Online Appendix for *Sleeping with the Enemy: Effective Representation under Dynamic Electoral Competition*[*]

Anderson Frey
University of Rochester

Gabriel López-Moctezuma
California Institute of Technology

Sergio Montero
University of Rochester

Contents

[*]This version: June 10, 2021.

A Parties' National Platforms

Figure A1 shows that voters (top) and experts (bottom left) widely agree that, nationally, PRD, PRI, and PAN can be placed in that order on a left-right ideology spectrum. Furthermore, using roll-call data from the 60th Legislature (2006-2009), the bottom-right panel demonstrates that federal legislators' policy positions in the lower house of the Mexican Congress (Cámara de Diputados) are consistent with their party's perceived ideology.

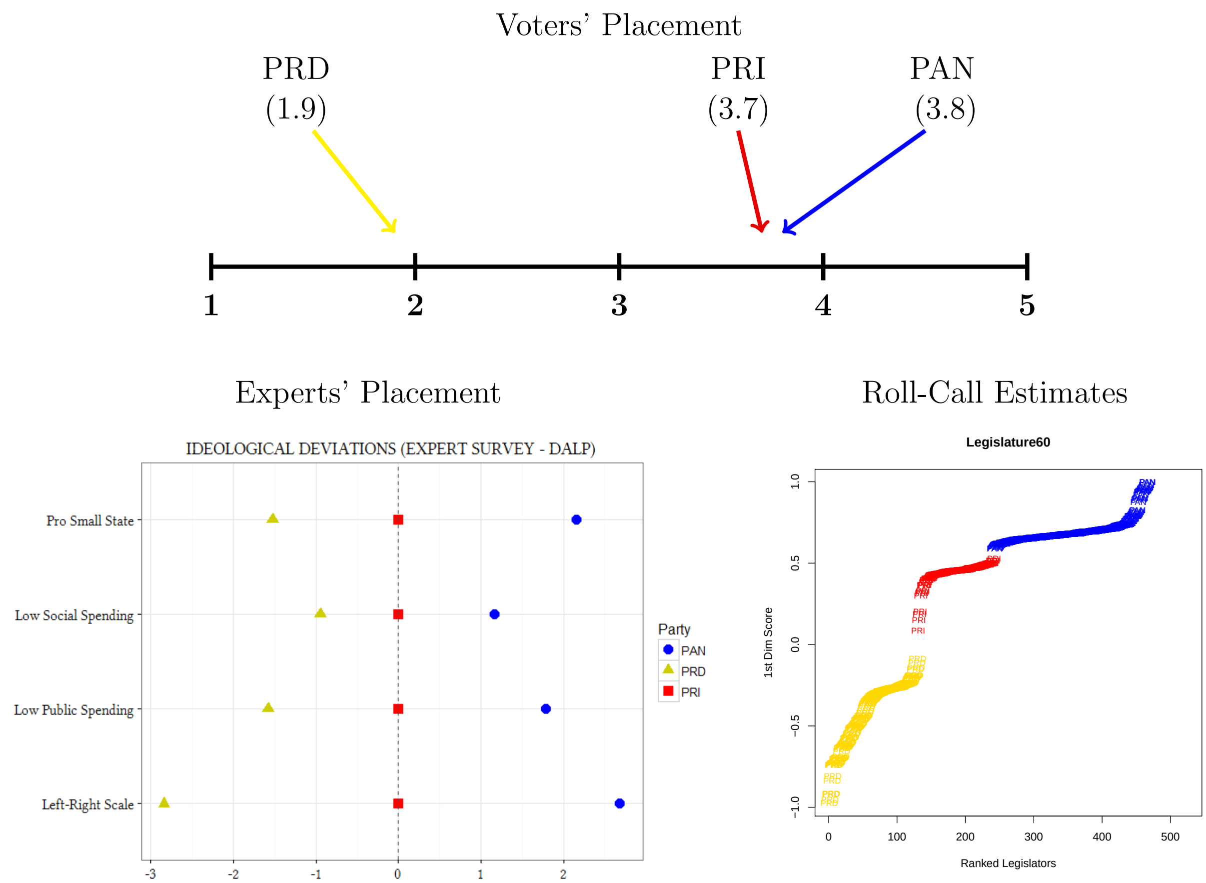

Figure A1: Ideology of Mexican Parties

Notes. The top panel shows results of a nationally representative survey of 1,000 registered voters who where asked in 2012 to place parties on a five-point, left-right ideology scale (arrows point to national averages). Source: Consulta Mitofsky (2012). The bottom-left panel shows (normalized) DALP expert ratings of parties on several policy issues as well as on a broad left-right ideology spectrum. Source: https://sites.duke.edu/democracylinkage/data/. The bottom-right panel plots the estimated first dimension of a Bayesian item-response model (Clinton, Jackman and Rivers, 2004). Legislators' policy positions and their ideological rank in the 60th Legislature are shown in yellow (PRD), red (PRI), and blue (PAN).

B Additional Figures and Tables

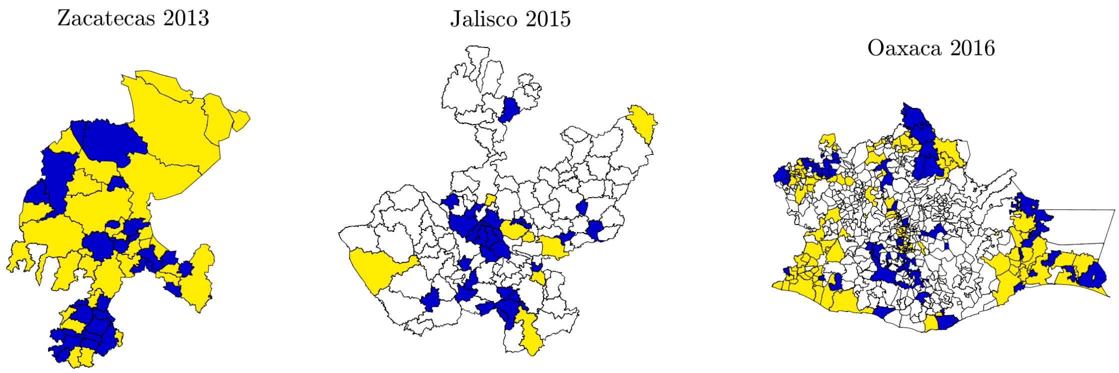

Figure B1: Coalition Choices in Selected States

Notes. This figure shows maps of the Mexican states of Zacatecas in 2013, Jalisco in 2015, and Oaxaca in 2016. Blue (yellow) municipalities correspond to mayoral races where PAN and PRD jointly nominated a PAN (PRD) candidate. White municipalities correspond to mayoral races where PAN and PRD ran independent candidates.

Table B1: Regression Discontinuity Design

Dependent Variable:	Social Spending (% of municipal budget)							
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
PAN	-1.724*	-1.742*	-1.794	-1.703	-1.598**	-1.571**	-2.020*	-2.082*
	(1.007)	(1.052)	(1.447)	(1.496)	(0.740)	(0.777)	(1.106)	(1.159)
PRD	2.615**	2.420**	2.637	1.898	3.578**	3.401**	2.742**	2.629**
	(1.157)	(1.225)	(1.670)	(1.787)	(0.884)	(0.945)	(1.251)	(1.337)
PAN:Coalition		-1.304		-4.389		-0.788		-0.532
		(3.389)		(5.784)		(2.558)		(3.831)
PRD:Coalition		2.577		6.963		2.316		2.030
		(3.386)		(4.763)		(2.643)		(3.624)
Intercept	30.116	30.628	30.294	30.666	29.602	30.049	30.184	30.758
Bandwidth	10	10	10	10	20	20	20	20
Observations	4025	4025	4025	4025	6275	6275	6275	6275
Polynomial	Linear	Linear	Quad.	Quad.	Linear	Linear	Quad.	Quad.

Notes. *$p < 0.1$, **$p < 0.05$. Standard errors are clustered by municipality and presented in parentheses. All regressions include year fixed effects and also control for (i) a dummy that indicates whether the election was held concurrently with state elections, (ii) the log of total voters in the municipality, and (iii) the share of all previous terms with a PRI mayor (starting in 1997). Observations are weighted by the uniform kernel.

