Online Appendix: Supporting Information for

Who Runs for Congress? A Study of State Legislators and Congressional Polarization

A   Robustness to Varying Cutoffs

In the main text, legislators are assigned to the pool of any congressional district that contains at least 1/3 of the voters in their state legislative district. Here, we show that our results continue to hold under other standards for defining the pools. We consider two alternative thresholds, 25 percent and 50 percent. Under the first, legislators are assigned to the pool of any congressional district that contains 25 percent of the voters in their state legislative district, and analogously for the second. The appendix tables replicate the principal analyses of the paper using these different thresholds for defining pools. Each table and figure is labeled with the number of the corresponding table or figure in the main text, followed by a suffix indicating the threshold used.

Table A1-25 – NP-Scores of State Legislators, by Party (Threshold to Define Pools = 25 Percent)

| | Democrats | | | Republicans | | |
Item	Mean NP-Score	Median NP-Score	# Legs.	Mean NP-Score	Median NP-Score	# Legs.
National Average/Median	-0.80	-0.81	43,452	0.75	0.76	44,333
			# Pools			# Pools
All Pools	-0.87	-0.88	3,913	0.74	0.79	3,704
Pools in Safe Districts	-1.16	-1.11	1,476	0.91	0.91	1,690
Pools in Competitive Dists	-0.85	-0.87	1,015	0.63	0.69	1,056
Pools With 1+ Cong Cand	-0.89	-0.94	274	0.77	0.79	358
Pools With 0 Cong Cands	-0.87	-0.87	3,639	0.74	0.79	3,346
National Std. Dev.	0.52			0.39		
Avg. Within-State Std. Dev.	0.32			0.28		
Avg. Within-Pool Std. Dev.	0.27			0.25		

In the first row the unit of observation is the individual legislator. In rows 2-6 the unit of observation is the pool.

Table A1-50 – NP-Scores of State Legislators, by Party (Threshold to Define Pools = 50 Percent)

	Democrats			Republicans		
Item	Mean NP-Score	Median NP-Score	# Legs.	Mean NP-Score	Median NP-Score	# Legs.
National Average/Median	-0.77	-0.78	34,682	0.74	0.75	35,696
			# Pools			# Pools
All Pools	-0.85	-0.86	3,665	0.73	0.77	3,428
Pools in Safe Districts	-1.16	-1.10	1,462	0.91	0.89	1,678
Pools in Competitive Dists	-0.83	-0.86	957	0.61	0.67	1,016
Pools With 1+ Cong Cand	-0.89	-0.93	239	0.77	0.77	320
Pools With 0 Cong Cands	-0.85	-0.85	3,426	0.73	0.77	3,108
National Std. Dev.	0.50			0.38		
Avg. Within-State Std. Dev.	0.32			0.28		
Avg. Within-Pool Std. Dev.	0.26			0.24		

In the first row the unit of observation is the individual legislator. In rows 2-6 the unit of observation is the pool.

Table A3-25 – Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not (Threshold to Define Pools = 25 Percent)

|	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.010	(0.018)	[4294]	0.090	(0.021)	[6347]
Open seats	0.013	(0.020)	[2319]	0.094	(0.023)	[3200]
Safe districts	-0.022	(0.022)	[1936]	0.098	(0.029)	[3552]
Competitive districts	0.021	(0.031)	[1275]	0.102	(0.037)	[2187]

Table A3-50 – Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not (Threshold to Define Pools = 50 Percent)

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.023	(0.021)	[3197]	0.083	(0.022)	[5150]
Open seats	0.002	(0.020)	[1720]	0.073	(0.021)	[2583]
Safe districts	-0.029	(0.026)	[1379]	0.079	(0.029)	[2906]
Competitive districts	0.005	(0.037)	[920]	0.095	(0.029)	[1759]

