electoral Experience of 2012 Chamber of Deputies Candidates

	Ran in 2009 (%)	Ran in 2006 (%)	Ran in 2003 (%)	Ran in 2003–2009 (%)
MP (FPTP)	7.7	8.0	4.7	17.7
MP (PR)	6.7	3.8	4.0	12.5
NA (FPTP)	3.7	4.7	1.0	9.3
NA (PR)	6.0	7.5	2.0	14.0
PVEM (FPTP Independent)	4.0	2.0	3.0	6.9
PVEM (FPTP Coalition)	0.0	9.3	2.3	11.6
PVEM (PR)	8.0	3.5	4.5	14.6
PRI (FPTP Independent)	2.0	5.9	5.0	12.9
PRI (FPTP Coalition)	3.2	4.5	8.3	14.7
PRI (PR)	4.5	9.0	2.5	14.5
PAN (FPTP)	2.3	7.0	3.7	10.7
PAN (PR)	4.0	9.5	9.0	17.5

*Notes.* This table summarizes, by party and election tier (i.e., first-past-the-post or proportional-representation), prior experience in federal legislative elections of 2012 Chamber of Deputies candidates. The first column reports the percentage of 2012 candidates who also ran (in any tier) in the 2009 Chamber of Deputies election. The second column corresponds to 2012 candidates who ran (in any tier) in 2006 for the Chamber of Deputies or the Senate. The third column corresponds to 2012 candidates who ran (in any tier) in the 2003 Chamber of Deputies election. The last column corresponds to 2012 candidates who participated in at least one federal legislative election between 2003 and 2009.

Table A7: Candidate-Choice Coefficient Estimates

	OLS (I)	BLP (II)	BLP (III)	OLS (IV)	BLP (V)	BLP (VI)
Spending	0.068 (0.022)	0.412 (0.240)	0.373 (0.223)	0.091 (0.020)	0.561 (0.363)	0.516 (0.340)
Spending Variance ( $\sigma_1$ )		0.000 (737.7)	0.000 (629.2)		0.000 (143.3)	0.000 (136.0)
Spending <sup>2</sup>	-0.002 (0.001)	-0.017 (0.019)	-0.016 (0.018)	-0.004 (0.001)	-0.025 (0.028)	-0.023 (0.026)
Spending <sup>2</sup> Variance ( $\sigma_2$ )		0.000 (265.4)	0.000 (244.5)		0.000 (23.16)	0.000 (28.26)
Spending in Neighboring Districts			0.026 (0.024)			0.016 (0.027)
Log-Lagged Vote Share	0.514 (0.023)	0.514 (0.031)	0.511 (0.030)	0.479 (0.029)	0.461 (0.038)	0.460 (0.036)
MP $\times$ Female	2.838 (0.756)	2.310 (1.211)	2.370 (1.169)	1.463 (0.669)	1.347 (1.283)	1.508 (1.222)
MP $\times$ Over 60	0.349 (0.895)	0.652 (1.265)	0.664 (1.211)	0.922 (0.733)	0.861 (1.269)	0.809 (1.188)
MP $\times$ Rural	-0.265 (0.137)	-0.417 (0.238)	-0.408 (0.225)	-0.534 (0.123)	-0.575 (0.243)	-0.555 (0.226)
NA $\times$ Female	-1.345 (0.955)	-1.630 (1.011)	-1.689 (1.003)	-1.924 (0.932)	-3.261 (1.133)	-3.353 (1.096)
NA $\times$ Over 60	0.665 (0.955)	0.374 (0.993)	0.465 (0.978)	0.855 (0.920)	0.914 (1.038)	1.001 (1.005)
NA $\times$ Rural	-0.087 (0.147)	0.104 (0.152)	0.098 (0.149)	-0.182 (0.143)	-0.094 (0.152)	-0.099 (0.148)
PVEM $\times$ Female	-1.395 (1.522)	-0.857 (1.788)	-1.187 (1.806)	-2.100 (1.459)	-0.290 (1.678)	-0.665 (1.736)
PVEM $\times$ Over 60	-0.563 (1.273)	-0.220 (1.522)	-0.177 (1.499)	-0.527 (1.218)	0.097 (1.578)	0.095 (1.509)
PVEM $\times$ Rural	0.675 (0.158)	1.043 (0.267)	0.997 (0.255)	0.462 (0.156)	0.818 (0.268)	0.775 (0.267)
PRI $\times$ Female	-1.735 (0.656)	-1.690 (1.065)	-1.506 (1.055)	-2.028 (0.667)	-1.757 (1.074)	-1.595 (1.091)
PRI $\times$ Over 60	0.832 (0.618)	0.661 (1.066)	0.602 (1.018)	1.016 (0.630)	1.385 (1.122)	1.266 (1.099)
PRI $\times$ Rural	0.290 (0.082)	0.424 (0.132)	0.450 (0.128)	0.228 (0.090)	0.380 (0.151)	0.409 (0.144)
PAN $\times$ Female	-1.614 (0.790)	-1.110 (1.127)	-1.164 (1.101)	-1.460 (0.790)	-1.311 (1.137)	-1.393 (1.098)
PAN $\times$ Over 60	3.020 (0.796)	1.202 (1.205)	1.263 (1.125)	3.113 (0.819)	1.555 (1.248)	1.738 (1.177)
PAN $\times$ Rural	-0.241 (0.119)	0.404 (0.373)	0.385 (0.321)	-0.258 (0.126)	0.426 (0.375)	0.379 (0.336)
Menu-Party F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Region F.E.	No	No	No	Yes	Yes	Yes
Observations	1,301	1,301	1,301	1,301	1,301	1,301

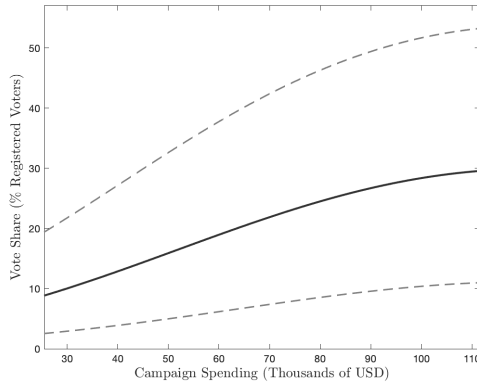
*Notes.* OLS and BLP estimates of coefficients driving candidate choice, with robust standard errors in parentheses. The first six rows correspond to  $(\alpha, \sigma)$ , which determine the effectiveness of campaign expenditures according to Equation (1). The remaining rows correspond to  $\beta$ , which characterizes baseline partisanship in Equation (1). Columns (III) and (VI) test for the presence of campaign spillovers across neighboring districts.