Table B2: Summary Statistics of Characteristics of Electoral Environment

	Obs.	Min.	Mean	Max.	St. Dev.	IQR
PRI_incumbency	4132	0.00	2.83	5.00	1.26	2.00
PAN_incumbency	4132	0.00	1.10	5.00	1.17	2.00
PRD_incumbency	4132	0.00	0.80	5.00	1.09	1.00
PRI_win	4132	0.00	0.49	1.00	0.50	1.00
PAN_win	4132	0.00	0.24	1.00	0.42	0.00
PRD_win	4132	0.00	0.15	1.00	0.35	0.00
PRD_coalition	4132	0.00	0.13	1.00	0.33	0.00
IND_coalition	4132	0.00	0.72	1.00	0.45	1.00
PAN_coalition	4132	0.00	0.15	1.00	0.36	0.00
conc_Congress	4132	0.00	0.41	1.00	0.49	1.00
conc_President	4132	0.00	0.19	1.00	0.39	0.00
conc_Governor	4132	0.00	0.56	1.00	0.50	1.00
rural	4059	0.00	0.45	1.00	0.50	1.00
pop_60	4059	3.80	17.87	43.66	5.52	7.00
poverty	4059	1.00	77.60	249.00	53.22	79.00
pop_female	4059	41.32	50.93	55.61	1.51	1.83
governor_party	4132	1.00	1.72	4.00	1.08	1.00
circ_1	4132	0.00	0.18	1.00	0.38	0.00
circ_2	4132	0.00	0.16	1.00	0.37	0.00
circ_3	4132	0.00	0.30	1.00	0.46	1.00
circ_4	4132	0.00	0.20	1.00	0.40	0.00
circ_5	4132	0.00	0.16	1.00	0.37	0.00

Notes. The unit of observation is municipality-electoral cycle. Variable descriptions: `PRI_incumbency`, `PAN_incumbency`, and `PRD_incumbency` measure the number of victories by each respective party in the past five electoral cycles; `PRI_win`, `PAN_win`, and `PRD_win` are binary indicators of election victory for each respective party; `PRD_coalition`, `IND_coalition`, and `PAN_coalition` are binary indicators of whether PAN and PRD nominated, respectively, a PRD coalition candidate, independent candidates, or a PAN coalition candidate; `conc_Governor`, `conc_Congress`, and `conc_President` are binary indicators of concurrent gubernatorial, congressional, and presidential elections; `rural` is a binary indicator of rural/urban status; `pop_60` and `pop_female` measure the percentage of the voting-age population that is over 60 and female, respectively; `poverty` is a poverty index; `governor_party` is a discrete variable with the party identity of the incumbent governor; and `circ_1` through `circ_5` are dummies corresponding to Mexico's five electoral regions as designated by the federal electoral authority.

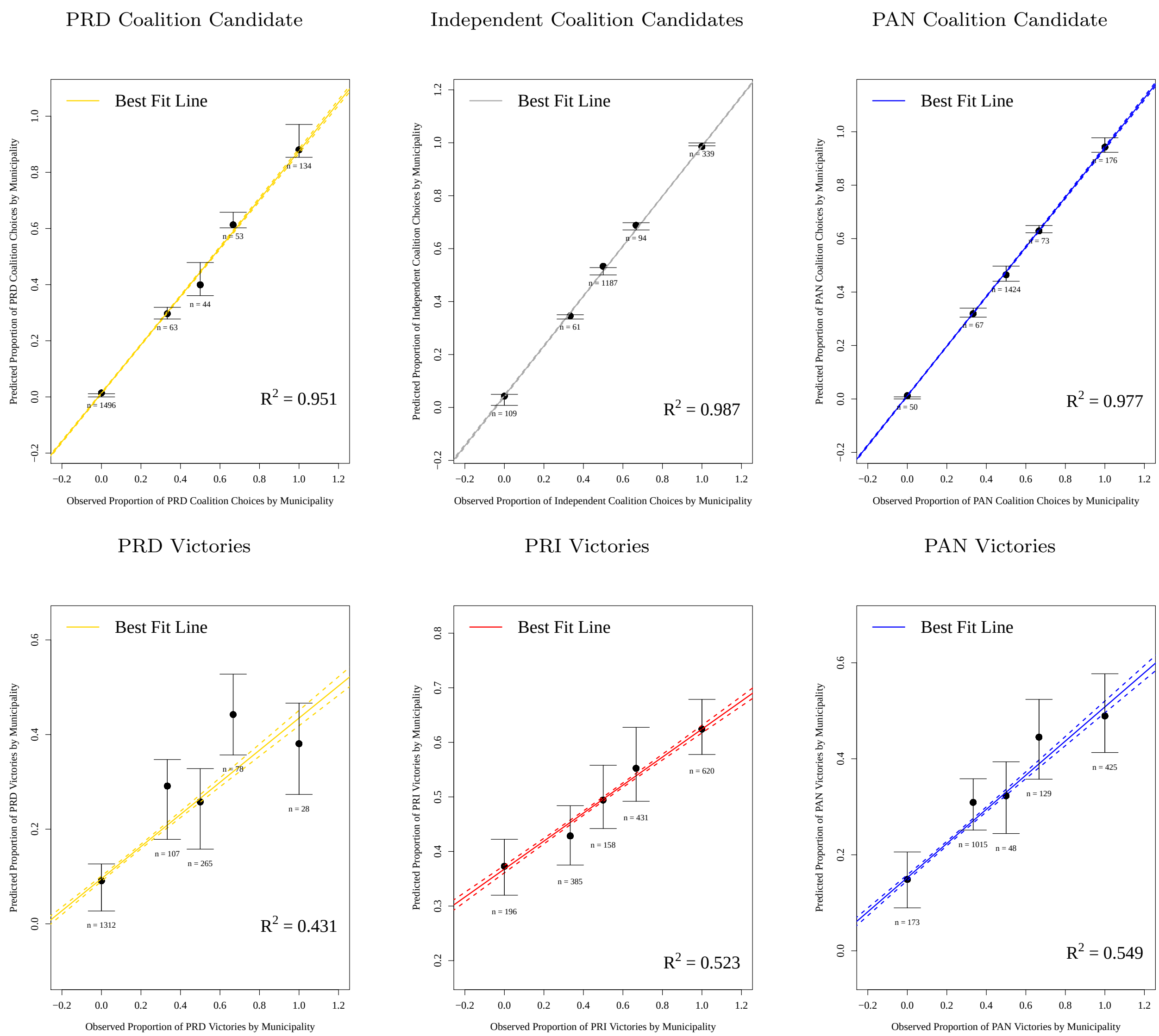

Figure B2: Model Fit

Notes. This figure summarizes goodness of fit of our structural model with regard to PAN-PRD coalition choices and parties' municipal election victories. The top panels plot, for each coalition choice—i.e., joint PRD candidate (left), independent candidates (center), PAN coalition candidate (right)—observed and predicted choice shares per municipality. The bottom panels plot, for PRD (left), PRI (center), and PAN (right), observed and predicted victory shares per municipality. Solid lines depict linear fit, and dashed lines delimit 95% confidence intervals. Solid points highlight the mean predicted probability for each observed share in the data, and whiskers delimit the corresponding interquartile range.

Table B3: (Exogenous) Characteristics of Electoral Environment and Party Performance

	PRD	PRI	PAN
conc_Governor	-0.076	-0.257**	-0.021
	(0.114)	(0.106)	(0.137)
conc_Congress	-0.546***	-0.848***	-1.154***
	(0.193)	(0.160)	(0.165)
conc_President	0.411**	0.799***	0.621***
	(0.193)	(0.187)	(0.176)
rural	0.300**	-0.017	0.012
	(0.147)	(0.108)	(0.142)
pop_60	0.054***	0.032**	0.052***
	(0.016)	(0.013)	(0.014)
pop_female	0.031	-0.019	-0.073
	(0.047)	(0.046)	(0.054)
poverty	0.001	0.002	0.002
	(0.001)	(0.001)	(0.001)
Region F.E.	Yes	Yes	Yes

Notes. $^{**}p < 0.05$, $^{***}p < 0.01$. Estimates correspond to coefficients $(\beta_x^p)_{p \in \{1,\ldots,P\}}$ in Equation (1), which governs parties' electoral prospects. The omitted category is $p =$ OTHER. Nonparametrically bootstrapped standard errors are reported in parentheses.

