Supplementary Materials:
Does Political Representation Increase Participation? Evidence from Party Candidate Lotteries in Mexico

Mathias Poertner
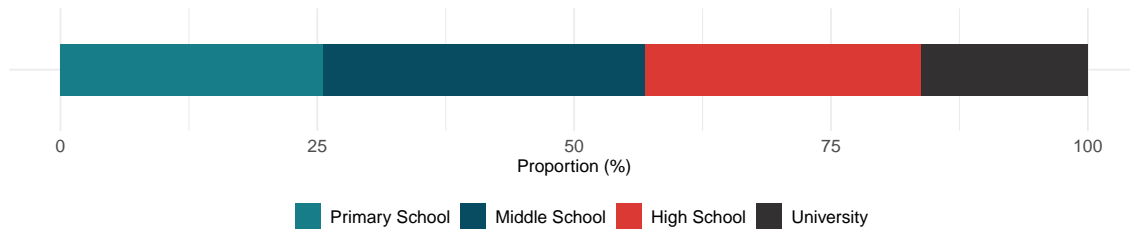
Table of Contents

S1. Constituents' Characteristics: Full National Sample

Figure A1. **Educational Background of Mexican Adult Population**

Note: Data from 2017 LAPOP survey (nationally representative sample).

Figure A2. **Occupational Background of Mexican Adult Population**
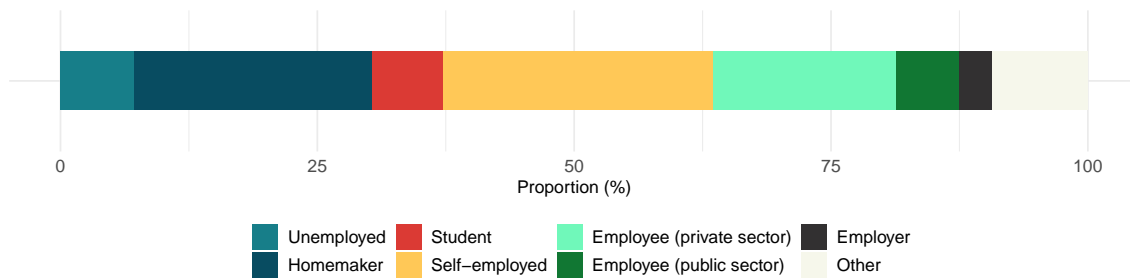
Note: Data from 2017 LAPOP survey (nationally representative sample).

S2. Educational Background of SMD Deputies

Figure A3. **Educational Background of SMD Deputies by Party**
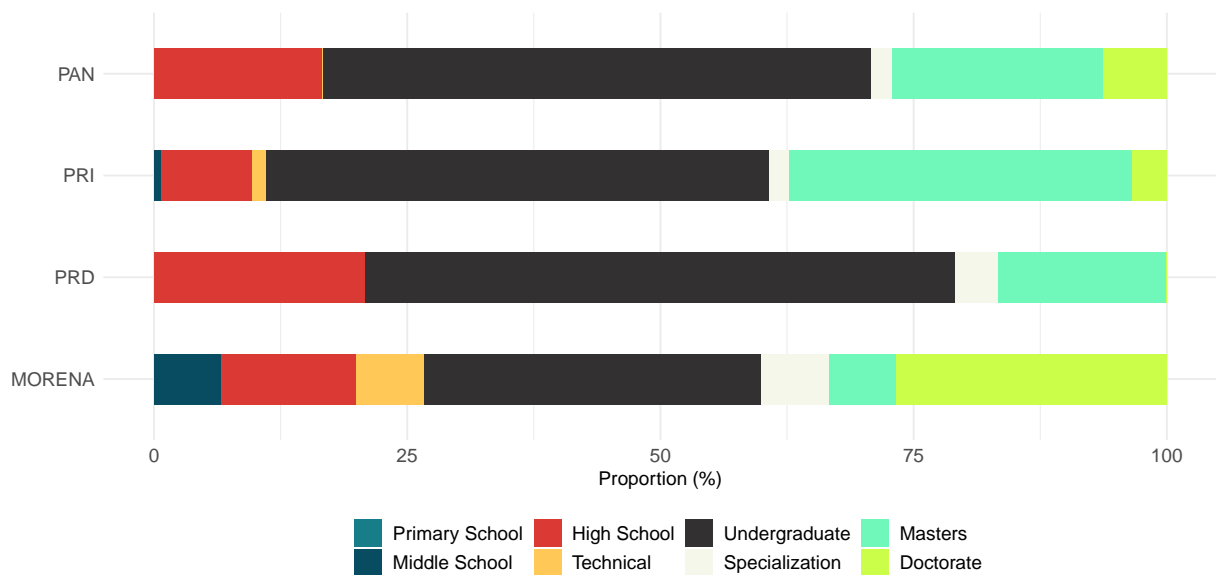
Note: Coded based on the information reported in the congressional biographies (LXIII Legislature; 2015–2018).