Figure A5-25 – Within-Pool Difference in NP-Score Between State Legislators Who Ran and Did Not Run Across Presidential Vote Share (Threshold to Define Pools = 25 Percent)
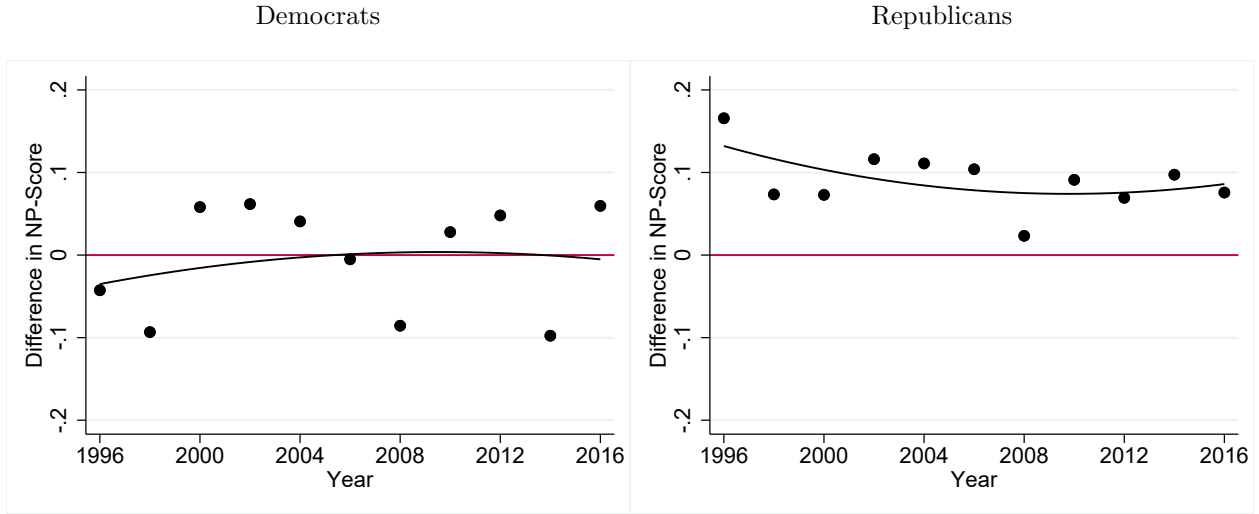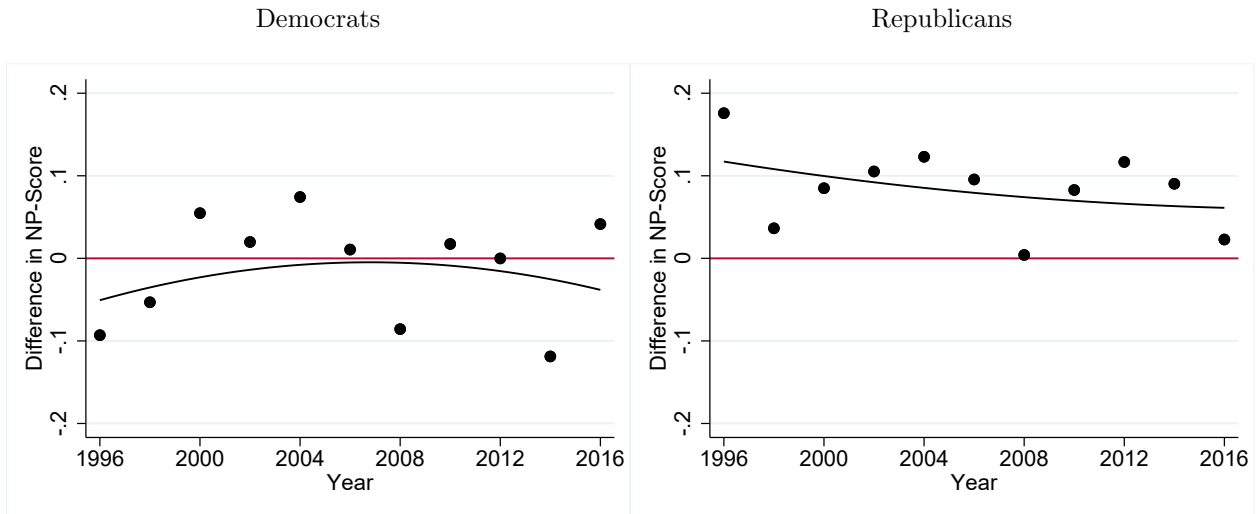
Democrats