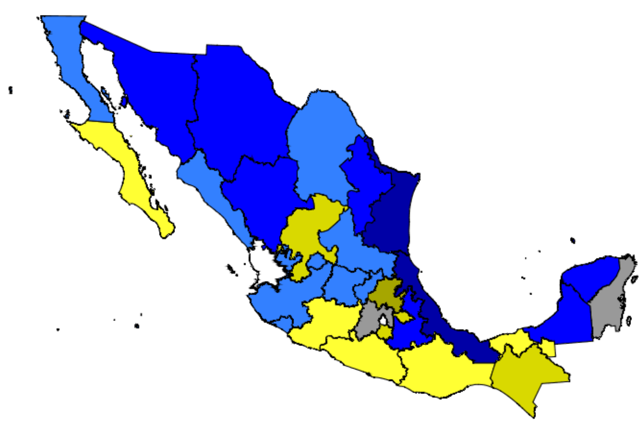

Figure B3: Bargaining Power across States

Notes. This figure shows estimates of PAN's bargaining power (λ_s) relative to PRD. States in blue (yellow) correspond to $\hat{\lambda}_s > 0.5$ ($\hat{\lambda}_s < 0.5$), with lighter shades indicating higher (lower) values. States in gray correspond to $\hat{\lambda}_s = 0.5$.

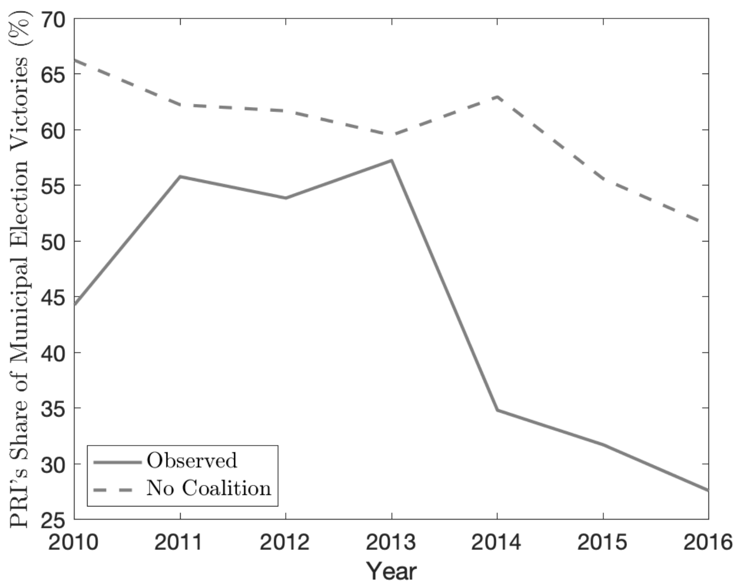

Figure B4: Evolution of Election Outcomes with and without PAN-PRD Coalitions

Notes. For municipalities in our main sample that experience at least one PAN-PRD coalition, this figure shows a three-year moving average of PRI's share of municipal election victories as observed in the data (solid) and under a counterfactual scenario with no PAN-PRD coalitions (dashed).

C PAN-PRD Coalitions: Descriptive Evidence

To motivate our model, we present some descriptive evidence that sheds light on the main predictors of coalition formation in the data. While these results do not account for parties' strategic motives when deciding whether to nominate common candidates, they reveal systematic empirical patterns that underlie key modeling choices.

The left-hand panel of Figure C1 plots estimates from a Bayesian logistic regression of a binary indicator of PAN-PRD coalition formation on a measure of PRI entrenchment in power, controlling for municipality and electoral-cycle random effects. As discussed in the paper, there are reasons to expect Mexican parties—especially the hegemonic PRI—may exploit their time in power to progressively build an electoral advantage over their rivals. In line with this perspective, we construct a measure of entrenched incumbency that goes beyond simply considering which party is in power at the time of the election. Rather, we wish to account for the entire recent history of incumbency by each party. Figure 2 in the paper makes clear that the PAN-PRD coalitions were deployed in full force beginning in 2010, which we take as the starting point for our analysis. We then measure entrenched incumbency as the share of the *past five* electoral cycles won by each party.[1]

As shown in Figure C1, PAN-PRD coalitions are more likely in PRI municipal strongholds than elsewhere. We find that, while the probability of a PAN-PRD coalition is only 20% in a municipality where PRI has not governed for the past five electoral cycles, this probability almost doubles to 40% in a municipality with five cycles of uninterrupted PRI rule.

To examine the nature of PAN-PRD coalitions—i.e., which party headlines the coalition—the right-hand panel of Figure C1 plots the predicted conditional probability that PAN and PRD jointly nominate a PAN candidate as a function of PAN entrenched incumbency. Perhaps unsurprisingly, we find that the party affiliation of PAN-PRD coalition candidates is largely determined by the relative strength of the coalition partners. In municipalities where PAN has not governed for the past five electoral cycles, the probability that a PAN candidate leads the PAN-PRD coalition is 25%. However, PAN almost surely leads the coalition in its municipal strongholds.

We also look at the electoral success of PAN-PRD coalition candidates. First, we fit a Bayesian multinomial logit model of parties' electability given our measures of entrenched incumbency, controlling for municipality and electoral-cycle random effects. The left-hand panel of Figure C2 shows the resulting predicted probabilities of victory for PAN, PRI, PRD, and other parties as a function of PRI entrenchment. As expected, PRI's probability of

[1] Our results are robust to alternative windows of incumbency and to expanding the sample to include the period 1999-2009, during which PAN and PRD proposed joint candidates in only 1.5% of municipal elections—see Online Appendix E.

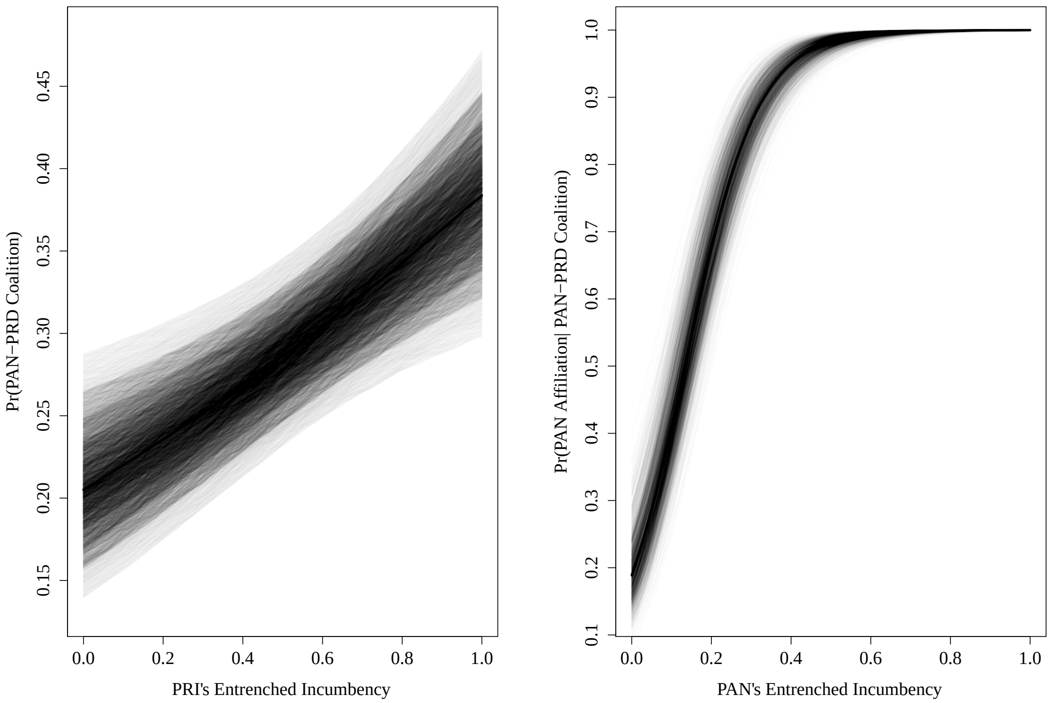

Figure C1: PAN-PRD Coalition Choices and Entrenched Incumbency

Notes. The left-hand plot shows the predicted probability of PAN-PRD coalition formation as a function of PRI's entrenchment. The right-hand plot shows the predicted conditional probability of a PAN candidate headlining the PAN-PRD coalition as a function of PAN's entrenchment. Estimates are from Bayesian logistic regressions, and shaded areas cover 95% confidence intervals. As we manipulate a party's incumbency history, we set the ratio of other parties' incumbency histories equal to the median in the data.

victory rises markedly as its entrenchment increases, ranging from 28% to almost 60%, which is considerable given that these are multi-candidate races.