Table A8: Proportional-Representation Party-Choice Coefficient Estimates

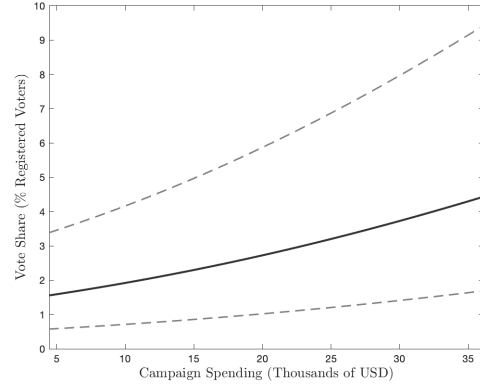
	(I)	(II)
Log-Lagged Vote Share	0.463 (0.055)	0.469 (0.067)
PVEM $\times$ Female	-1.348 (1.283)	-1.141 (1.468)
PVEM $\times$ Over 60	0.882 (1.311)	0.888 (1.242)
PVEM $\times$ Rural	0.518 (0.248)	0.320 (0.206)
PRI $\times$ Female	-2.108 (0.760)	-1.239 (0.836)
PRI $\times$ Over 60	3.208 (0.751)	2.703 (0.689)
PRI $\times$ Rural	0.147 (0.118)	0.091 (0.103)
Menu-Party F.E.	Yes	Yes
Region F.E.	No	Yes
Observations	398	398

*Notes.* OLS estimates of  $\beta^{\text{ST}}$ , which drives second-tier choice for PRI-PVEM coalition supporters of how to allocate their PR vote according to Equation (3), with robust standard errors in parentheses. Outside option is 50-50 vote split between the two partners.

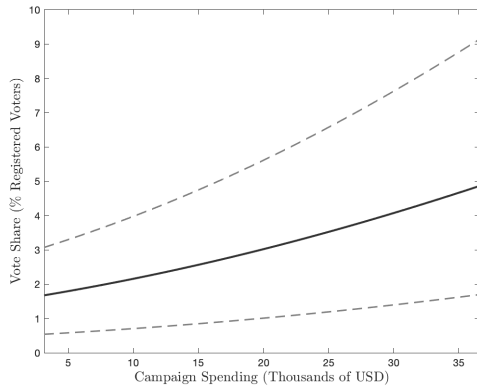




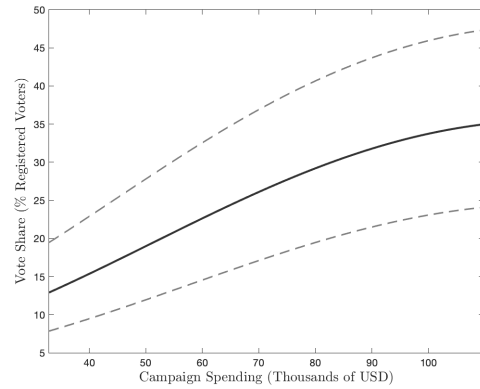
(a) MP



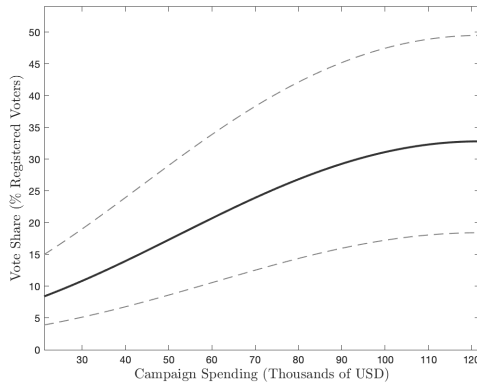
(b) NA



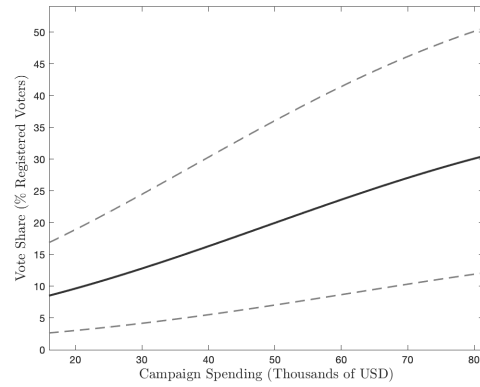
(c) PVEM (alone)



(d) PRI (alone)



(e) PRI+PVEM



(f) PAN

Figure A4: Electoral Impact of Campaign Expenditures

*Notes.* The horizontal axis of each panel is centered at the party's observed average spending and ranges by plus/minus two standard deviations. The vertical axes measure FPTP district vote shares as a percentage of registered voters. Solid lines plot averages, and dashed lines delimit 90% confidence intervals, taking into account the empirical distribution of district characteristics and observed spending by competing parties.

## B Campaign-Stage Details

**Equivalence between alternative formulations of campaign stage.** Since there is no evidence in Table A7 of heterogeneous voter responsiveness to campaign efforts, candidate  $j$ 's vote share in Equation (4), given menu  $M_d$ , can be written simply as

$$s_{jd}^{M_d}(c_{jd}, c_{-j,d}) = \frac{\exp(\alpha_1 c_{jd} + \alpha_2 c_{jd}^2 + (x_{jd}^{M_d})' \beta + \xi_{jd}^{M_d})}{1 + \sum_{j' \in M_d} \exp(\alpha_1 c_{j'd} + \alpha_2 c_{j'd}^2 + (x_{j'd}^m)' \beta + \xi_{j'd}^{M_d})}, \quad (\text{B1})$$

where  $c_{-j,d}$  denotes the profile of spending in district  $d$  by  $j$ 's rivals. Party or coalition  $j$ 's payoff in district  $d$ —up to a constant in  $c_{jd}$ —is given by

$$\pi_{jd}^{M_d}(c_{jd}, c_{-j,d}) = \tilde{\gamma}_j^{M_d} \log(s_{jd}^{M_d}(c_{jd}, c_{-j,d})) - c_{jd},$$

with

$$\tilde{\gamma}_j^{M_d} = \begin{cases} \gamma_{\text{PRI}} + \gamma_{\text{PVEM}} & \text{if } M_d \in \{M^{\text{PRI}}, M^{\text{PVEM}}\} \text{ and } j \in \{\text{PRI}, \text{PVEM}\}, \\ \gamma_j & \text{otherwise.} \end{cases}$$

As discussed in the paper, I assume parties face a flexible national budget constraint, which implies that they effectively play independent complete-information campaign spending games across districts. A (pure-strategy) Nash equilibrium in district  $d$  is a profile of spending,  $c_d^*$ , such that  $c_{jd}^* \in \arg \max_{c_{jd} \in [0, \infty)} \pi_{jd}^{M_d}(c_{jd}, c_{-j,d}^*)$  for all  $j \in M_d$ . In equilibrium, assuming positive spending by all parties as observed in the data,  $j$ 's spending satisfies the first-order condition

$$\frac{\partial \pi_{jd}^{M_d}(c_{jd}^*, c_{-j,d}^*)}{\partial c_{jd}} = \tilde{\gamma}_j^{M_d} [1 - s_{jd}^{M_d}(c_{jd}^*, c_{-j,d}^*)] (\alpha_1 + 2\alpha_2 c_{jd}^*) - 1 = 0. \quad (\text{B2})$$

The term  $\tilde{\gamma}_j^{M_d} [1 - s_{jd}^{M_d}(c_{jd}^*, c_{-j,d}^*)] (\alpha_1 + 2\alpha_2 c_{jd}^*)$  represents the marginal value for  $j$  of an additional dollar of spending, which is thus equalized across districts.