S3. Occupational Background of PR Candidates

Figure A4. **Occupational Background of PR Candidates by Party**
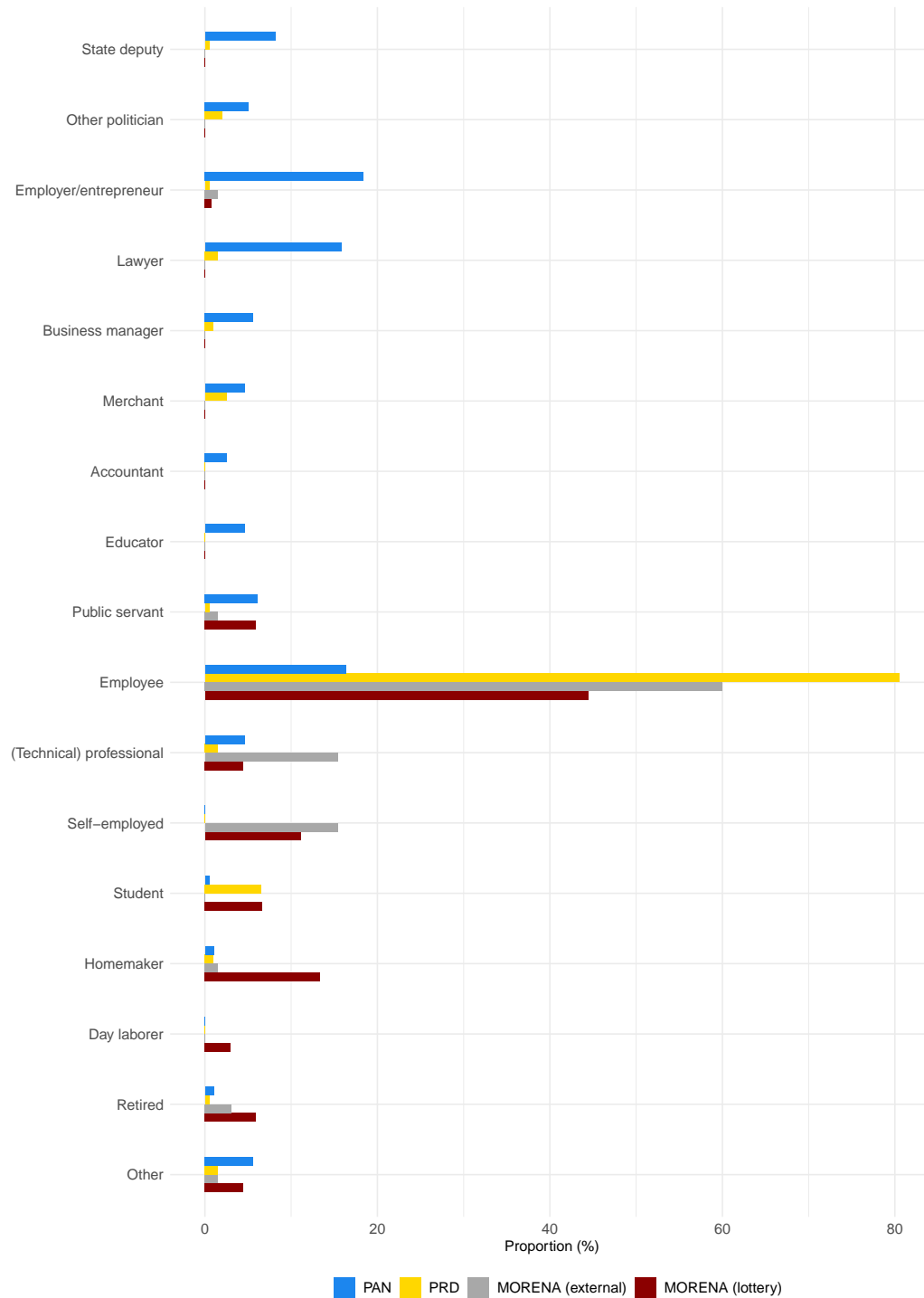
Note: Coded based on the information provided on the candidate registration forms (LXIII Legislature; 2015–2018).