Focusing on the electability of PAN and PRD candidates, the right-hand panel of Figure C2 plots the predicted probability of victory of coalition versus independent candidates as a function of PRI entrenchment. While independent candidates fare worse in PRI strongholds, coalition candidates' prospects are relatively better given high levels of PRI entrenchment. Although we cannot disentangle the causes with these reduced-form regressions, the evidence is consistent with the dynamic tradeoff at the heart of our argument. In line with standard spatial-voting intuition, PAN and PRD pay a substantial cost at the polls from forming an ideologically incompatible coalition in races against the hegemonic PRI that are relatively competitive. However, when PRI entrenchment is high, voters are seemingly willing to put their ideological tastes aside to temporarily support PAN-PRD coalition candidates.

To analyze the persistence of coalition choices, we estimate the influence of previous coalitions and their electoral success on current coalition choices using linear probability models

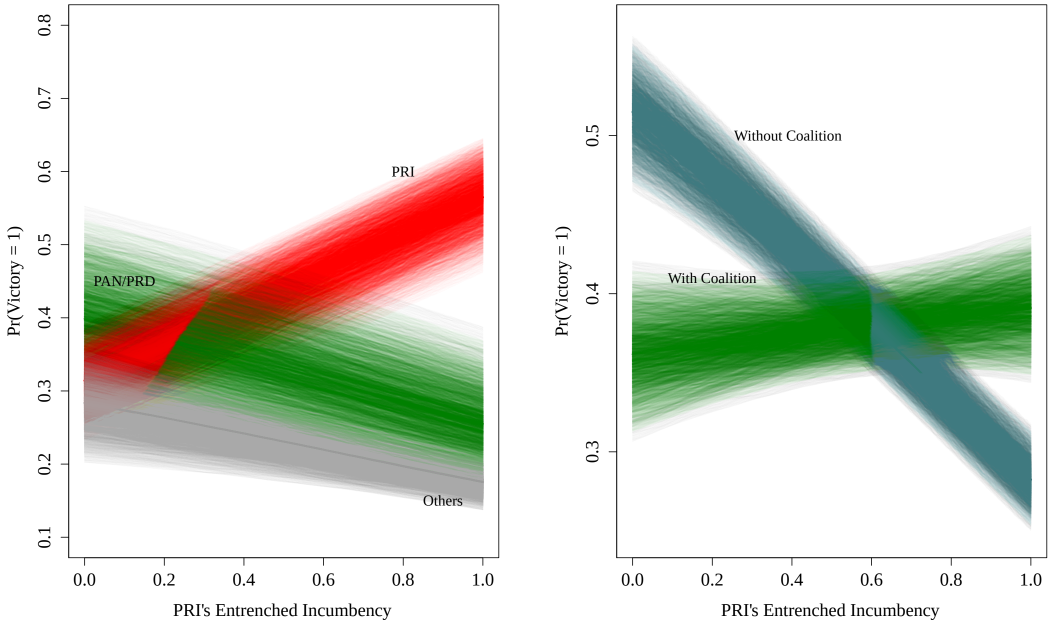

Figure C2: Probability of Victory in Municipal Elections

Notes. The left-hand panel shows the predicted probability of victory by each party (PRI, PAN, PRD, others) as a function of PRI's entrenchment. The right-hand panel focuses on the probability of victory by a PAN or PRD candidate conditional on the parties either forming or not an electoral coalition. Estimates are from Bayesian multinomial logit regressions, and shaded areas cover 95% confidence intervals. As we manipulate PRI's incumbency history, we set the ratio of other parties' histories equal to the median in the data.

that account for municipality and electoral-cycle fixed effects. Table C1 presents our results. Column (I) shows the estimated effect of the number of previous PAN-PRD coalitions in the last five electoral cycles (`sum_coalition`) on the likelihood of coalition formation. As reported, the presence of an additional coalition in the past reduces the probability of a current coalition by 34 percentage points. Column (II) shows that this negative effect is amplified by the presence of previous coalition victories in the past five electoral cycles (`one_coalition_win` and `two_coalition_wins`). While the effect of an additional coalition is -37 percentage points without a coalition victory, this effect almost doubles, to -65, in the presence of two electorally successful coalitions.[2]

Overall, this evidence is again consistent with our central argument. The two ideological rivals, PAN and PRD, are significantly less likely to nominate a joint candidate once they have successfully depleted the hegemonic PRI's entrenched incumbency advantage.

In column (III) of Table C1, we probe whether there is evidence of a direct relationship between the number of races per state and coalition choices. Under a purely static account of coalition formation in which coalition partners are myopic but may trade municipalities to al-

[2]Results are very similar if we only control for coalition decisions and outcomes in the previous cycle.

Table C1: PAN-PRD Coalition Likelihood Given Past Coalition Choices and State Size

	(I)	(II)	(III)
`sum_coalition`	−0.342***	−0.368***	−0.342***
	(0.017)	(0.023)	(0.017)
`one_coalition_win`		0.072	
		(0.059)	
`two_coalition_wins`		1.330***	
		(0.170)	
`state_size` (log)			−0.019
			(0.027)
`sum_coalition`× `one_coalition_win`		0.006	
		(0.050)	
`sum_coalition`× `two_coalition_wins`		−0.650***	
		(0.054)	
Election Cycle FE	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes
Observations	4,670	4,670	4,670
R^2	0.280	0.286	0.280
F Statistic	146.086***	95.639***	127.849***

Notes. *p<0.1; **p<0.05; ***p<0.01. Clustered standard errors in parentheses.

locate resources more efficiently and thus improve their current electoral prospects, one would expect coalitions to be more likely to emerge where there are more races to be traded. Contrary to this pure race-trading motive, we do not find a statistically-significant or substantively-meaningful relationship between the number of available races to trade and the likelihood of a PAN-PRD coalition. In fact, the associated coefficient is negative.

With regard to coalition survival across electoral cycles, we exploit the duration of observed PAN-PRD coalitions within municipalities to show further evidence of the dynamic tradeoffs at play. First, we provide evidence that PAN-PRD coalitions are predominantly short-lived. Second, we show that the rate at which these electoral coalitions are dissolved from one election to the next, along with their expected duration, depend systematically on past coalition choices and their success in eroding PRI's entrenched incumbency.

As we focus on the period 2010-2016, PAN-PRD coalitions can last between one (if short-lived) and three electoral cycles (if uninterrupted). Table C2 shows non-parametric Kaplan-Meier estimates of PAN-PRD coalitions' survival function for each potential duration, along with 95% confidence intervals. While we find that the expected probability of observing a PAN-PRD coalition that lasts at least one electoral cycle is 35%, the coalition survival is reduced to 23% for coalitions lasting at least two cycles and is cut in half, to 18%, for coalitions lasting more than two electoral cycles.

More importantly, we find that the expected duration of PAN-PRD coalitions is a function of parties' past coalition choices (measured as above with `sum_coalition`) and their electoral success (`one_coalition_win` and `two_coalition_wins`). To show this relationship, we estimate both a semiparametric Cox proportional hazard model, in which the dependent variable

Table C2: Kaplan-Meier PAN-PRD Coalition Survival Curve

Duration (Cycles)	N at Risk	N	N Censored	Survival	S.E	Lower	Upper
1 +	2196	1430	766	0.35	0.01	0.33	0.37
2 +	673	232	441	0.23	0.01	0.21	0.25
3 +	133	26	107	0.18	0.01	0.16	0.21

Notes. This table presents the survival curve of PAN-PRD coalitions, along with 95% confidence intervals.

is given by the hazard rate, and a parametric duration model, in which the dependent variable is the time until a coalition dissolution (i.e., when parties decide to nominate independent candidates), via a Weibull duration model.[3] In both cases, we find that the expected duration of PAN-PRD coalitions is inversely related to their past electoral success.

The estimates from the Cox proportional hazard model in the top panel of Table C3 show that, when a municipality has not experienced a successful PAN-PRD coalition in the past, the expected hazard of a coalition dissolution in year t decreases around 87% in the presence of one additional coalition (i.e., $(1 - 0.13) \times 100$). However, previous coalition victories significantly increase the failure rate of PAN-PRD coalitions. In particular, for a coalition with two victories in the past five electoral cycles, the presence of one additional coalition increases the coalition failure rate by 74%.

The estimates from the Weibull duration model in the bottom panel of Table C3 also imply that expected duration is significantly shorter for coalitions with past victories. For instance, we find that, conditional on observing three electoral cycles with PAN-PRD coalitions, the expected duration of a coalition with no previous victories is 1.7 and 2.5 times longer, respectively, than coalitions with one and two previous victories.

This evidence on coalition persistence and duration provides additional support for the importance of parties' dynamic incentives as key factors behind PAN-PRD coalition choices. These alliances are short-lived, and, once they manage to erode the entrenchment of the centrist PRI, they are likely to be abandoned.