With a hard national budget constraint, consider an alternative formulation of  $j$ 's problem wherein it seeks to maximize its aggregate electoral payoff,  $\sum_d \tilde{\gamma}_j^{M_d} \log(s_{jd}^{M_d}(c_{jd}, c_{-j,d}^*))$ , subject to  $\sum_d c_{jd} \leq \bar{c}_j$ , where  $\bar{c}_j$  denotes  $j$ 's budget. Equation (B2) in this case would be replaced with

the corresponding first-order condition for the Lagrangian,  $L_j$ , of  $j$ 's constrained optimization problem:

$$\frac{\partial L_j(c_j^*, c_{-j}^*)}{\partial c_{jd}} = \tilde{\gamma}_j^{M_d} [1 - s_{jd}^{M_d}(c_{jd}^*, c_{-j,d}^*)] (\alpha_1 + 2\alpha_2 c_{jd}^*) - \lambda_j = 0. \quad (\text{B3})$$

Assuming a binding budget constraint,  $j$ 's Lagrange multiplier  $\lambda_j > 0$ . Dividing Equation (B3) by  $\lambda_j$  then yields

$$\frac{\tilde{\gamma}_j^{M_d}}{\lambda_j} [1 - s_{jd}^{M_d}(c_{jd}^*, c_{-j,d}^*)] (\alpha_1 + 2\alpha_2 c_{jd}^*) - 1 = 0,$$

which is identical to Equation (B2) up to a renormalization of  $j$ 's payoff. The two versions of the campaign stage are in this sense observationally equivalent. For computational convenience, I adopt the independent-games version of the model, but the estimates of parties' campaign-stage payoffs in Table 1 can be interpreted as capturing all relevant opportunity costs of campaign expenditures.

**Games with strategic complementarities.** While I refer the reader to Echenique and Edlin (2004) for a formal definition of games with strict strategic complementarities (GSSC), I discuss here properties of the parties' payoff functions, satisfied at the estimated parameter values in Tables A7 and 1, which imply that the district spending games belong to this class. First, since  $\alpha_1 > 0 > \alpha_2$ , the effect of  $c_{jd}$  on candidate  $j$ 's vote share in Equation (B1) is maximized at  $\bar{c} = -\alpha_1/(2\alpha_2)$ . It then follows that spending more than  $\bar{c}$  is a strictly dominated strategy for all players in the spending games. Thus, the effective strategy space for each party is  $[0, \bar{c}]$ , a compact interval, which satisfies condition 1 of the definition of GSSC in Echenique and Edlin (2004). Second, given any  $(\tilde{c}_{jd}, \tilde{c}_{-j,d}) \in [0, \bar{c}]^{M_d}$  and  $j' \neq j$ ,

$$\frac{\partial^2 \pi_{jd}^{M_d}(\tilde{c}_{jd}, \tilde{c}_{-j,d})}{\partial c_{j'd} \partial c_{jd}} = \tilde{\gamma}_j^{M_d} s_{jd}^{M_d}(\tilde{c}_{jd}, \tilde{c}_{-j,d}) s_{j'd}^{M_d}(\tilde{c}_{jd}, \tilde{c}_{-j,d}) (\alpha_1 + 2\alpha_2 \tilde{c}_{jd}) (\alpha_1 + 2\alpha_2 \tilde{c}_{j'd}) > 0.$$

That is,  $j$ 's incentive to raise its spending is strictly increasing in its rivals' spending. This implies the remaining conditions of the definition of GSSC.

As noted in the paper, GSSC have three useful properties. First, existence of equilibrium is guaranteed (Vives, 1990). Second, mixed-strategy equilibria are unstable, so their omission is justified (Echenique and Edlin, 2004). Lastly, Echenique (2007) provides a simple and fast algorithm for computing the set of all pure-strategy equilibria. This set has an additional key property. It has a largest and a smallest equilibrium, providing a simple test of uniqueness: if the largest and smallest equilibria coincide, the resulting strategy profile is the unique equilibrium of the game. These extremal equilibria can be easily computed through best-response iteration. The smallest (largest) equilibrium is obtained by iterating best responses until convergence starting from the strategy profile with  $\underline{c}_{jd} = 0$  ( $\bar{c}_{jd} = \bar{c}$ ) for all  $j \in M_d$ . At the estimated parameter values, the largest and smallest equilibria of the campaign spending games always coincide.

## C Estimation Details

As summarized in the paper, the estimation strategy mirrors the model’s three-stage structure. Step 1 recovers the voting-stage parameters in Equations (1) and (3). Step 2 obtains payoff coefficient  $\gamma_p$  for each party  $p$  by matching the spending levels observed in the data with the model’s predictions from the campaign stage. Finally, ex-ante coalition surplus maximization is exploited in Step 3 to recover  $\theta$ , which characterizes the partners’ (dis)utility from not fielding a candidate.

**Step 1.** With heterogenous voter impressionability ( $\sigma \neq 0$ ), the simple linear regression estimator of voters’ preferences described in the paper is no longer feasible. However, Berry, Levinsohn and Pakes (BLP, 1995) show that predicted vote shares can still be implicitly “inverted” in this case after matching observed vote shares exactly. That is, given Equation (4) and any value of  $\sigma$ , there exists a unique vector of mean utilities,  $\delta_d^{M_d}(\sigma) = (\delta_{jd}^{M_d}(\sigma))_{j \in M_d}$ , such that  $\hat{s}_{jd}^{M_d} = s_{jd}^{M_d}(\delta_d^{M_d}(\sigma), \sigma)$  for all  $j \in M_d$ .<sup>1</sup> Unobserved candidate valence consistent

---

<sup>1</sup>To compute predicted vote shares given Equation (4), I use sparse-grid integration as implemented by Heiss and Winschel (2008).

with  $\delta_{jd}^{M_d}(\sigma)$  can then be computed using Equation (2), for any trial value of  $\varphi = (\alpha, \beta, \sigma)$ , as

$$\xi_{jd}^{M_d}(\varphi) = \delta_{jd}^{M_d}(\sigma) - \alpha_1 c_{jd} - \alpha_2 c_{jd}^2 - (x_{jd}^{M_d})' \beta.$$

Given a vector  $z_{jd}$  of valid instruments—i.e.,

$$E[z_{jd} \xi_{jd}^{M_d}(\varphi)] = 0 \quad \text{if and only if} \quad \varphi = \varphi_0, \quad (\text{C1})$$

where  $\varphi_0$  denotes the true value of the parameters—a Generalized Method of Moments (GMM) estimator can be obtained by minimizing the quadratic form  $Q_N(\varphi) = [\frac{1}{N} Z' \xi(\varphi)]' W_N [\frac{1}{N} Z' \xi(\varphi)]$ . Here,  $Z$  and  $\xi(\varphi)$  are vertical stackings of  $z'_{jd}$  and  $\xi_{jd}^{M_d}(\varphi)$  across candidates and districts,  $N$  denotes the total number of observations, and  $\frac{1}{N} Z' \xi(\varphi)$  is the sample analog of moment condition (C1).