Republicans
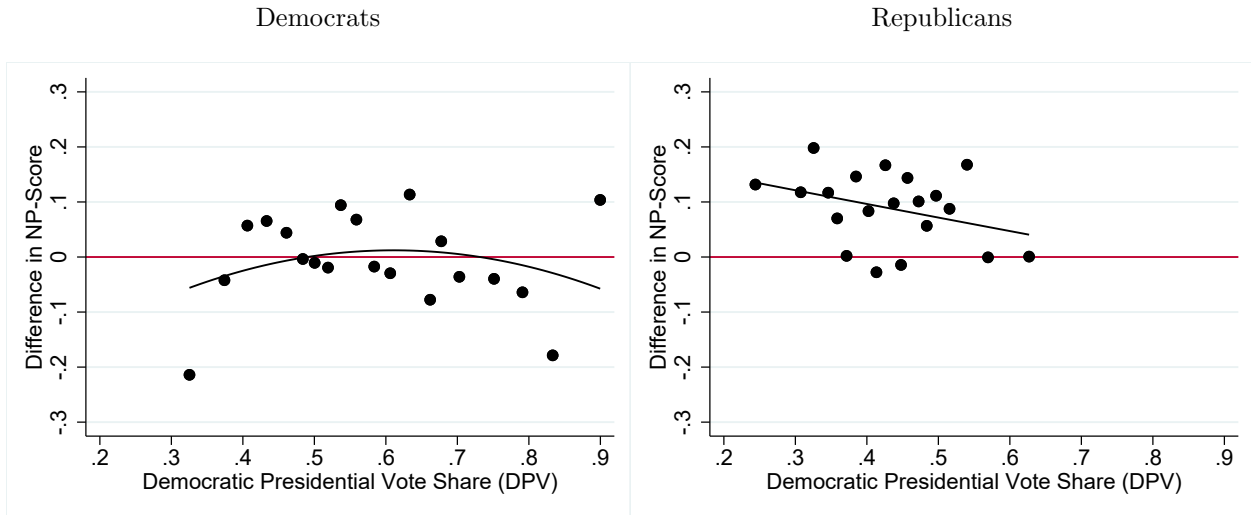
Figure A5-50 – Within-Pool Difference in NP-Score Between State Legislators Who Ran and Did Not Run Across Presidential Vote Share (Threshold to Define Pools = 50 Percent)

Democrats

Republicans
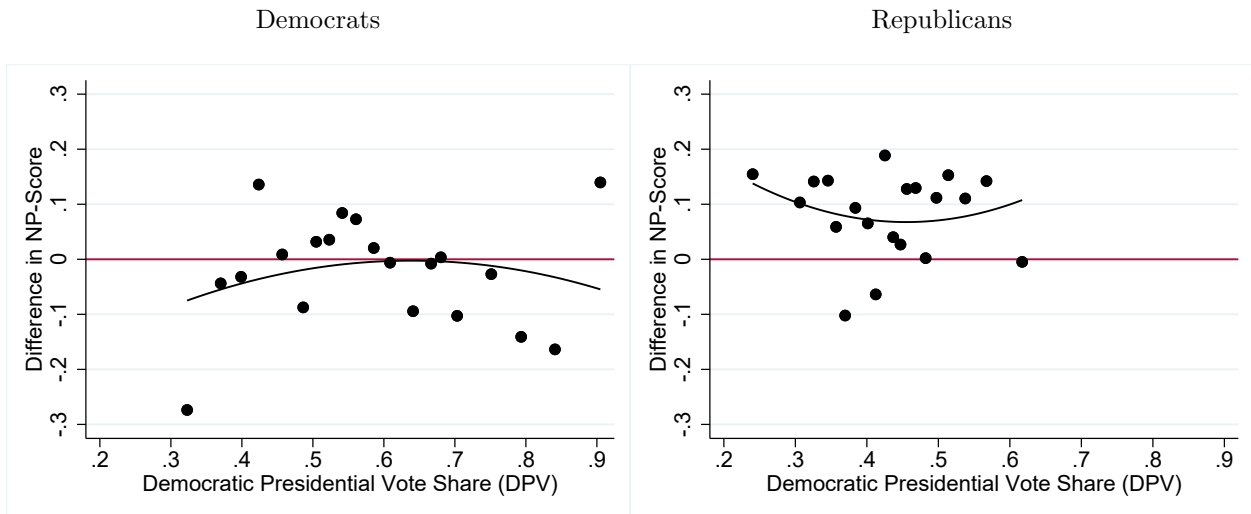
Figure A6-25 – Ideology of State Legislators Who Ran and Did Not Run Over Time (Threshold to Define Pools = 25 Percent)

Democrats   Republicans

Figure A6-50 – Ideology of State Legislators Who Ran and Did Not Run Over Time (Threshold to Define Pools = 50 Percent)

Democrats   Republicans

Table A4-25 – Within-Pool Comparisons: NP-Score of State Legislators Who Win Primary or General Elections vs. Others (Threshold to Define Pools = 25 Percent)