S4. Occupational Background of SMD Candidates and Deputies

Figure A5. **Occupational Background of SMD Candidates by Party**
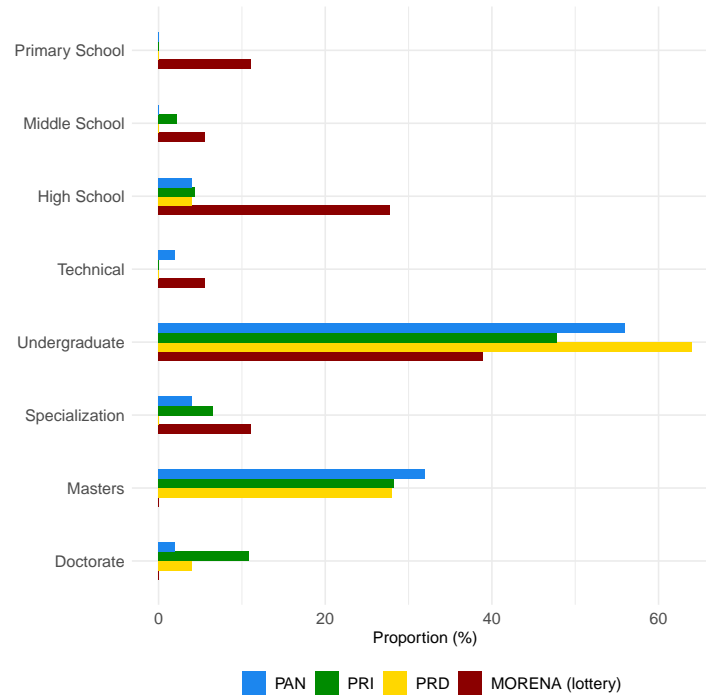
Note: Coded based on the information provided on the candidate registration forms (LXIII Legislature; 2015–2018).

Figure A6. **Occupational Background of SMD Candidates by Party**

Note: Coded based on the information provided on the candidate registration forms (LXIII Legislature; 2015–2018).

Figure A7. **Occupational Background of SMD Deputies by Party**
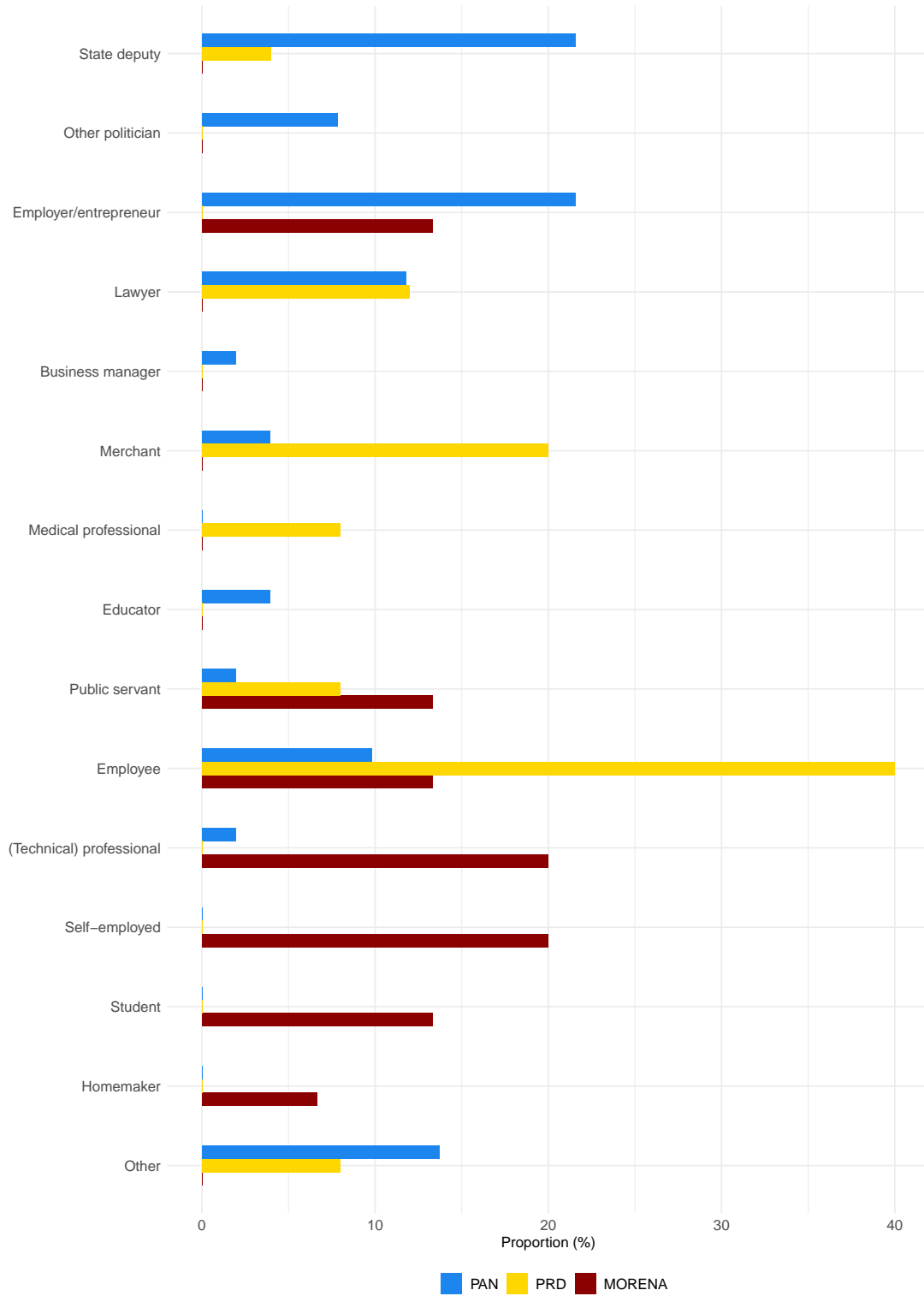
Note: Coded based on the information provided on the candidate registration forms (LXIII Legislature; 2015–2018).