Under standard regularity conditions (Hansen, 1982; Berry, Levinsohn and Pakes, 1995), this GMM estimator,  $\hat{\varphi}$ , satisfies

$$\sqrt{N}(\hat{\varphi} - \varphi_0) \xrightarrow{d} \mathcal{N}(0, (G'WG)^{-1}G'W\Omega W'G(G'W'G)^{-1})$$

as the sample size  $N \rightarrow \infty$ , where

$$G = E[z_{jd} \nabla_{\varphi} \xi_{jd}^{M_d}(\varphi_0)] \quad \text{and} \quad \Omega = E[z_{jd} \xi_{jd}^{M_d}(\varphi_0) \xi_{jd}^{M_d}(\varphi_0)' z'_{jd}]$$

are the gradient and variance, respectively, of the moment conditions defined by Equation (C1), and  $W_N \xrightarrow{p} W$ . Notice that the optimal weighting matrix  $W^* = \Omega^{-1}$  minimizes the asymptotic variance of the estimator, which then simplifies to  $(G'\Omega^{-1}G)^{-1}$ . This suggests a two-step estimation approach, which I follow. In a first step, a consistent but inefficient estimate  $\hat{\varphi}_I$  can be obtained by minimizing  $Q_N(\varphi)$  using any positive-definite weighting matrix.<sup>2</sup> Then, allowing for arbitrary heteroskedasticity, the optimal weighting matrix can be consis-

---

<sup>2</sup>I employ an approximation of  $\Omega^{-1}$  using residuals from the homogeneous version of the model with  $\sigma = 0$ .

tently estimated as  $\hat{W}^* = \hat{\Omega}^{-1} = \left(\frac{1}{N} Z' V_{\xi}(\hat{\varphi}_I)' Z\right)^{-1}$ , where  $(V_{\xi}(\hat{\varphi}_I))_{jj'} = \xi_j(\hat{\varphi}_I) \xi_{j'}(\hat{\varphi}_I) \mathbf{1}_{j=j'}$ . In a second step, reestimating the model using  $\hat{W}^*$  delivers a consistent and efficient estimate  $\hat{\varphi}$ . For robust inference, again allowing for arbitrary heteroskedasticity, a consistent estimate of the asymptotic variance of  $\hat{\varphi}$  can be obtained simply as  $(\hat{G}' \hat{\Omega}^{-1} \hat{G})^{-1}$ , where  $\hat{G} = Z' \nabla_{\varphi} \xi(\hat{\varphi})$  and  $\hat{\Omega} = Z' V_{\xi}(\hat{\varphi}) Z$ .<sup>3</sup>

BLP propose an estimation algorithm that proceeds by iterating over two nested loops. This algorithm, however, can be computationally inefficient and sensitive to convergence criteria. Instead, I follow the Mathematical Programming with Equilibrium Constraints (MPEC) approach of Dubé, Fox and Su (2012). The key idea is to impose the “equilibrium conditions” of the model,  $\hat{s}_{jd}^{M_d} = s_{jd}^{M_d}(\delta_d^{M_d}(\sigma), \sigma)$ , as explicit constraints on the GMM program, relying on recent advances in constrained optimization algorithms for improved numerical performance. Specifically, I compute  $\hat{\varphi}$  by solving the following mathematical program with equilibrium constraints:

$$\begin{aligned} \min_{\varphi, \xi, \psi} \psi' W \psi \quad \text{subject to} \\ \psi = Z' \xi \quad \text{and} \end{aligned} \tag{C2}$$

$$s_{jd}^{M_d}(\delta_d^{M_d}, \sigma) = \hat{s}_{jd}^{M_d} \quad \text{for all } j, d, \text{ where} \tag{C3}$$

$$\delta_{jd}^{M_d} = \alpha_1 c_{jd} + \alpha_2 c_{jd}^2 + (x_{jd}^{M_d})' \beta + \xi_{jd}^{M_d}. \tag{C4}$$

Dubé, Fox and Su (2012) show that this MPEC and the traditional BLP algorithm yield theoretically identical estimates, but the MPEC approach delivers superior numerical performance. While the computational cost of estimation may seem to increase by treating  $\xi$  and the moment conditions,  $\psi$ , as auxiliary variables—and thus expanding the size of the optimization problem—note that (C2) and (C4) are linear constraints, and  $(\varphi, \xi)$  no longer enter the objective function directly. This, together with the sparsity that results from  $\xi_{jd}^{M_d}$  having no effect on vote shares outside of  $j$ ’s district, adds to the computational advantage over the

---

<sup>3</sup>This can also easily accommodate clustering by district, letting  $(V_{\xi}(\hat{\varphi}))_{jj'} = \xi_j(\hat{\varphi}) \xi_{j'}(\hat{\varphi})$  if  $j$  and  $j'$  compete in the same district, and  $(V_{\xi}(\hat{\varphi}))_{jj'} = 0$  otherwise. Results are nearly identical.

traditional BLP approach.<sup>4</sup>

A necessary order condition for the instrument vector,  $z_{jd}$ , is that it must include at least as many variables as there are parameters to be estimated. The choice of instruments to identify  $(\alpha, \beta)$  follows standard intuition from linear models: the exogenous covariates in  $x_{jd}^{M_d}$  constitute valid—in fact, optimal—instruments to identify  $\beta$ , and the lagged-spending instruments, as described in the paper, identify  $\alpha$ . On the other hand, the impressionability variance parameters,  $\sigma$ , determine nonlinear features of the model and are, in many applications, hard to estimate precisely (Gordon and Hartmann, 2013; Gillen et al., 2019; Gandhi and Houde, 2023). Part of the difficulty stems from finding the right source of variation to pin down the effects of these parameters on model predictions. The standard approach has been to heuristically construct nonlinear transformations of other available instruments in an attempt to match the nonlinear features of the model. Recent work by Gandhi and Houde (2023) shows that this approach, while well-intended, can produce very weak instruments if the transformations don’t involve the right ingredients. In particular, the coefficients in  $\sigma$  shape patterns of substitutability across candidates, relaxing the Independence of Irrelevant Alternatives property that is otherwise imposed on the homogeneous-voters version of the model by the TIEV distribution. Since substitutability is determined, empirically, by how close alternatives are in terms of their relevant attributes, Gandhi and Houde argue that a flexible function of attribute differences across candidates provides the right source of variation to identify  $\sigma$ . Accordingly, I use a second-degree polynomial of observed differences across candidates in  $x_{jd}^{M_d}$  and the (two-stage least squares) fitted value of  $c_{jd}$  (using the lagged-spending instruments).