[3]For the Weibull duration model, we control for municipality random effects and right-censoring.

Table C3: PAN-PRD Coalition Duration

	Cox Proportional Hazard Model				
	Coef (exp)	S.E (exp)	*p*-value	95% Lower C.I.	95% Upper C.I.
`sum_coalition`	0.13	0.08	0.00	0.11	0.15
`one_coalition_win`	0.17	0.27	0.00	0.10	0.29
`two_coalition_wins`	0.01	1.62	0.00	0.00	0.22
`sum_coalition` × `one_coalition_win`	3.28	0.22	0.00	2.12	5.07
`sum_coalition` × `two_coalition_wins`	16.39	0.76	0.00	3.71	72.45
	Weibull Duration Model				
	Coef (exp)	S.E (exp)	*p*-value	95% Lower C.I.	95% Upper C.I.
`sum_coalition`	0.69	0.01		0.68	0.70
`one_coalition_win`	0.33	0.02		0.29	0.38
`two_coalition_wins`	1.13	0.16		0.77	1.41
`sum_coalition` × `one_coalition_win`	-0.28	0.02		-0.31	-0.24
`sum_coalition` × `two_coalition_wins`	-0.69	0.08		-0.82	-0.52

Notes. The top panel presents estimates of a Cox proportional hazard model, in which the hazard rate is a function of the number of past coalition choices interacted with the number of past coalition victories (zero, one, or two victories). The bottom panel shows estimates from a Bayesian Weibull regression, controlling for municipality random effects.

D Estimation Details

D.1 Likelihood of the Data

In each municipality $m \in \{1, \ldots, M\}$, for $T_m \geq 1$ electoral cycles, we observe the sequence $d_m = \{d_{m1}, \ldots, d_{mT_m}\}$ of coalition configuration choices by PAN and PRD. We also observe the sequence $I_m = \{I_{m1}, \ldots, I_{m,T_m+1}\}$ of incumbency histories as well as exogenous municipal characteristics, $x_m = \{x_{m1}, \ldots, x_{mT_m}\}$. To economize on notation, we subsume dependence on x_{mt} under the description of the municipal electoral cycle itself, and we accordingly write $\pi(\xi_{mt}|x_{mt})$ simply as $\pi_{mt}(\xi_{mt})$.

The likelihood of observing the pair $(d_{mt}, z_{m,t+1})$ given z_{mt} is

$$\mathcal{L}_{mt}(d_{mt}, z_{m,t+1}|z_{mt}; \varphi) = \prod_{j \in \mathcal{J}} \Big[l_{jmt}(z_{mt}; \varphi) g_{jmt}(z_{m,t+1}|z_{mt}; \varphi) \Big]^{d_{jmt}},$$

where

$$l_{jmt}(z_{mt}; \varphi) = \frac{\exp\left[v_{jmt}(z_{mt}; \varphi)\right]}{\sum_{j' \in \mathcal{J}} \exp\left[v_{j'mt}(z_{mt}; \varphi)\right]} \tag{D1}$$

and $g_{jmt}(z_{m,t+1}|z_{mt}; \varphi) = \pi_{m,t+1}(\xi_{m,t+1}) \sum_{p \in C_j} \tilde{w}^p_{mt} f^p_{jmt}(z_{mt}; \beta)$, provided that $I^p_{mt} = (w^p_{m,t-N}, \ldots, w^p_{m,t-1})$ and $\tilde{I}^p_{m,t+1} = (\tilde{w}^p_{m,t-N+1}, \ldots, \tilde{w}^p_{mt})$ are mutually consistent (i.e., $w^p_{m\tau} = \tilde{w}^p_{m\tau}$ for all $t-1 \geq \tau \geq t-N+1$; otherwise, $g_{jmt}(z_{m,t+1}|z_{mt}; \varphi) = 0$.) The intuition is straightforward. At the start of electoral cycle t, the coalition partners observe the state of the election, z_{mt}. Given φ, optimal dynamic behavior, as described by Equations (3) and (4) in the paper, compels the parties to select—from the perspective of the researcher, who doesn't observe ϵ_{mt}—coalition arrangement $j \in \mathcal{J}$ with probability $l_{jmt}(z_{mt}; \varphi)$. Conditional on this choice, the outcome of the election in period t is determined by the probabilities of victory, $f^p_{jmt}(z_{mt}; \beta)$. Parties' incumbency histories then evolve to $I^p_{m,t+1}$, a new set of candidates, $\xi_{m,t+1}$, is drawn from $\pi_{m,t+1}$, and thus the state of the election transitions to $z_{m,t+1} = (I_{m,t+1}, \xi_{m,t+1})$ with probability $g_{jmt}(z_{m,t+1}|z_{mt}; \varphi)$.

Since ξ_{mt} is unobserved by the researcher, it must be integrated out to obtain the likelihood of the data. Given that unobserved valence is independently distributed over time according to π, by the law of iterated expectations we have

$$\mathcal{L}_{mt}(d_{mt}, I_{m,t+1}|I_{mt}; \varphi) = \sum_{\xi_{mt} \in \Xi} \pi_{mt}(\xi_{mt}) \prod_{j \in \mathcal{J}} \left[l_{jmt}(z_{mt}; \varphi) \sum_{p \in C_j} w^p_{mt} f^p_{jmt}(z_{mt}; \beta) \right]^{d_{jmt}}, \tag{D2}$$

where Ξ denotes the set of all possible realizations of ξ_{mt}. Thus, the likelihood of the observed pair (d_m, I_m) given initial incumbency history I_{m1} can be written as

$$\mathcal{L}_m(d_m, I_m | I_{m1}; \varphi) = \prod_{t=1}^{T_m} \mathcal{L}_{mt}(d_{mt}, I_{m,t+1} | I_{mt}; \varphi),$$

and the log-likelihood of the sample is given by

$$\sum_{m=1}^{M} \log \left[\mathcal{L}_m(d_m, I_m | I_{m1}; \varphi) \right]. \tag{D3}$$

D.2 Estimation Procedure

Maximum likelihood estimates of the model parameters, φ, could be obtained by directly maximizing (D3). However, this approach poses two considerable computational challenges. First, given (D2), evaluating the log-likelihood of the data involves taking logs of sums of probabilities, which can give rise to numerical instability. While this may be addressed, as is standard, with an application of the Expectation-Maximization (EM) algorithm, computation of the conditional value functions, $v_{jmt}(z_{mt}; \varphi)$, in (D1) relies on very costly fixed-point calculations based on Equation (4). To sidestep this burden, we follow the two-stage estimation procedure proposed by Arcidiacono and Miller (2011).

D.2.1 First Stage

In the first stage, the model's computationally-expensive conditional choice probabilities, $l_{jmt}(z_{mt}; \varphi)$, can be replaced with nonparametric or semiparametric estimates, $\rho_{jmt}(z_{mt})$.[4] The EM algorithm can then be employed to obtain consistent and asymptotically normal estimates of β and π—which determine the state transition probabilities, $g_{jmt}(z_{m,t+1} | z_{mt}; \varphi)$—and of the conditional choice probabilities, ρ. The algorithm proceeds by iteratively alternating between an Expectation Step and a Maximization Step, which we describe in turn.

Expectation Step. This step involves computing an estimate of the posterior distribution of unobserved valence conditional on the data and a preliminary estimate of (β, π, ρ). The posterior is then used to compute the conditional expected log-likelihood of the data that is maximized in the next step.

[4] In our implementation, we set

$$\rho_{jmt}(z_{mt}; \zeta) = \frac{\exp[h_j(z_{mt}, x_{mt}; \zeta)]}{\sum_{j' \in \mathcal{J}} \exp[h_{j'}(z_{mt}, x_{mt}; \zeta)]},$$

where h_j is a flexible function (second-degree polynomial) of z_{mt} and x_{mt}, parameterized by ζ.