Figure A8. **Occupational Background of SMD Deputies by Party**
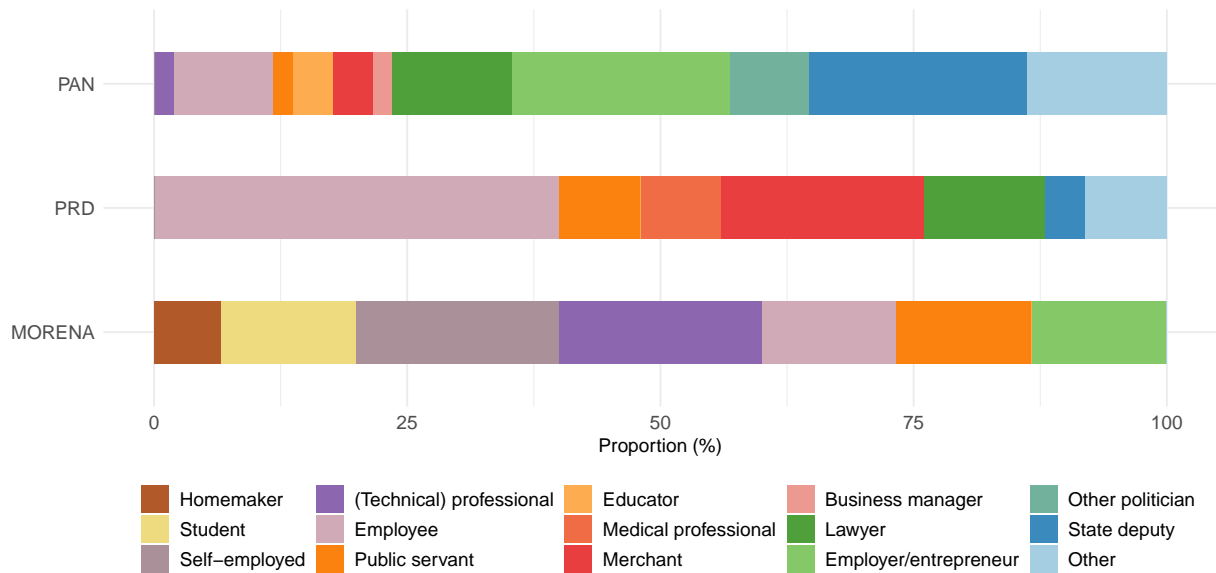
Note: Coded based on the information provided on the candidate registration forms (LXIII Legislature; 2015–2018).

S5. Balance Tests for Candidates on Party Lists

Table A1. **Balance Statistics for Candidates on Party Lists**

Variable	Not elected	Elected	Diff. of means	p-value	n
Female	0.5128	0.5000	-0.0128	0.9222	135
Years of Residence	4.0983	4.7500	0.6517	0.4773	135
Age	45.6496	44.3889	-1.2607	0.7084	135

Note: The table entries are estimated through t tests comparing characteristics of candidates who ended up in office to those who did not. Years of Residence refers to years residence in the municipality.

S6. Balance Tests for Municipalities with Fixed Effects

Table A2. **Balance Statistics for Municipalities**

Outcome	Effect of Represent.	SE	p-value	n
Population Ages 0-29	-0.0025	0.0733	0.9751	100
Population Ages 30-49	-0.0143	0.0378	0.7313	100
Population Ages 50+	0.0167	0.1103	0.8890	100
Primary Sector Workers	-0.0204	0.0704	0.7914	100
Industrial Workers	-0.0007	0.0077	0.9369	100
Comercial Sector Workers	0.0194	0.0131	0.2346	100
Service Sector Workers	0.0210	0.0567	0.7361	100
Income Vulnerable Population	0.0132	0.0070	0.1575	100
Turnout (Baseline)	0.0068	0.0042	0.2057	100
Vote Share for PRD (Baseline)	0.0080	0.0077	0.3754	100
Vote Share for PRI (Baseline)	-0.0067	0.0231	0.7897	100
Vote Share for PAN (Baseline)	-0.0225	0.0196	0.3358	100
PRD Mayor	0.0646	0.0560	0.3324	102

Note: Balance statistics for municipality characteristics with fixed effects (FE) for constituency and the number of candidates from a municipality on the list, comparing comparing electoral returns (for 2012), mayor party in office at baseline (beginning of 2015), and other municipality characteristics (2015) between municipalities with a MORENA lottery candidate who was elected (in 2015) (treatment) to those with a lottery candidate who was not elected (control). The vote share for the PRI also includes support for the PVEM and the vote share for the PRD also includes support for the PT and Movimiento Ciudadano since they ran in electoral alliances in 2012. Standard errors are clustered by state and number of lottery candidates (for "PRD Mayor," constituency and number of lottery candidates due to model constraints); p-values are two-tailed.