**Step 2.** The GMM estimator of the campaign-stage parameters is analogous to that in Step 1, with  $\hat{c}_{jd} - c_{jd}(\gamma, \hat{\varphi})$  playing the role of  $\xi_{jd}^{M_d}(\varphi)$  above. The only difference is that inference in this case must account for estimation uncertainty in  $\hat{\varphi}$ . I rely on standard results for

---

<sup>4</sup>Realizing these gains, however, requires state-of-the-art optimization software, capable of handling large problems with nonlinear constraints. I rely on the industry-leading Knitro’s (<https://www.artelys.com/en/optimization-tools/knitro>) Interior-Point/Direct algorithm, to which I provide exact first and second derivatives of the objective and constraints.

two-step GMM estimation (Newey and McFadden, 1994). Specifically, a consistent estimate of the joint asymptotic variance of  $(\hat{\varphi}, \hat{\gamma})$  is given by  $(\hat{G}'\hat{\Omega}^{-1}\hat{G})^{-1}$ , as above, where  $\hat{G}$  and  $\hat{\Omega}$  correspond in this case to estimates of the gradient and variance, respectively, of the joint moment restrictions

$$E \left[ \begin{pmatrix} z_{jd}(\hat{c}_{jd} - c_{jd}(\gamma_0, \varphi_0)) \\ z_{jd}\xi_{jd}^{M_d}(\varphi_0) \end{pmatrix} \right] = 0.$$

**Step 3.** Lastly, as described in the paper,  $\theta$  can be estimated by maximizing the log-likelihood

$$\sum_d \log (\mathcal{L}_d(M_d; \theta, \hat{\gamma}, \hat{\varphi})).$$

Again, standard errors must be adjusted to account for estimation uncertainty in  $(\hat{\gamma}, \hat{\varphi})$ . I rely once more on two-step GMM inference noting that Maximum Likelihood estimation here is equivalent to GMM estimation based on the moment (first-order) conditions

$$E[\nabla_{\theta} \log (\mathcal{L}_d(M_d; \theta, \gamma_0, \varphi_0))] = 0 \quad \text{if and only if} \quad \theta = \theta_0,$$

where  $\theta_0$  denotes the true value of the parameters.

The coalition formation stage of the model can be alternatively formulated without introducing idiosyncratic bargaining shocks. In this case, only an average (dis)utility of not fielding a candidate can be identified for each party, with Equation (7) simplifying to

$$\bar{\pi}_{pd}^m(\theta, \gamma, \varphi) = \theta_p \mathbf{1}_{j \neq p} + E[\pi_{pd}^m(\gamma, \varphi)].$$

Analogous to nonnegative-profit market entry conditions, joint surplus maximization by PRI and PVEM implies the following moment inequalities:

$$\bar{\pi}_{\text{PRI},d}^{M_d}(\theta, \gamma, \varphi) + \bar{\pi}_{\text{PVEM},d}^{M_d}(\theta, \gamma, \varphi) \geq \bar{\pi}_{\text{PRI},d}^m(\theta, \gamma, \varphi) + \bar{\pi}_{\text{PVEM},d}^m(\theta, \gamma, \varphi) \quad (\text{C5})$$

for all  $m \in \{M^{\text{PRI}}, M^{\text{PVEM}}, M^{\text{IND}}\}$ .



Shi and Shum (2015) propose a simple inference procedure for models with such a structure—i.e., models where a subset of parameters ( $\gamma$  and  $\varphi$ ) are point identified and estimated in a preliminary stage (Steps 1 and 2), and the remaining parameters are related to the point-identified parameters via inequality/equality restrictions. To implement their procedure, which requires both equalities and inequalities, I introduce slackness parameters: for each  $m$ , condition (C5) becomes an equality restriction,

$$\bar{\pi}_{\text{PRI},d}^{M_d}(\theta, \gamma, \varphi) + \bar{\pi}_{\text{PVEM},d}^{M_d}(\theta, \gamma, \varphi) - [\bar{\pi}_{\text{PRI},d}^m(\theta, \gamma, \varphi) + \bar{\pi}_{\text{PVEM},d}^m(\theta, \gamma, \varphi)] + \kappa_m = 0,$$

and the slackness parameters must satisfy  $\kappa_m \geq 0$ . A criterion function is constructed as follows. With a slight abuse of notation, let  $\beta$  be a vector collecting the output of Steps 1 and 2, and let  $\theta = (\theta_{\text{PRI}}, \theta_{\text{PVEM}}, \kappa_{\text{IND}}, \kappa_{\text{PRI}}, \kappa_{\text{PVEM}})$ . Define  $g^e(\theta, \beta) = (g_m^e(\theta, \beta))_{m \in \{M^{\text{PRI}}, M^{\text{PVEM}}, M^{\text{IND}}\}}$  by

$$g_m^e(\theta, \beta) = \bar{\pi}_{\text{PRI},d}^{M_d}(\theta, \gamma, \varphi) + \bar{\pi}_{\text{PVEM},d}^{M_d}(\theta, \gamma, \varphi) - [\bar{\pi}_{\text{PRI},d}^m(\theta, \gamma, \varphi) + \bar{\pi}_{\text{PVEM},d}^m(\theta, \gamma, \varphi)] + \kappa_m,$$

and let  $g^{ie}(\theta) = (g_m^{ie}(\theta))_{m \in \{M^{\text{PRI}}, M^{\text{PVEM}}, M^{\text{IND}}\}} = (\kappa_m)_{m \in \{M^{\text{PRI}}, M^{\text{PVEM}}, M^{\text{IND}}\}}$ . Thus,  $g^e$  summarizes the equality restrictions involving all parameters of the model, and  $g^{ie}$  summarizes the inequality restrictions involving only  $\theta$ . Letting  $\beta_0$  denote the true value of  $\beta$ , the identified set of  $\theta$ —i.e., the set of parameter values consistent with (or not rejected by) the data—is

$$\Theta_0 = \{\theta : g^e(\theta, \beta_0) = 0 \text{ and } g^{ie}(\theta) \geq 0\}.$$

Given  $Q(\theta, \beta; W) = g^e(\theta, \beta)'Wg^e(\theta, \beta)$ , where  $W$  is a positive definite matrix, it follows that  $\Theta_0 = \arg \min_{\theta} Q(\theta, \beta_0; W)$  subject to  $g^{ie}(\theta) \geq 0$ .