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
Primary Election Winners vs. Others in Pool						
All pools	-0.033	(0.026)	[2645]	0.102	(0.023)	[3690]
Safe districts	-0.046	(0.029)	[1044]	0.162	(0.033)	[2114]
Competitive districts	-0.022	(0.042)	[874]	0.054	(0.033)	[1173]
General Election Winners vs. Others in Pool						
All pools	-0.039	(0.027)	[1358]	0.130	(0.035)	[2035]
Safe districts	-0.049	(0.029)	[1013]	0.169	(0.038)	[1554]
Competitive districts	-0.019	(0.065)	[285]	0.023	(0.059)	[475]

Table A4-50 – Within-Pool Comparisons: NP-Score of State Legislators Who Win Primary or General Elections vs. Others (Threshold to Define Pools = 50 Percent)

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
Primary Election Winners vs. Others in Pool						
All pools	-0.047	(0.032)	[1955]	0.106	(0.027)	[2935]
Safe districts	-0.061	(0.033)	[690]	0.141	(0.038)	[1660]
Competitive districts	-0.030	(0.059)	[626]	0.071	(0.032)	[964]
General Election Winners vs. Others in Pool						
All pools	-0.049	(0.032)	[908]	0.123	(0.040)	[1551]
Safe districts	-0.055	(0.036)	[669]	0.152	(0.046)	[1160]
Competitive districts	-0.044	(0.090)	[196]	0.049	(0.057)	[385]

B   Predicting Probability of Running from Ideology

These analyses reverse the main analyses in Tables 3 and 4, predicting the probability of state legislators running for office and winning from their ideology. To be clear, we do not assign any causal interpretation to these estimates; instead, they simply demonstrate that the relationships between running for office and ideology still hold when the models are presented in the more intuitive way. Each table is labeled with the number of the corresponding table in the body of the paper, followed by a 1 for tables using a linear probability model and a 2 for tables using a logistic regression model.

Table B3-1 – **Within-Pool Comparisons: Predicting Probability of Running from Ideology**

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.006	(0.015)	[3898]	0.063	(0.016)	[5993]
Open seats	0.015	(0.019)	[2104]	0.077	(0.019)	[3001]
Safe districts	-0.019	(0.024)	[1566]	0.071	(0.026)	[2910]
Competitive districts	0.020	(0.021)	[1527]	0.062	(0.022)	[2720]

Table B4-1 – **Within-Pool Comparisons: Predicting Probability of Winning Primary or General Election from Ideology**

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
Primary Election Winners vs. Others in Pool						
All pools	-0.022	(0.019)	[2379]	0.077	(0.019)	[3488]
Safe districts	-0.058	(0.025)	[832]	0.098	(0.029)	[1675]
Competitive districts	0.011	(0.030)	[930]	0.069	(0.029)	[1570]
General Election Winners vs. Others in Pool						
All pools	-0.034	(0.024)	[1184]	0.090	(0.027)	[1921]
Safe districts	-0.061	(0.027)	[806]	0.095	(0.033)	[1218]
Competitive districts	0.017	(0.060)	[324]	0.080	(0.044)	[697]

Table B3-2 – Within-Pool Comparisons: Logistic Regression of Running on Ideology

Sample	Democrats			Republicans		
	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.084	(0.197)	[3898]	1.080	(0.232)	[5993]
Open seats	0.203	(0.239)	[2104]	1.157	(0.262)	[3001]
Safe districts	-0.208	(0.267)	[1566]	1.101	(0.369)	[2910]
Competitive districts	0.323	(0.306)	[1527]	1.205	(0.289)	[2720]

Table B4-2 – Within-Pool Comparisons: Logistic Regression of Winning Primary or General Election on Ideology

Sample	Democrats			Republicans		
	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
Primary Election Winners vs. Others in Pool						
All pools	-0.341	(0.304)	[2379]	1.287	(0.238)	[3488]
Safe districts	-0.859	(0.372)	[832]	1.985	(0.443)	[1675]
Competitive districts	0.162	(0.421)	[930]	0.994	(0.322)	[1570]
General Election Winners vs. Others in Pool						
All pools	-0.455	(0.339)	[1184]	1.623	(0.347)	[1921]
Safe districts	-0.914	(0.368)	[806]	1.852	(0.464)	[1218]
Competitive districts	0.194	(0.620)	[324]	1.243	(0.486)	[697]