S7. Effect of Representation on Vote Shares for Other Parties

Table A3. **Effect of Representation on Vote Shares for Other Parties**

Outcome	Control	Treatment	Effect of Represent.	SE	*p*-value	n
Vote Share for PRD	0.2891	0.2419	-0.0472	0.0311	0.0663	102
Vote Share for PRI	0.2277	0.2163	-0.0114	0.0024	0.0000	102

Note: Estimates of effect of representation on electoral support using IPW, comparing electoral returns (in the 2018 election) from municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) to those with a lottery candidate who was not elected (control). The vote share for the PRI also includes support for the PVEM and Nueva Alianza and the vote share for the PRD also includes support for the PAN and Movimiento Ciudadano since they ran in electoral alliances in 2018. Standard errors are clustered by constituency and number of lottery candidates; *p*-values are one-tailed to account for directional hypothesis expecting that other parties would lose vote support.

S8. Alternative Specifications for Main Outcomes

Table A4. **Effect of Representation on Turnout & Party Vote Shares (FE Models)**

Outcome	Effect of Represent.	SE	*p*-value	n
Turnout	0.0151	0.0051	0.0293	102
Vote Share for MORENA	0.0349	0.0138	0.0429	102
Vote Share for PRD	-0.0406	0.0227	0.0857	102
Vote Share for PRI	0.0064	0.0250	0.3990	102

Note: Estimates of effect of representation on electoral participation and support with fixed effects (FE) for constituency and the number of candidates from a municipality on the list, comparing electoral returns (in the 2018 election) from municipalities with a MORENA lottery candidate who was elected (in 2015) (treatment) to those with a lottery candidate who was not elected (control). Standard errors are clustered by state and number of lottery candidates (for PRI, only by state due to model constraints); *p*-values are one-tailed to account for directional hypotheses.

Table A5. **Effect of Representation on Turnout & Party Vote Shares (Boot-strapped Estimates)**

	Turnout	Vote Share for MORENA	Vote Share for PRD	Vote Share for PRI
	(1)	(2)	(3)	(4)
Represented	0.054***	0.069*	-0.070**	0.002
	(0.014)	(0.041)	(0.028)	(0.02)
Registered Voters	-0.007***	-0.002	0.006*	-0.006***
	(0.003)	(0.004)	(0.004)	(0.002)
Constant	0.646***	0.473***	0.278***	0.242***
	(0.011)	(0.018)	(0.016)	(0.010)

Note: (Non-parametric) bootstrapped estimates of effect of representation on electoral participation and support (controlling for the number of registered voters in a given municipality prior to the lottery (per 100,000 voters)) based on 10,000 replications each (two-tailed p-values), comparing electoral returns (in the 2018 election) from municipalities with a MORENA lottery candidate who was elected (in 2015) (treatment) to those with a lottery candidate who was not elected (control).

Table A6. **Effect of Representation on Turnout & Party Vote Shares (RI Estimates)**

Outcome	ATE	p-value
Turnout	0.0537	0.0108
Vote Share for MORENA	0.0627	0.0977
Vote Share for PRD	-0.0702	0.0376
Vote Share for PRI	0.0085	0.6621

Note: Estimates of effect of representation on electoral participation and support (adjusting for the number of registered voters in a given municipality prior to the lottery) using randomization inference (two-tailed p-values, based on 50,000 simulations), comparing electoral returns (in the 2018 election) from municipalities with a MORENA lottery candidate who was elected (in 2015) (treatment) to those with a lottery candidate who was not elected (control).

S9. Difference-in-Differences Models of Main Outcomes

Table A7. **Effect of Representation on Turnout & Party Vote Shares (DiD Models)**

	Turnout	MORENA Vote Share	PRI Vote Share	PRD/PAN Vote Share
	(1)	(2)	(3)	(4)
Constant	0.625***	0.226***	0.258***	0.484***
	(0.010)	(0.026)	(0.018)	(0.029)
Represented	0.006***	0.012	−0.010	0.002
	(0.002)	(0.014)	(0.009)	(0.021)
Time	0.005	0.244***	−0.030**	−0.195***
	(0.007)	(0.013)	(0.014)	(0.044)
Represented x Time	0.013**	0.039***	−0.001	−0.049
	(0.006)	(0.014)	(0.007)	(0.055)
Observations	202	202	202	202

Note: Estimates of difference-in-differences in electoral participation and support using IPW, comparing electoral returns from 2018 to last election prior to the lottery (previous midterm election) in 2012 from municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) with those in municipalities with a lottery candidate who was not elected (control). Since MORENA did not yet contest the 2012 election, electoral support for the PRD is used for this election. Since the PRD and PAN ran in an electoral coalition in 2018 and their individual vote shares are not available, their vote shares are added for the 2012 election. Standard errors are clustered by constituency and number of of lottery candidates. *p<0.1; **p<0.05; ***p<0.01, two-tailed tests.