Shi and Shum show that the following is a confidence set of level  $\alpha \in (0, 1)$  for  $\Theta_0$ :

$$CS = \{\theta : g^{ie}(\theta) \geq 0 \text{ and } Q(\theta, \hat{\beta}, \hat{W}) \leq \chi_{(3)}^2(\alpha)/N\},$$

where  $\chi^2_{(3)}(\alpha)$  is the  $\alpha$ -th quantile of the  $\chi^2$  distribution with 3 degrees of freedom (the number of restrictions in  $g^e$ ),  $\hat{\beta}$  a consistent estimator of  $\beta_0$  (obtained from Steps 1 and 2),  $N$  is the number of observations used to estimate  $\hat{\beta}$ , and

$$\hat{W} = \left[ G(\theta, \hat{\beta}) \hat{V}_{\beta} G(\theta, \hat{\beta})' \right]^{-1},$$

with  $G(\theta, \hat{\beta}) = \nabla_{\beta} g^e(\theta, \hat{\beta})$  and  $\hat{V}_{\beta}$  a consistent estimate of the asymptotic variance of  $\hat{\beta}$ . Figure C1 shows the projection of this confidence set, focusing on  $(\theta_{\text{PRI}}, \theta_{\text{PVEM}})$ . As  $g^e(\theta, \beta)$  and  $g^{ie}(\theta)$  are in fact linear in  $\theta$ ,  $Q(\theta, \hat{\beta}; \hat{W})$  has a unique minimizer subject to  $g^{ie}(\theta) \geq 0$ , which provides a useful “point estimate,” highlighted in Figure C1. As discussed by Shi and Shum, the slackness parameters,  $\kappa_m$ , are nuisance parameters, which may lead to conservative confidence sets for the parameters of interest. This does not seem to be a problem in this application, however, given that the depicted confidence set is fairly tight. Furthermore, the identified values of  $\theta_p$  broadly agree with the mean of  $w'_{pd}\theta$  from the version of the coalition formation stage in the paper, although the former naturally miss considerable heterogeneity.

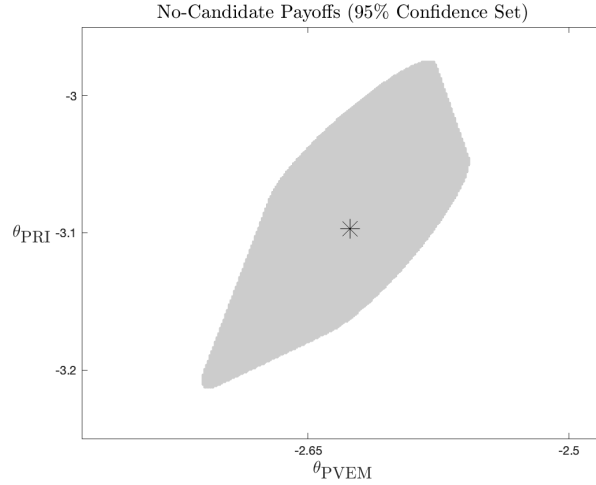


Figure C1: Confidence Set for Parameters from Alternative Formulation of Coalition Formation Stage with No Idiosyncratic Bargaining Shocks

## D Alternative Specifications

Tables D1 (voting-stage candidate choice), D2 (voting-stage party choice), and D3 (coalition stage) present coefficient estimates from alternative specifications aimed at addressing several potential concerns. For easy reference, all tables reproduce the baseline estimates in the paper or Online Appendix A.

**Governors.** As discussed in the paper, related research has found that Mexican governors can be very influential with regard to federal legislators. Columns (II) and (VII) of Table D1, and columns (II) and (VI) of Tables D2 and D3, report coefficient estimates from alternative model specifications that allow for potential effects of incumbent governors on voting behavior and coalition formation incentives. Although Tables D1 and D2 suggest governors may have some impact on same-party candidates' vote shares and on voters' PR party choice, these results are not robust to controlling for electoral region fixed effects. Moreover, Table D3 indicates governors have no substantively or statistically significant influence on coalition formation considerations.

**Measurement error in campaign expenditures.** In 2012, the Chamber of Deputies election took place concurrently with the Senate and presidential contests. The victorious PRI-PVEM presidential candidate was accused of using Chamber of Deputies campaign expenditures as a way of skirting presidential campaign spending limits.<sup>5</sup> This raises serious concerns about the reliability of reported spending in each FPTP district as a measure of campaign efforts in direct support of the corresponding candidate for the Chamber of Deputies. To address this, I conduct two related analyses. Since 2009 was a mid-term election year, cross-election contamination concerns surrounding campaign expenditures do not apply. To identify the districts where cross-election contamination is most likely to have occurred in 2012, I calculate the percentage increase from 2009 to 2012 in joint PRI-PVEM spending for

---

<sup>5</sup><https://www.animalpolitico.com/analisis/invitados/para-entender-el-prorratio-de-los-gastos-de-campana> (in Spanish).

each district. I then drop from the sample all districts in the top 5%. Columns (III) and (VIII) of Table D1, and columns (III) and (VII) of Tables D2 and D3, report coefficient estimates using this restricted sample. Similarly, columns (IV) and (IX) of Table D1, and columns (IV) and (VIII) of Tables D2 and D3, report coefficient estimates after dropping districts in the top 10%. Throughout, results are virtually identical to their baseline counterparts.

**Campaign spending instruments.** Finally, to address concerns about the validity of 2009 spending as an instrument for expenditures in 2012, columns (V) and (X) of Table D1 report coefficient estimates using an alternative set of instruments. First, since Table A7 rules out meaningful spillovers across districts in campaign efforts, I use 2009 spending in neighboring districts rather than in the district itself to instrument for spending in 2012. This should alleviate concerns about any unobservables affecting 2009 spending and 2012 election outcomes in a district not already captured by lagged vote shares. Second, 2012 was the first electoral cycle in Mexico in which the Internet seemed to play an important role because it enabled direct communication between candidates and voters, lowering campaign costs (Díaz Cayeros et al., 2012). Assuming parties anticipated this and tailored campaign expenditures accordingly, the share of households in a district with Internet access (available from the 2010 census) should provide another valid instrument for 2012 spending after controlling for other observed district characteristics. I similarly use a measure, computed by the electoral authority, of the average travel time it takes to visit all election precincts in a district. Reassuringly, point estimates in columns (V) and (X) of Table D1 are consistent with their baseline counterparts. Estimates of  $\alpha_1$  (first row) and  $\alpha_2$  (third row) are less precise, however, which is unsurprising given that the alternative instruments are weaker as they are less directly related to variation in spending *by each party* in each particular district.