At the nth iteration of the EM algorithm, let $q^{(n)}_{\xi mt}$ denote the posterior probability that $\xi_{mt} = \xi \in \Xi$ in municipality m and electoral cycle t conditional on the data and current estimates $(\beta^{(n-1)}, \pi^{(n-1)}, \rho^{(n-1)})$ of (β, π, ρ). Given (D2), the joint likelihood of $\xi_{mt} = \xi$ and the data can be written as $\pi^{(n-1)}_{mt}(\xi)\mathcal{L}_{mt}(d_{mt}, I_{m,t+1}|I_{mt}, \xi_{mt} = \xi; \beta^{(n-1)}, \pi^{(n-1)}, \rho^{(n-1)})$, where

$$\mathcal{L}_{mt}(d_{mt}, I_{m,t+1}|I_{mt}, \xi_{mt} = \xi; \beta^{(n-1)}, \pi^{(n-1)}, \rho^{(n-1)}) = \prod_{j\in\mathcal{J}} \left[\rho^{(n-1)}_{jmt}(I_{mt}, \xi) \sum_{p\in C_j} w^p_{mt} f^p_{jmt}(I_{mt}, \xi; \beta^{(n-1)}) \right]^{d_{jmt}}.$$

By Bayes' rule, it follows that

$$q^{(n)}_{\xi mt} = \frac{\pi^{(n-1)}_{mt}(\xi)\mathcal{L}_{mt}(d_{mt}, I_{m,t+1}|I_{mt}, \xi_{mt} = \xi; \beta^{(n-1)}, \pi^{(n-1)}, \rho^{(n-1)})}{\sum_{\xi_{mt}\in\Xi} \pi^{(n-1)}_{mt}(\xi_{mt})\mathcal{L}_{mt}(d_{mt}, I_{m,t+1}|I_{mt}, \xi_{mt}; \beta^{(n-1)}, \pi^{(n-1)}, \rho^{(n-1)})}. \tag{D4}$$

Note that the right-hand side of (D4) can be easily computed using the data and current estimates of (β, π, ρ).

This posterior can also be used to update the estimated unconditional, or ex-ante, distribution of unobserved valence. By the law of iterated expectations,

$$\pi^{(n)}(\xi|x) = \frac{\sum_{m=1}^{M} \sum_{t=1}^{T_m} q^{(n)}_{\xi mt} \mathbf{1}\{x_{mt} = x\}}{\sum_{m=1}^{M} \sum_{t=1}^{T_m} \mathbf{1}\{x_{mt} = x\}}. \tag{D5}$$

If x is high-dimensional, a smoothing kernel can be introduced in (D5). In our implementation, we allow $\pi(\cdot|x)$ to vary only by electoral region, thus overcoming the curse of dimensionality without raising the computational cost of estimation.

Maximization Step. Having updated the ex-ante and posterior distributions of unobserved valence, updated estimates $(\beta^{(n)}, \rho^{(n)})$ of (β, ρ) can be computed by maximizing the conditional expected log-likelihood of the data:

$$\max_{\beta,\rho} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{\xi\in\Xi} \sum_{j\in\mathcal{J}} q^{(n)}_{\xi mt} d_{jmt} \left[\log\left(\rho_{jmt}(I_{mt}, \xi)\right) + \sum_{p\in C_j} w^p_{mt} \log\left(f^p_{jmt}(I_{mt}, \xi; \beta)\right) \right]. \tag{D6}$$

Notice that (D6) is additively separable in β and ρ.

In sum, given starting values $(\beta^{(0)}, \pi^{(0)}, p^{(0)})$, the EM algorithm proceeds as follows: at each iteration n, the Expectation Step yields updates $q^{(n)}$ and $\pi^{(n)}$ using (D4) and (D5), respectively, and the Maximization Step yields updates $\beta^{(n)}$ and $\rho^{(n)}$ solving (D6). As in any

EM implementation, the log-likelihood of the data (D3) increases after every iteration and is guaranteed to converge to a (local) maximum.[5]

D.2.2 Second Stage

In the second stage, having obtained estimates $\hat{\beta}$, $\hat{\pi}$, and $\hat{\rho}$, a Generalized Method of Moments (GMM) estimator of the remaining parameters, θ and β_λ, can be constructed by exploiting Equation (D1). Since $\hat{\rho}_{jmt}(z_{mt})$ is consistent, it follows from (D1) that (asymptotically)

$$\log\left(\hat{\rho}_{jmt}(z_{mt})\right) - \log\left(\hat{\rho}_{0mt}(z_{mt})\right) - \left[v_{jmt}(z_{mt};\varphi_0) - v_{0mt}(z_{mt};\varphi_0)\right] = 0, \tag{D7}$$

where φ_0 denotes the true value of the model parameters. Equation (D7)—which must hold for all alternatives $j \neq 0$ and states z_{mt}—yields moments conditions that identify θ and β_λ by ensuring that the relative value of alternative j as described by the payoff difference $v_{jmt}(z_{mt};\varphi) - v_{0mt}(z_{mt};\varphi)$ is consistent with the corresponding log-odds, $\log\left(\hat{\rho}_{jmt}(z_{mt})\right) - \log\left(\hat{\rho}_{0mt}(z_{mt})\right)$, recovered from the data. Estimation and inference then follow standard GMM practice—see Arcidiacono and Miller (2011) for a formal discussion.

Implementation of this GMM estimator, however, still requires computation of the conditional value functions, $v_{jmt}(z_{mt};\varphi)$, which, as noted, involves costly fixed-point calculations. Yet, having estimated $(\hat{\beta}, \hat{\pi}, \hat{\rho})$ in the first stage, it is possible to instead approximate $v_{jmt}(z_{mt};\varphi)$ via forward simulation. As shown by Arcidiacono and Miller (2011),

$$\begin{aligned} v_{jmt}(z_{mt};\varphi) &= U_{jmt}(z_{mt};\varphi) + \\ &\sum_{\tau=t+1}^{\infty}\sum_{z_{m\tau}} \delta^{\tau-t}\left[U_{0m\tau}(z_{m\tau};\varphi) - \log\left(\rho_{0m\tau}(z_{m\tau})\right)\right]\kappa_{m\tau}(z_{m\tau}|z_{mt};\varphi), \end{aligned} \tag{D8}$$

where

$$\kappa_{m\tau}(z_{m\tau}|z_{mt};\varphi) = \begin{cases} g_{jmt}(z_{m,t+1}|z_{mt};\varphi) & \text{if } \tau = t+1, \\ \sum_{z_{m,\tau-1}} g_{0m,\tau-1}(z_{m\tau}|z_{m,\tau-1};\varphi)\kappa_{m,\tau-1}(z_{m,\tau-1}|z_{mt};\varphi) & \text{if } \tau > t+1. \end{cases}$$

In other words, $v_{jmt}(z_{mt};\varphi)$ can be written as the expected discounted sum of payoffs the coalition would obtain if it were to choose $\tilde{d}_{jmt} = 1$ and $\tilde{d}_{0m\tau} = 1$ for all $\tau > t$, with payoffs adjusted by $-\log\left(\rho_{0m\tau}(z_{m\tau})\right)$ to account for the potential suboptimality of fixing $\tilde{d}_{0m\tau} = 1$.[6]

[5] To address concerns regarding convergence to local maxima, we initialize the EM algorithm using multiple random starting values $(\beta^{(0)}, \pi^{(0)}, p^{(0)})$.

[6] Setting $\tilde{d}_{0m\tau} = 1$ is arbitrary but convenient. Similar expressions can be derived using any other sequence of choices.

Equation (D8) has two key advantages. First, expected payoffs can be easily approximated via Monte Carlo integration. Furthermore, note that κ only depends on the state transition law, g, which in turn only depends on β and π, not on the parameters that are to be estimated in the second stage. Thus, sample paths used to simulate $v_{jmt}(z_{mt};\varphi)$, as explained below, can be drawn once prior to the GMM optimization and remain fixed throughout. This considerably reduces the computational burden of estimation.