S10. Effect of Representation on Voting Behavior by Whether (Also) Represented by Other Deputy

Table A8. **Effect of Representation on Voting Behavior by Whether (Also) Represented by Other Deputy**

Outcome	Other MP	Effect of Represent.	SE	*p*-value	n
Vote Share for MORENA	No	0.0216	0.0260	0.2043	75
Vote Share for MORENA	Yes	0.0911	0.0348	0.0074	27
Vote Share for MORENA	Difference	0.0695	0.0368	0.0310	102
Vote Share for PRD	No	-0.0477	0.0206	0.0117	75
Vote Share for PRD	Yes	-0.0739	0.0462	0.0612	27
Vote Share for PRD	Difference	-0.0262	0.0402	0.2581	102
Vote Share for PRI	No	0.0356	0.0173	0.0217	75
Vote Share for PRI	Yes	-0.0226	0.0170	0.0971	27
Vote Share for PRI	Difference	-0.0582	0.0335	0.0426	102

Note: Estimates of effect of representation on vote shares using IPW, comparing electoral returns (in the 2018 election) from municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) to those with a lottery who was not elected (control). Standard errors are clustered by constituency and number of lottery candidates; *p*-values are one-tailed to account for directional hypotheses. *Other MP* indicates whether municipalities in the sample are (also) represented by another federal deputy from that municipality and or whether an estimate for difference in ATEs between the two groups is presented.

S11. Calculations: MORENA's Proportion of Turnout Increase

MORENA's increased electoral support in areas with a lottery deputy could theoretically be the result of mobilization of constituents who are eligible to vote but who would not have turned out to vote absent lottery deputy representation (*new voters*) and/or the winning over of voters who otherwise would have voted for another party (*swayed voters*); most importantly for the PRD (p), the closest programmatic and organizational competitor of MORENA (m).[1]

$$VoteIncrease_m = NewVoters_m + SwayedVoters_m \tag{1}$$

The increase in turnout that went to MORENA could then be defined as the electoral support gained by MORENA minus the electoral support it swayed from the PRD.

$$TurnoutIncrease_m = VoteIncrease_m - SwayedVoters_m \tag{2}$$

If we suppose that—in the most extreme case—all voters lost by the PRD moved to MORENA, the increase in turnout that went to MORENA would then be:

$$TurnoutIncrease_m = VoteIncrease_m - VoteDecrease_p \tag{3}$$

The proportion of the increase in turnout that went to MORENA would then simply be the electoral support gained by MORENA minus the electoral support it swayed from the PRD divided by the overall increase in turnout:

$$\frac{VoteIncrease_m - VoteDecrease_p}{TurnoutIncrease_{overall}} \tag{4}$$

---

[1] I am focusing on the PRD here because most models find no significant effect of lottery representation on the vote share for the PRI (see Tables A4–A7).

Or to be more precise:

$$= \frac{\overline{v}_m(1)\overline{t}(1) - \overline{v}_m(0)\overline{t}(0) - [\overline{v}_p(0)\overline{t}(0) - \overline{v}_p(1)\overline{t}(1)]}{\overline{t}(1) - \overline{t}(0)} \tag{5}$$

In this case, $\overline{v}_j$ represents the mean vote share for party j and $\overline{t}$ represents the mean turnout rate. j can take on two values: m for MORENA and p for the PRD. $\overline{v}_j(1)$ and $\overline{t}(1)$ give the values of these outcomes given representation by a lottery deputy (treatment), and $\overline{v}_j(0)$ and $\overline{t}(1)$ are the vote shares/turnout rates in areas without representation by a lottery deputy (control).

Using the main estimates for the parties' mean vote shares and the mean turnout rates reported in the paper, we can then simply solve the equation:

$$\frac{52.17\% \cdot 64.93\% - 47.07\% \cdot 63.03\% - [28.91\% \cdot 63.03\% - 24.19\% \cdot 64.93\%]}{64.93\% - 63.03\%} \approx 88.97\% \tag{6}$$

Therefore, the proportion of the increase in turnout that went to MORENA is at least 88.97%. To the extent that not all voters lost by the PRD moved to MORENA but rather stayed at home or voted for a third party, this proportion would be even larger.

S12. Alternative Specifications for Allocation of Federal Transfers

Table A9. **Effect of Representation on Allocation of Federal Transfers (Per Capita) (Cross-Section)**

Outcome	Control	Treatment	Effect of Represent.	SE	p-value	n
Transfers 2015 (Per Capita)	1532	1729	197	161	0.2256	96
Transfers 2016 (Per Capita)	1689	1790	100	269	0.7108	96
Transfers 2017 (Per Capita)	1826	1989	163	278	0.5604	96
Transfers 2018 (Per Capita)	2103	2162	58	324	0.8573	96

Note: Estimates of effect of representation on federal transfers (in Mexican pesos) to federal entities and municipalities (Ramo 28) using IPW, comparing per capita transfers to municipalities with a lottery candidate who was elected in 2015 (treatment) to those with a lottery candidate who was not elected (control). Standard errors are clustered by constituency; p-values are two-tailed.