Table D1: Candidate-Choice Coefficient Estimates

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)
Spending	0.412 (0.240)	0.406 (0.428)	0.434 (0.253)	0.412 (0.246)	0.281 (0.274)	0.561 (0.363)	0.455 (0.534)	0.605 (0.375)	0.413 (0.320)	0.557 (0.352)
Spending Variance ( $\sigma_1$ )	0.000 (737.6)	0.000 (6101)	0.000 (857.6)	0.000 (910.7)	0.000 (481.5)	0.000 (143.3)	0.000 (5967)	0.000 (142.2)	0.000 (317.6)	0.000 (137.4)
Spending <sup>2</sup>	-0.017 (0.019)	-0.018 (0.046)	-0.018 (0.020)	-0.017 (0.020)	-0.012 (0.024)	-0.025 (0.028)	-0.019 (0.051)	-0.027 (0.029)	-0.018 (0.026)	-0.024 (0.028)
Spending <sup>2</sup> Variance ( $\sigma_2$ )	0.000 (265.4)	0.007 (0.046)	0.000 (313.5)	0.000 (325.9)	0.000 (240.1)	0.000 (23.14)	0.000 (205.6)	0.000 (27.11)	0.000 (68.92)	0.000 (22.93)
Log-Lagged Vote Share	0.514 (0.031)	0.397 (0.050)	0.503 (0.031)	0.503 (0.029)	0.515 (0.031)	0.461 (0.038)	0.431 (0.045)	0.453 (0.039)	0.473 (0.035)	0.477 (0.038)
Incumbent Governor		0.716 (0.377)					0.263 (0.528)			
MP $\times$ Female	2.310 (1.211)	0.364 (1.267)	2.408 (1.250)	2.519 (1.305)	3.091 (1.078)	1.347 (1.283)	1.018 (1.212)	1.460 (1.332)	1.834 (1.158)	0.751 (1.180)
MP $\times$ Over 60	0.652 (1.265)	0.344 (1.248)	0.418 (1.261)	0.255 (1.280)	0.158 (1.069)	0.861 (1.269)	0.535 (1.175)	0.392 (1.247)	0.129 (1.085)	1.346 (1.179)
MP $\times$ Rural	-0.417 (0.238)	-0.447 (0.249)	-0.404 (0.256)	-0.365 (0.264)	-0.243 (0.193)	-0.575 (0.243)	-0.506 (0.250)	-0.590 (0.260)	-0.500 (0.220)	-0.677 (0.208)
NA $\times$ Female	-1.630 (1.011)	-1.927 (1.097)	-1.610 (1.019)	-1.397 (1.037)	-1.392 (0.963)	-3.261 (1.133)	-3.161 (1.172)	-3.281 (1.155)	-2.582 (1.052)	-3.107 (1.147)
NA $\times$ Over 60	0.374 (0.993)	0.416 (0.932)	0.568 (1.060)	0.790 (1.073)	0.273 (0.947)	0.914 (1.038)	0.874 (0.991)	1.026 (1.130)	1.020 (1.035)	0.874 (1.025)
NA $\times$ Rural	0.104 (0.152)	-0.010 (0.156)	0.106 (0.155)	0.101 (0.166)	0.080 (0.142)	-0.094 (0.152)	-0.109 (0.152)	-0.099 (0.158)	-0.123 (0.158)	-0.030 (0.150)
PVEM $\times$ Female	-0.857 (1.788)	-0.087 (2.816)	-0.511 (1.823)	-0.204 (1.896)	1.049 (2.938)	-0.290 (1.678)	-0.340 (2.006)	-0.101 (1.748)	-0.286 (1.729)	-0.766 (2.090)
PVEM $\times$ Over 60	-0.220 (1.522)	-0.875 (1.992)	0.005 (1.566)	0.015 (1.569)	0.611 (1.958)	0.097 (1.578)	-0.222 (1.920)	0.202 (1.659)	0.204 (1.556)	0.190 (1.672)
PVEM $\times$ Rural	1.043 (0.267)	1.005 (0.328)	1.044 (0.275)	1.045 (0.274)	0.827 (0.220)	0.818 (0.268)	0.765 (0.292)	0.840 (0.279)	0.716 (0.237)	0.848 (0.247)
PRI $\times$ Female	-1.690 (1.065)	-0.007 (2.384)	-1.455 (1.105)	-1.133 (1.057)	-1.393 (1.335)	-1.757 (1.074)	-0.890 (2.654)	-1.616 (1.128)	-1.305 (1.005)	-2.208 (1.219)
PRI $\times$ Over 60	0.661 (1.066)	0.545 (1.501)	0.199 (0.908)	0.069 (0.882)	0.713 (0.837)	1.385 (1.122)	1.060 (1.453)	0.832 (0.988)	0.569 (0.826)	1.378 (1.022)
PRI $\times$ Rural	0.424 (0.132)	0.666 (0.311)	0.434 (0.138)	0.444 (0.142)	0.385 (0.117)	0.380 (0.151)	0.506 (0.410)	0.375 (0.157)	0.376 (0.142)	0.352 (0.147)
PAN $\times$ Female	-1.110 (1.127)	-2.550 (1.309)	-1.206 (1.143)	-1.198 (1.141)	-1.339 (1.111)	-1.311 (1.137)	-2.294 (1.296)	-1.290 (1.164)	-1.186 (1.037)	-1.647 (1.097)
PAN $\times$ Over 60	1.202 (1.205)	2.771 (1.584)	1.321 (1.179)	1.279 (1.164)	1.865 (0.969)	1.555 (1.248)	2.555 (1.916)	1.617 (1.278)	2.325 (1.109)	1.653 (1.183)
PAN $\times$ Rural	0.404 (0.373)	0.217 (0.394)	0.414 (0.404)	0.370 (0.385)	0.151 (0.281)	0.426 (0.375)	0.162 (0.437)	0.478 (0.411)	0.202 (0.334)	0.388 (0.320)
Menu-Party F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region F.E.	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	1,301	1,301	1,241	1,181	1,301	1,301	1,301	1,241	1,181	1,301

*Notes.* BLP estimates of coefficients driving candidate choice, with robust standard errors in parentheses. The first four rows correspond to  $(\alpha, \sigma)$ , which determine the effectiveness of campaign expenditures according to Equation (1). The remaining rows correspond to  $\beta$ , which characterizes baseline partisanship in Equation (1). For reference, columns (I) and (VI) reproduce the baseline estimates in Table A7. Columns (II) and (VII) add as a control a binary indicator of whether the incumbent governor was from the corresponding party. Columns (III) and (VIII) report estimates after dropping from the sample districts in the top 5% of joint PRI-PVEM spending increases relative to 2009. Columns (IV) and (IX) do the same after dropping the top 10%. Columns (V) and (X) use lagged spending in neighboring districts, Internet availability, and average travel time as alternative instruments.