Given $j \in \mathcal{J}$, $\xi_{mt} = \xi$, and (θ, β_λ), we simulate $v_{jmt}(I_{mt}, \xi; \varphi)$ as follows. For sample path $l = 1, \ldots, L$, we first draw election outcome $w_{mt}^{(l)} = (w_{mt}^{p(l)})_{p=1}^{P}$ according to $f_{jmt}^{p}(I_{mt}, \xi; \hat{\beta})$.[7] Then, drawing $\xi_{m,t+1}^{(l)}$ from $\pi_{m,t+1}$, we obtain state $z_{m,t+1}^{(l)} = (I_{m,t+1}^{(l)}, \xi_{m,t+1}^{(l)})$, where $I_{m,t+1}^{p(l)} = (w_{m,t-N+1}^{p}, \ldots, w_{m,t-1}^{p}, w_{mt}^{p(l)})$. Next, election outcome $w_{m,t+1}^{(l)}$ is drawn according to $f_{0m,t+1}^{p}(z_{m,t+1}^{(l)}; \hat{\beta})$, and $\xi_{m,t+2}^{(l)}$ is drawn from $\pi_{m,t+2}$, which yields $z_{m,t+2}^{(l)} = (I_{m,t+2}^{(l)}, \xi_{m,t+2}^{(l)})$, where $I_{m,t+2}^{p(l)} = (w_{m,t-N+2}^{p}, \ldots, w_{m,t-1}^{p}, w_{mt}^{p(l)}, w_{m,t+1}^{p(l)})$. Carrying on recursively for $\tau = t+3, \ldots, T^*$, drawing election outcomes according to $f_{0m,\tau-1}^{p}(z_{m,\tau-1}^{(l)}; \hat{\beta})$, we obtain the sequence $\{z_{m\tau}^{(l)}\}_{\tau=t+1}^{T^*}$, where T^* is a finite approximation to the infinite-horizon problem chosen so that future payoff differences become negligible given the discount factor, δ.[8] Thus, $v_{jmt}(I_{mt}, \xi; \varphi)$ can be approximated by

$$\hat{v}_{jmt}(I_{mt}, \xi; \theta, \beta_\lambda) = \frac{1}{L} \sum_{l=1}^{L} \hat{v}_{jmt}^{(l)}(I_{mt}, \xi; \theta, \beta_\lambda),$$

where

$$\begin{aligned}\hat{v}_{jmt}^{(l)}(I_{mt}, \xi; \theta, \beta_\lambda) &= \sum_{p \in C_j} \left[\lambda_{s(m)}(\beta_\lambda)\theta_{\text{PAN},p} + (1 - \lambda_{s(m)}(\beta_\lambda))\theta_{\text{PRD},p}\right] f_{jmt}^{p}(I_{mt}, \xi; \hat{\beta}) \\ &+ \sum_{\tau=t+1}^{T^*} \delta^{\tau-t} \left\{ \sum_{p \in C_0} \left[\lambda_{s(m)}(\beta_\lambda)\theta_{\text{PAN},p} + (1 - \lambda_{s(m)}(\beta_\lambda))\theta_{\text{PRD},p}\right] f_{0m\tau}^{p}(z_{m\tau}^{(l)}; \hat{\beta}) - \log\left(\hat{\rho}_{0m\tau}(z_{m\tau}^{(l)})\right) \right\}\end{aligned}$$

and

$$\lambda_s(\beta_\lambda) = \frac{\exp\left(\beta_\lambda \sum_{m \in M_s} \frac{\iota(I_{m1}^{\text{PAN}},1) - \iota(I_{m1}^{\text{PRD}},1)}{|M_s|}\right)}{1 + \exp\left(\beta_\lambda \sum_{m \in M_s} \frac{\iota(I_{m1}^{\text{PAN}},1) - \iota(I_{m1}^{\text{PRD}},1)}{|M_s|}\right)}.$$

To prevent simulation error from propagating, we draw independent sets of sample paths for each municipality m and electoral cycle t.

Our GMM estimator of (θ, β_λ) is based on the following sample analogs of the identifying moment conditions obtained from (D7). Let $G(\theta, \beta_\lambda)$ be the $2|\Xi| \times 1$ vector of moments defined

[7] In our implementation, we set $L = 200$.

[8] In our implementation, we set $\delta = 0.9$ and $T^* = 15$ (45 years). Our results are robust to alternative choices of the discount factor—see Online Appendix E.

by

$$G(\beta_\lambda, \theta) = \begin{pmatrix} \bar{p}_{-1}(\xi) - \bar{p}_0(\xi) - [\bar{v}_{-1}(\xi; \theta, \beta_\lambda) - \bar{v}_0(\xi; \theta, \beta_\lambda)] \\ \bar{p}_1(\xi) - \bar{p}_0(\xi) - [\bar{v}_1(\xi; \theta, \beta_\lambda) - \bar{v}_0(\xi; \theta, \beta_\lambda)] \end{pmatrix}_{\xi \in \Xi},$$

where

$$\bar{p}_j(\xi) - \bar{p}_0(\xi) = \frac{1}{MT} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \left[\log \left(\hat{\rho}_{jmt}(I_{mt}, \xi) \right) - \log \left(\hat{\rho}_{0mt}(I_{mt}, \xi) \right) \right],$$

$$\bar{v}_j(\xi; \theta, \beta_\lambda) - \bar{v}_0(\xi; \theta, \beta_\lambda) = \frac{1}{MT} \sum_{m=1}^{M} \sum_{t=1}^{T_m} \left[\hat{v}_{jmt}(I_{mt}, \xi; \theta, \beta_\lambda) - \hat{v}_{0mt}(I_{mt}, \xi; \theta, \beta_\lambda) \right],$$

and $T = \sum_{m=1}^{M} T_m$. Then, given (D7), we obtain our estimates $(\hat{\theta}, \hat{\beta}_\lambda)$ by minimizing the quadratic form

$$\min_{\theta,\, \beta_\lambda} G(\theta, \beta_\lambda)' G(\theta, \beta_\lambda). \tag{D9}$$

Since the conditional value functions, $\hat{v}_{jmt}(I_{mt}, \xi; \theta, \beta_\lambda)$, are linear in θ, it is straightforward to solve explicitly for $\hat{\theta}(\beta_\lambda)$ from the least-squares first-order conditions, simplifying the GMM optimization to

$$\min_{\beta_\lambda} G(\hat{\theta}(\beta_\lambda), \beta_\lambda)' G(\hat{\theta}(\beta_\lambda), \beta_\lambda).$$

E Alternative Model Specifications

To explore the robustness of our main results to key modeling and sample choices, we present in Table E1 estimates of coefficients $\beta = ((\beta_x^p)_{p\in\{1,\ldots,P\}}, \beta_I, \alpha)$ in Equation (1) in the paper, which governs parties' electoral prospects, from four alternative model specifications. Results in column (I) are obtained using an expanded sample covering the period 1999-2016. In column (II), using our baseline sample, we reestimate our model setting $N = 4$. In column (III), we restrict unobserved valence so that $\xi^p \in \{0, 1\}$. In column (IV), we add as a control in x_{mt} the party affiliation of the incumbent state governor. Finally, in column (V), we control for whether the party formed a coalition in the previous election—either with its own candidate (i.e., senior partner) or by supporting another candidate (i.e., junior partner)—and for whether the incumbent mayor was supported by the party as their junior coalition partner. These specifications are otherwise identical to our baseline model. As shown in Table E1, our main results are virtually unchanged. Notably, column (V) indicates that there are no direct electoral benefits from being a junior coalition partner, which again underscores the importance of dynamic incentives for understanding the PAN-PRD alliance.

Furthermore, Table E2 shows that our estimates of parties' payoffs, θ, are robust to alternative choices of the coalition's discount factor, δ. Crucially, both PAN and PRD would prefer PRI to be in power instead of each other regardless of the choice of δ. The last row of Table E2 reports the minimized value of the GMM criterion in (D9), normalized by the sample size. According to this statistic, setting $\delta \geq 0.9$ fits the data best, which indicates that the coalition partners indeed are forward-looking. (In our main specification, we set $\delta = 0.95$.)