C   Relaxing the Cardinality Assumption

In order to ensure that our findings do not rely on treating distance on the NP-Score scale as a cardinal measure of ideology, here we implement alternative specifications of our analysis in Tables 2 and 3. Our first approach, the rank specification, orders all legislators in our dataset from most to least conservative within their party and then rescales this rank from 0 to 1 (so 0 represents the legislator with the lowest NP-Score in their party, 1 represents the legislator with the highest NP-Score in their party, and all legislators are evenly spaced along the 0-1 interval). Our second approach, the within-party dummy specification, assigns each legislator a value of 0 if they are below the median NP-Score across all observations belonging to their party in the dataset and 1 if they are above this median NP-score in their party. Our third specification similarly uses a dummy variable, this time coded 1 if the legislator is above the median NP-Score in their pool and 0 otherwise (dropping all pools with fewer than three members).[1] Each table is labeled with the number of the corresponding table in the body of the paper, followed by a 1 for tables using the rank specification, a 2 for tables using the within-party dummy specification, and a 3 for tables using the within-pool dummy specification.

Table C2-1 – NP-Score of State Legislators Who Run vs. Those Who Do Not, Rank of Ideology

	Democrats		Republicans	
Item	Mean NP-Score Rank	# Obs.	Mean NP-Score Rank	# Obs.
Ran for Congress	0.429	306	0.569	404
Did Not Run	0.501	35,171	0.499	36,199

---

[1]While the first two specifications only count legislators placed in multiple pools once, as Table 2 does, this one counts them in each of their pools, which results in a slightly larger sample size even after we drop pools with one or two members.

Table C2-2 – **NP-Score of State Legislators Who Run vs. Those Who Do Not, Dummy of Ideology**

| | Democrats | | Republicans | |
| | Mean | | Mean | |
Item	NP-Score Dummy	# Obs.	NP-Score Dummy	# Obs.
Ran for Congress	0.389	306	0.592	404
Did Not Run	0.501	35,171	0.498	36,199

Table C2-3 – **NP-Score of State Legislators Who Run vs. Those Who Do Not, Dummy of Ideology Relative to Pool**

| | Democrats | | Republicans | |
| | Mean NP-Score | | Mean NP-Score | |
Item	Dummy Relative	# Obs.	Dummy Relative	# Obs.
Ran for Congress	0.450	307	0.630	395
Did Not Run	0.478	38,763	0.477	40,306

Table C3-1 – **Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not, DV = Rank of Ideology**

| | Democrats | | | Republicans | | |
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.008	(0.010)	[3898]	0.060	(0.015)	[5993]
Open seats	0.001	(0.012)	[2104]	0.063	(0.016)	[3001]
Safe districts	-0.012	(0.014)	[1718]	0.068	(0.021)	[3358]
Competitive districts	0.004	(0.021)	[1169]	0.064	(0.022)	[2066]

Table C3-2 – **Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not, DV = Dummy of Ideology**

| | Democrats | | | Republicans | | |
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	0.002	(0.024)	[3898]	0.082	(0.028)	[5993]
Open seats	0.020	(0.027)	[2104]	0.094	(0.029)	[3001]
Safe districts	-0.013	(0.030)	[1718]	0.103	(0.037)	[3358]
Competitive districts	0.049	(0.047)	[1169]	0.082	(0.039)	[2066]

Table C3-3 – Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not, DV = Dummy of Ideology Relative to Pool

Sample	Democrats			Republicans		
	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.025	(0.033)	[3878]	0.179	(0.038)	[5973]
Open seats	-0.036	(0.043)	[2094]	0.165	(0.043)	[2991]
Safe districts	-0.088	(0.042)	[1704]	0.179	(0.052)	[3352]
Competitive districts	0.081	(0.073)	[1167]	0.204	(0.048)	[2056]

D   Including Retired State Legislators in the Pools

Here, we show that our results are robust to including state legislators who are not currently serving at the time they run for Congress. One challenge in matching retired state legislators to congressional districts is that the more distant in time their service is from their candidacy, the less likely a mapping from state legislative to congressional districts is to be available. Accordingly, here we focus on one-term-out state legislators, or those who most recently served in the two-year legislative session during the preceding electoral cycle. (In states with odd-year legislative elections, we include all legislators who served within four years of their candidacy for Congress.) For example, state legislators who served in 2007, 2008, 2009, or 2010 are in the expanded pool for the congressional elections of 2010. In our data, 81.7% of state legislators who ever run for Congress run either as incumbents or within two years after leaving. We replicate Tables 2, 3, and 4 for all states and years for which we have legislator information and district mappings available for this expanded pool. Even with this modest expansion of the data, we lose information for a good number of pools, resulting in a slightly smaller sample of congressional candidates on net. Overall, the qualitative results appear quite similar when using this broader definition of the pool.