Table D2: Proportional-Representation Party-Choice Coefficient Estimates

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Log-Lagged Vote Share	0.463 (0.055)	0.502 (0.062)	0.468 (0.056)	0.491 (0.057)	0.469 (0.067)	0.510 (0.078)	0.479 (0.068)	0.508 (0.071)
Incumbent Governor		-0.123 (0.061)				-0.123 0.075		
PVEM $\times$ Female	-1.348 (1.283)	-1.348 (1.284)	-1.619 (1.298)	-1.851 (1.326)	-1.141 (1.468)	-1.224 (1.476)	-1.409 (1.490)	-1.986 (1.562)
PVEM $\times$ Over 60	0.882 (1.311)	0.895 (1.311)	0.783 (1.409)	-0.018 (1.493)	0.888 (1.242)	0.961 (1.247)	0.728 (1.326)	0.312 (1.397)
PVEM $\times$ Rural	0.518 (0.248)	0.546 (0.247)	0.488 (0.250)	0.549 (0.257)	0.320 (0.206)	0.335 (0.206)	0.302 (0.211)	0.369 (0.229)
PRI $\times$ Female	-2.108 (0.760)	-2.395 (0.786)	-2.218 (0.752)	-2.247 (0.751)	-1.239 (0.836)	-1.715 (0.891)	-1.302 (0.837)	-1.449 (0.858)
PRI $\times$ Over 60	3.208 (0.751)	3.159 (0.746)	2.967 (0.799)	2.261 (0.807)	2.703 (0.689)	2.798 (0.678)	2.468 (0.726)	2.043 (0.766)
PRI $\times$ Rural	0.147 (0.118)	0.126 (0.117)	0.102 (0.118)	0.092 (0.120)	0.091 (0.103)	0.031 (0.108)	0.059 (0.104)	0.062 (0.119)
Menu-Party F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region F.E.	No	No	No	No	Yes	Yes	Yes	Yes
Observations	398	398	368	338	398	398	368	338

*Notes.* OLS estimates of  $\beta^{\text{ST}}$ , which drives second-tier choice for PRI-PVEM coalition supporters of how to allocate their PR vote according to Equation (3), with robust standard errors in parentheses. Outside option is 50-50 vote split between the two partners. For reference, columns (I) and (V) reproduce the baseline estimates in Table A8. Columns (II) and (VI) add as a control a binary indicator of whether the incumbent governor was from the corresponding party. Columns (III) and (VII) report estimates after dropping from the sample districts in the top 5% of joint PRI-PVEM spending increases relative to 2009. Columns (IV) and (VIII) do the same after dropping the top 10%.

Table D3: Estimates of Parties' Coalition-Stage Payoffs

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Prob. of Winning FPTP Race Alone	-4.919 (2.734)	-4.870 (2.660)	-4.728 (2.586)	-4.851 (2.583)	-4.563 (2.674)	-4.519 (2.645)	-4.473 (2.583)	-4.960 (2.746)
Incumbent Governor		-0.076 (0.370)				-0.068 (0.457)		
PVEM $\times$ Female					-1.370 (4.946)	-1.387 (4.951)	-0.960 (4.826)	-0.181 (4.759)
PVEM $\times$ Over 60					-6.841 (5.484)	-6.825 (5.496)	-6.304 (5.465)	-4.262 (5.451)
PVEM $\times$ Rural					-1.048 (0.678)	-1.049 (0.678)	-0.898 (0.672)	-0.829 (0.701)
PRI $\times$ Female					-1.826 (7.495)	-1.895 (7.558)	-1.431 (7.414)	-4.005 (7.406)
PRI $\times$ Over 60					0.837 (7.300)	0.930 (7.386)	1.389 (7.483)	4.912 (7.735)
PRI $\times$ Rural					-3.626 (1.140)	-3.662 (1.188)	-3.486 (1.145)	-4.298 (1.023)
Party F.E.	Yes	Yes	Yes	Yes	No	No	No	No
Party-Region F.E.	No	No	No	No	Yes	Yes	Yes	Yes
Log-Likelihood	-285.8	-285.8	-275.3	-264.0	-258.1	-258.1	-250.1	-238.3
Observations	300	300	285	270	300	300	285	270

*Notes.* ML estimates of  $\theta$ , which characterizes PRI and PVEM's (dis)utility from standing down in a district to support their partner's candidate as defined by Equation (7), with standard errors in parentheses. For reference, columns (I) and (V) reproduce the baseline estimates in Table 2. Columns (II) and (VI) add as a control a binary indicator of whether the incumbent governor was from the corresponding party. Columns (III) and (VII) report estimates after dropping from the sample districts in the top 5% of joint PRI-PVEM spending increases relative to 2009. Columns (IV) and (VIII) do the same after dropping the top 10%.

## References

- Berry, Steven, James Levinsohn and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica* 63(4):841–890.
- Díaz Cayeros, Alberto, Beatriz Magaloni, Jorge Olarte and Edgar Franco. 2012. La geografía electoral de 2012. Policy brief ([https://cddrl.fsi.stanford.edu/publications/la\\_geografia\\_electoral\\_de\\_2012\\_méxico](https://cddrl.fsi.stanford.edu/publications/la_geografia_electoral_de_2012_méxico)). México Evalúa, Program on Poverty and Governance, and Center for U.S.-Mexican Studies.
- Dubé, Jean-Pierre, Jeremy T. Fox and Che-Lin Su. 2012. “Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation.” *Econometrica* 80(5):2231–2267.
- Echenique, Federico. 2007. “Finding All Equilibria in Games of Strategic Complements.” *Journal of Economic Theory* 135(1):514–532.
- Echenique, Federico and Aaron Edlin. 2004. “Mixed equilibria are unstable in games of strategic complements.” *Journal of Economic Theory* 118(1):61–79.
- Gandhi, Amit and Jean-François Houde. 2023. “Measuring Substitution Patterns in Differentiated-Products Industries.” Working Paper.
- Gillen, Benjamin J., Sergio Montero, Hyungsik Roger Moon and Matthew Shum. 2019. “BLP-2LASSO for Aggregate Discrete Choice Models with Rich Covariates.” *Econometrics Journal* 22(3):262–281.
- Gordon, Brett R. and Wesley R. Hartmann. 2013. “Advertising Effects in Presidential Elections.” *Marketing Science* 32(1):19–35.
- Hansen, Lars P. 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica* 50(4):1029–1054.
- Heiss, Florian and Viktor Winschel. 2008. “Likelihood approximation by numerical integration on sparse grids.” *Journal of Econometrics* 144(1):62–80.
- Newey, Whitney K. and Daniel McFadden. 1994. Large Sample Estimation and Hypothesis Testing. In *Handbook of Econometrics*, ed. Robert F. Engle and Daniel L. McFadden. Vol. IV Elsevier chapter 36, pp. 2111–2245.
- Shi, Xiaoxia and Matthew Shum. 2015. “Simple Two-Stage Inference for a Class of Partially Identified Models.” *Econometric Theory* 31(3):493–520.
- Vives, Xavier. 1990. “Nash equilibrium with strategic complementarities.” *Journal of Mathematical Economics* 19(3):305–321.