S13. Placebo Tests: Political Attitudes (LAPOP)

Table A10. **Effect of Representation on Unrelated Outcomes (LAPOP)**

	Outcome:					
	Indigenous	Female	Working	Leftist Ideology (Self-placement)	Leftist Policy Preferences	Support for Democracy
	(1)	(2)	(3)	(4)	(5)	(6)
Represented	−0.010	0.003	−0.028	0.039	0.024	0.132
	(0.016)	(0.015)	(0.042)	(0.048)	(0.025)	(0.126)
Constant	0.093***	0.498***	0.505***	0.620***	0.587***	4.284***
	(0.017)	(0.013)	(0.021)	(0.046)	(0.011)	(0.123)
Clusters	27	27	27	27	27	27
Observations	421	506	504	453	498	479

Note: Estimates of effect of representation on survey outcomes using IPW, comparing respondents in municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) with those in municipalities with a lottery candidate who was not elected (control). "Indigenous," "Female," "Working," "Leftist Ideology (Self-placement)," and "Leftist Policy Preferences" are binary variables indicating that the characteristic applies; "Support for Democracy" is coded on a 7-point scale with higher values indicating more support for democracy. Standard errors are clustered by municipality and number of of lottery candidates. *p<0.1; **p<0.05; ***p<0.01, one-tailed tests.

S14. Effect of Representation on Political Attitudes by Education Level & Past Voting Status

Table A11. **Effect of Representation on Political Attitudes by Education Level**

	Outcome:		
	Respect for Pol. Institutions	Political Efficacy	Political Interest
	(1)	(2)	(3)
Represented	0.083	−0.104	−0.053
	(0.405)	(0.135)	(0.057)
No Postsec. Education	0.328	0.717**	−0.270***
	(0.334)	(0.379)	(0.050)
Represented x No Postsec. Education	0.298	0.274*	0.138*
	(0.420)	(0.182)	(0.085)
Constant	3.724***	2.441***	0.582***
	(0.262)	(0.298)	(0.019)
Clusters	27	27	27
Observations	494	492	500

Note: Estimates of effect of representation on survey outcomes using IPW, comparing respondents in municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) with those in municipalities with a lottery candidate who was not elected (control). Standard errors are clustered by municipality and number of lottery candidates. *p<0.1; **p<0.05; ***p<0.01, one-tailed tests to account for directional hypothesis.

Table A12. **Effect of Representation on Political Attitudes by Past Voting Status**

	Outcome:		
	Respect for Pol. Institutions	Political Efficacy	Political Interest
	(1)	(2)	(3)
Represented	0.143***	−0.119	−0.020
	(0.023)	(0.132)	(0.044)
Non-Voter (in 2012)	−0.263	−0.269*	−0.080
	(0.209)	(0.177)	(0.080)
Represented x Non-Voter (in 2012)	0.536***	0.741***	0.190***
	(0.198)	(0.230)	(0.035)
Constant	4.075***	3.116***	0.383***
	(0.087)	(0.105)	(0.048)
Clusters	27	27	27
Observations	495	492	500

Note: Estimates of survey outcomes using IPW, comparing respondents in municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) with those in municipalities with a lottery candidate who was not elected (control). Standard errors are clustered by municipality and number of lottery candidates. *p<0.1; **p<0.05; ***p<0.01, one-tailed tests to account for directional hypothesis.

S15. Google Trends Data

S15.1. Collection of Google Trends Data

Google Trends data offers a representative sample of search requests made to Google that "allows us to display interest in a particular topic from around the globe or down to city-level geography" (Rogers 2016).[2] In recent years, these data have been analyzed increasingly in public health and economics to capture hard to observe individual-level behavior, ranging from financial trading behavior (Preis, Susannah Moat, and Stanley 2013), over the filing of unemployment claims (Choi and Varian 2012), to the contagion with infectious diseases (Carneiro and Mylonakis 2009; Seifter et al. 2010).

We searched for the keywords of interest in Google Trends (https://trends.google.com) one at a time during the stated time period in Mexico. The supplied relative search scores (ranging from 1-100) at the city-level (including low search volume regions) were then exported to capture the relative proportion of searches for each term across different cities in Mexico. These scores capture the proportion of all queries from that location (e.g., a city in our case) during the specified time period. Higher values indicate a higher proportion of all queries from that location. The Google News Lab explains this normalization of search terms as follows:

> When we look at regional search interest for a topic, we're looking at the search interest for that topic in a given region as a proportion for all searches on all topics on Google in that same place and time. For instance, if we look at the Trends around Bernie Sanders, we can see that Vermont has the highest search interest in the current senator. This is because of all states, Vermont has the highest percentage of searches for Sanders out of all searches in that state. If we had looked at raw data rather normalized values, we would've seen larger states with higher population rise to the top of the rank. ... By normalizing our data, we can make deeper insights: comparing different dates, different countries or different cities.
> (Rogers 2016)

Thereby, holding the time period constant, these relative search scores allow for easy cross-sectional comparisons of interest in a given search term across different locations.