Table D2 – NP-Score of State Legislators Who Run vs. Those Who Do Not – Expanded Pool Including Retired State Legislators

	Democrats		Republicans	
Item	Mean. NP-Score	# Obs.	Mean NP-Score	# Obs.
Ran for Congress	-0.949	265	0.858	367
Did Not Run	-0.783	35,403	0.738	37,017

Table D3 – Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not – Expanded Pool Including Retired State Legislators

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
All pools	-0.022	(0.023)	[3909]	0.090	(0.022)	[6523]
Open seats	-0.002	(0.021)	[2039]	0.080	(0.022)	[3267]
Safe districts	-0.012	(0.026)	[1597]	0.108	(0.029)	[3656]
Competitive districts	0.009	(0.030)	[1280]	0.078	(0.023)	[2467]
Large pools	-0.024	(0.024)	[3049]	0.109	(0.021)	[5215]
Small pools	-0.019	(0.047)	[860]	0.066	(0.035)	[1308]
High-variance pools	-0.019	(0.039)	[2294]	0.116	(0.032)	[4078]
Low-variance pools	-0.025	(0.018)	[1615]	0.063	(0.017)	[2445]

Table D4 – Within-Pool Comparisons: NP-Score of State Legislators Who Win Primary or General Elections vs. Others – Expanded Pool Including Retired State Legislators

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
Primary Election Winners vs. Others in Pool						
All pools	-0.045	(0.035)	[2387]	0.109	(0.026)	[3672]
Safe districts	-0.030	(0.031)	[779]	0.182	(0.035)	[2110]
Competitive districts	-0.037	(0.044)	[851]	0.055	(0.035)	[1311]
General Election Winners vs. Others in Pool						
All pools	-0.038	(0.032)	[1049]	0.134	(0.038)	[2056]
Safe districts	-0.025	(0.033)	[756]	0.194	(0.040)	[1504]
Competitive districts	-0.079	(0.083)	[230]	0.006	(0.059)	[546]

E   Heterogeneous Effects by Pool Size and Dispersion

This table examines whether the patterns we find vary based on pool characteristics, in particular whether the results differ for large pools and small pools, or for pools with and without substantial ideological variation. Specifically, we split all pools for each party into two groups based on the median pool size and two groups based on the median within-pool standard deviation in NP-Score for that party. We then re-run the analyses in Table 3 for these subsets of the overall sample. Although the observed difference between runners and non-runners among Republicans appears to be greater in large pools than in small pools, and in pools with a larger standard deviation relative to pools with less variance, the difference between these coefficients is not statistically significant.

Table E3 – Within-Pool Comparisons: NP-Score of State Legislators Who Run vs. Those Who Do Not, Additional Results on Heterogeneity

	Democrats			Republicans		
Sample	Coeff.	Std. Error	# Obs.	Coeff.	Std. Error	# Obs.
Large pools	-0.015	(0.023)	[3105]	0.107	(0.024)	[4663]
Small pools	0.002	(0.037)	[793]	0.075	(0.034)	[1330]
High-variance pools	-0.000	(0.034)	[2285]	0.110	(0.031)	[3759]
Low-variance pools	-0.016	(0.016)	[1613]	0.074	(0.019)	[2234]

F   Missing Pools

This table examines whether the missing pools are unrepresentative in a way that might skew our results. As discussed in the main text, we are able to incorporate pools from all 49 states that constitute our analysis—which excludes Nebraska—and drop state-years where mappings could not be constructed for all districts, alleviating potential concerns about geographical representativeness. One remaining issue is that the available years for the incomplete states might systematically differ in some way from the missing ones. Accordingly, in Table F1, we regress measures of district-level partisanship (*Democratic Presidential Vote*) and competitiveness (the absolute difference between *Democratic Presidential Vote* and 0.5, which will be zero in a district that mirrors the national popular vote and grow greater as a district becomes less competitive) on an indicator variable for whether that district-year is missing from our data, including state and year fixed effects. The results show that the missing districts do not appear to be outliers: *Democratic Presidential Vote* and the absolute presidential vote margin relative to the national vote in these districts are on average only a fraction of a percentage point different from their values in the years that are available from the same state, after accounting for nationwide differences in district partisanship and competitiveness between years.

Table F1 – Missing vs. Non-Missing Districts

Variable	Vote Share	Absolute Margin
Missing	0.003	-0.002
	(0.005)	(0.005)
Observations	4785	4785

State and year fixed effects included in both specifications. Standard errors, clustered by state, in parentheses.