Table E1: Robustness to Alternative Sample and Model Specifications

		(I)	(II)	(III)	(IV)	(V)
PRD:	conc_Governor	-0.059	-0.069	-0.056	-0.047	-0.052
		(0.102)	(0.142)	(0.113)	(0.124)	(0.150)
	conc_Congress	-1.147	-0.541	-0.583	-0.808	-0.819
		(0.139)	(0.161)	(0.159)	(0.154)	(0.167)
	conc_President	0.723	0.424	0.404	0.323	0.386
		(0.147)	(0.187)	(0.160)	(0.200)	(0.176)
	rural	0.245	0.286	0.284	0.190	0.186
		(0.112)	(0.157)	(0.139)	(0.164)	(0.163)
	pop_60	0.023	0.054	0.054	0.057	0.061
		(0.014)	(0.016)	(0.014)	(0.017)	(0.017)
	pop_female	0.025	0.037	0.029	0.061	0.074
		(0.040)	(0.058)	(0.055)	(0.065)	(0.059)
	poverty	0.002	0.001	0.001	0.001	0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PRI:	conc_Governor	-0.367	-0.247	-0.247	-0.264	-0.271
		(0.084)	(0.117)	(0.095)	(0.117)	(0.115)
	conc_Congress	-0.998	-0.840	-0.841	-0.875	-0.808
		(0.116)	(0.129)	(0.140)	(0.136)	(0.139)
	conc_President	0.590	0.796	0.772	0.842	0.818
		(0.113)	(0.146)	(0.147)	(0.195)	(0.129)
	rural	-0.069	0.009	-0.015	0.021	0.043
		(0.105)	(0.131)	(0.123)	(0.139)	(0.127)
	pop_60	0.042	0.032	0.033	0.034	0.032
		(0.011)	(0.014)	(0.014)	(0.013)	(0.011)
	pop_female	-0.116	-0.023	-0.019	-0.005	-0.022
		(0.031)	(0.048)	(0.043)	(0.050)	(0.046)
	poverty	0.002	0.002	0.002	0.002	0.002
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
PAN:	conc_Governor	0.002	-0.023	-0.041	-0.009	-0.029
		(0.080)	(0.128)	(0.116)	(0.131)	(0.139)
	conc_Congress	-0.917	-1.149	-1.147	-1.570	-1.375
		(0.112)	(0.161)	(0.154)	(0.168)	(0.143)
	conc_President	0.762	0.593	0.621	0.446	0.637
		(0.119)	(0.185)	(0.168)	(0.199)	(0.163)
	rural	-0.112	-0.010	0.010	-0.017	-0.005
		(0.116)	(0.147)	(0.129)	(0.146)	(0.126)
	pop_60	0.041	0.051	0.053	0.061	0.059
		(0.012)	(0.015)	(0.015)	(0.015)	(0.013)
	pop_female	-0.040	-0.065	-0.070	-0.050	-0.047
		(0.030)	(0.055)	(0.049)	(0.057)	(0.052)
	poverty	0.003	0.002	0.002	0.002	0.002
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
β_I		0.371	0.244	0.213	0.248	0.275
		(0.030)	(0.022)	(0.025)	(0.033)	(0.028)
α		0.807	1.000	0.967	0.961	0.955
		(0.039)	(0.032)	(0.039)	(0.056)	(0.039)
Governor					0.410	
					(0.056)	
(Lagged) Senior Coalition Partner						-0.112
						(0.138)
(Lagged) Junior Coalition Partner						-1.018
						(0.273)
Junior-Partner Incumbent						-0.326
						(0.318)
Region F.E.		Yes	Yes	Yes	Yes	Yes
Observations		9917	4059	4059	4059	4059

Notes. Estimates correspond to coefficients $\beta = ((\beta_x^p)_{p\in\{1,\ldots,P\}}, \beta_I, \alpha)$ in Equation (1) in the paper, which governs parties' electoral prospects. The omitted category is $p =$ OTHER. Nonparametrically bootstrapped standard errors are reported in parentheses. Estimates in column (I) use our expanded sample covering 1999-2016. In column (II), we set $N = 4$. In column (III), we restrict unobserved valence to $\xi^p \in \{0, 1\}$. In column (IV), we control for whether the incumbent governor is from the same party. In column (V), we control for whether the party formed a coalition in the previous election—either with its own candidate (senior partner) or by supporting another candidate (junior partner)—and for whether the incumbent mayor was supported by the party as their junior coalition partner.

Table E2: Robustness to Choice of Coalition's Discount Factor

	$\delta = 0.75$	$\delta = 0.80$	$\delta = 0.85$	$\delta = 0.90$	$\delta = 0.95$
$\theta_{\text{PAN,PRD}}$	-485.107	-479.961	-474.216	-432.279	-433.111
$\theta_{\text{PRD,PRI}}$	-463.574	-453.454	-441.896	-418.891	-405.716
$\theta_{\text{PRD,PAN}}$	-991.015	-977.127	-960.987	-914.052	-897.102
(Normalized) GMM Criterion	32.588	32.695	32.836	32.367	32.526

Notes. Estimates correspond to coefficients $\theta_{pp'}$ measuring the payoff party p derives whenever party p' is in power. Recall that $\theta_{pp} = 0$ and $\theta_{\text{PAN,PRI}} = \theta_{\text{PRD,PRI}}$. The last row corresponds to the minimized value of the GMM criterion defined by (D9), normalized by the sample size.

F Unconventional Coalitions in Brazil

Lastly, we provide descriptive evidence from electoral coalitions in Brazilian municipalities similar to the Mexican case. Two of the three largest parties in Brazil, PT (left) and PSDB (center-right), have been the main rivals in national politics for the past few decades. Their candidates finished in the top two in all six presidential elections between 1994-2014 as well as in many gubernatorial races. They are also among the most programmatic parties in the country. Similar to the case of Mexico, the literature has established that voters and politicians alike clearly recognize PT and PSDB's opposing ideological positions and that these positions are relevant for voters' choices (Desai and Frey, 2021; Samuels and Zucco Jr., 2014; Power and Zucco Jr., 2009).

Despite not having formed a national alliance since 1989, PT and PSDB often form coalitions in municipal elections, supporting each other's mayoral candidates. What is more, we show that these coalitions are much more likely to occur in municipalities where neither party has been successful in the past or where a large centrist party has a strong presence (similar to the PAN-PRD case).

To make the analysis as comparable as possible to the Mexican case, we focus on PT-PSDB local coalitions and past incumbency of the largest centrist party in Brazil, MDB.[9] We focus on the 2016 mayoral elections for which we can observe incumbency status in each municipality for the previous four mayoral tenures (our sample starts with the 2000 electoral cycle, and mayors are elected for four-year terms).

In 666 municipalities, we observe a candidate from MDB running alongside a candidate from at least one of the two relevant parties, PT and PSDB. Our outcome variable is a binary indicator of whether or not the MDB candidate faced a PT-PSDB coalition candidate. Coalition candidates can be observed in 9% of these municipalities. We then regress our dependent variable on the following measures of past incumbency: (i) the number of previous administrations in which the mayor was *not* from PT or PSDB, (ii) a binary indicator of whether the municipality did not have a PT or PSDB mayor in any of the past four terms, (iii) the number of previous administrations in which the mayor was from MDB, and (iv) a binary indicator of whether the municipality had an MDB mayor in at least one of the past four terms.

Table F1 presents our results. There is a robust association between past incumbency and the probability of a PT-PSDB coalition: the coalition is much more likely to occur where MDB has a strong grip on power or where PT and PSDB have not held office. For example,

[9] MDB's ideological placement is to the right of PT and slightly to the left of PSDB (Power and Zucco Jr., 2009). The party has employed an extensive patronage machine to elect the largest number of mayors in this period. At the federal level, MDB has always joined the governing coalition led by either PT or PSDB.

column (III) indicates that the probability of a PT-PSDB coalition is nearly 7 times higher in locations where MDB has been in power for four consecutive terms than in locations where MDB has never been the incumbent. Overall, this exercise provides a suggestive example of ends-against-the-middle coalitions in another large, developing democracy consistent with the mechanism elucidated by our model.

Table F1: Probability of a PT-PSDB Coalition and Past Incumbency

	(I)	(II)	(III)	(IV)
Effect on coalition probability	0.024**	0.055**	0.049**	0.070**
	(0.010)	(0.025)	(0.013)	(0.020)
Baseline coalition probability	0.114	0.122	0.035	0.041
Observations	666	666	666	666

Notes. $^{**}p < 0.05$. Standard errors are heteroskedasticity robust and presented in parenthesis. The explanatory variable in each column is: (I) number of previous municipal administrations in which the mayor was *not* from PT or PSDB, (II) a binary indicator of whether the municipality did not have a PT or PSDB mayor in any of the past four terms, (III) number of past municipal administrations with an MDB mayor, (IV) a binary indicator of whether the municipality had an MDB mayor in at least one of the past four terms.

References

Arcidiacono, Peter and Robert A. Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity." *Econometrica* 79(6):1823–1867.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The statistical analysis of roll call data." *American Political Science Review* 98(2):355–370.

Consulta Mitofsky. 2012. "*Geometría electoral en México* (Electoral geometry in Mexico)." Public Opinion Poll, http://www.consulta.mx/web/images/MexicoOpina/2013/NA_GEOMETRIA_ELECTORAL.pdf (in Spanish).

Desai, Zuheir and Anderson Frey. 2021. "Can Descriptive Representation Help The Right Win Votes From The Poor? Evidence From Brazil." *American Journal of Political Science* (forthcoming).

Power, Timothy J. and Cesar Zucco Jr. 2009. "Estimating Ideology of Brazilian Legislative Parties, 1990-2005: A Research Communication." *Latin American Research Review* 44(1):218–246.

Samuels, David and Cesar Zucco Jr. 2014. "The Power of Partisanship in Brazil: Evidence from Survey Experiments." *American Journal of Political Science* 58(1):212–225.