---

[2]Google Trends uses representative, random samples from its whole, raw search volume data for computational reasons (Rogers 2016).

Table A13. **Balance Tests: Effect of Representation on Google Searches (2010-2014)**

	Outcome:			
	Chamber of Deputies	Deputy	Candidates	MORENA
	(1)	(2)	(3)	(4)
Represented	−3.878	−2.034***	−3.859	4.530***
	(2.416)	(0.127)	(5.217)	(0.370)
Constant	6.458**	11.548***	8.316**	15.102***
	(2.502)	(3.122)	(3.911)	(3.393)
Municipalities	95	95	95	95

Note: Estimates of effect of representation on the popularity of different Google search terms between January 1, 2010 and December 31, 2014 using IPW, comparing respondents in municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) with those in municipalities with a lottery candidate who was not elected (control). Mexico City is not part of the sample because here searches can only be observed at the "state" level, not the municipal (*demarcaciones territoriales*) level. Standard errors are clustered by constituency and number of lottery candidates. *p<0.1; **p<0.05; ***p<0.01, two-tailed tests.

Table A14. **Effect of Representation on Google Searches (with Controls)**

	Outcome:			
	Chamber of Deputies	Deputy	Candidates	MORENA
	(1)	(2)	(3)	(4)
Represented	6.353***	9.198***	16.352***	8.843
	(2.173)	(2.890)	(5.170)	(6.793)
Chamber of Deputies (2010-2014)	0.670***			
	(0.193)			
Deputy (2010-2014)		0.734***		
		(0.109)		
Candidates (2010-2014)			0.654***	
			(0.091)	
MORENA (2010-2014)				0.688***
				(0.065)
Constant	−0.001	−0.185	3.894***	9.393***
	(0.426)	(0.669)	(0.640)	(3.477)
Municipalities	95	95	95	95

Note: Estimates of effect of representation on the popularity of different Google search terms between June 15, 2015 and June 30, 2018 using IPW controlling for Google search term popularity in previous years (2010-2014), comparing respondents in municipalities with a MORENA lottery candidate who was elected in 2015 (treatment) with those in municipalities with a lottery candidate who was not elected (control). Higher values indicate a higher proportion of all queries from that location. Mexico City is not part of the sample because here searches can only be observed at the "state" level, not the municipal (*demarcaciones territoriales*) level. Standard errors are clustered by constituency and number of lottery candidates. *p<0.1; **p<0.05; ***p<0.01, two-tailed tests.

S16. Alternative Specifications for Candidate Background Tests

Table A15. **Difference in Electoral Participation by Candidate Background (FE Models)**

	Estimate	SE	p-value
Low Education (vs. High Education)	0.0159	0.0033	0.0164
Org. Ties (vs. No Org. Ties)	0.0240	0.0001	0.0000

Note: Estimates of difference in electoral participation by candidate background with fixed effects (FE) for constituency and the number of candidates from a municipality on the list, comparing electoral returns (in the 2018 elections) from municipalities with a MORENA lottery deputy without a post-secondary education (Low Class) or with ties to societal organizations (Org. Ties) to those with a lottery deputy with a post-secondary education and without organizational ties respectively. Standard errors are clustered by state and number of lottery candidates; p-values are two-tailed.

S17. Ethical Practices Concerning Human Participants

The interviews with members of the MORENA leadership and federal deputies that are part of this study were conducted in compliance with the ethical and transparency obligations described in APSA's *Principles and Guidance for Human Subjects Research* (2020) and APSA's *A Guide to Professional Ethics in Political Science* (2012). Furthermore, the research was conducted in compliance with the applicable laws and regulations and the research protocol was reviewed and deemed exempt from full committee and subcommittee review by the Office for Protection of Human Subjects at the University of California (Protocol Number: 2014-07-6538).

All interviews were conducted by the author himself. Informed and voluntary consent was obtained prior to the beginning of each interview and no compensation was paid to interviewees. No deception was used and interviewees were given the option to keep their identities confidential. No individual interviewee is identified in the publication of the data, unless they have given me explicit permission to use their name in my publications. In order to minimize any risk, study participants were informed of their rights not to answer any questions and to stop the interview at any time.

References

Carneiro, Herman Anthony, and Eleftherios Mylonakis. 2009. "Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks". *Clinical Infectious Diseases* 49 (10): 1557–1564.

Choi, Hyunyoung, and Hal Varian. 2012. "Predicting the Present with Google Trends". *Economic Record* 88 (s1): 2–9.

Preis, Tobias, Helen Susannah Moat, and H. Eugene Stanley. 2013. "Quantifying Trading Behavior in Financial Markets Using Google Trends". *Scientific Reports* 3:1864.

Rogers, Simon. 2016. *What is Google Trends data - and what does it mean?* Medium (July), Google News Lab. Visited on 07/22/2021. `https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8`.

Seifter, Ari, Alison Schwarzwalder, Kate Geis, and John Aucott. 2010. "The utility of "Google Trends" for epidemiological research: Lyme disease as an example". *Geospatial Health* 4 (2): 135–137.