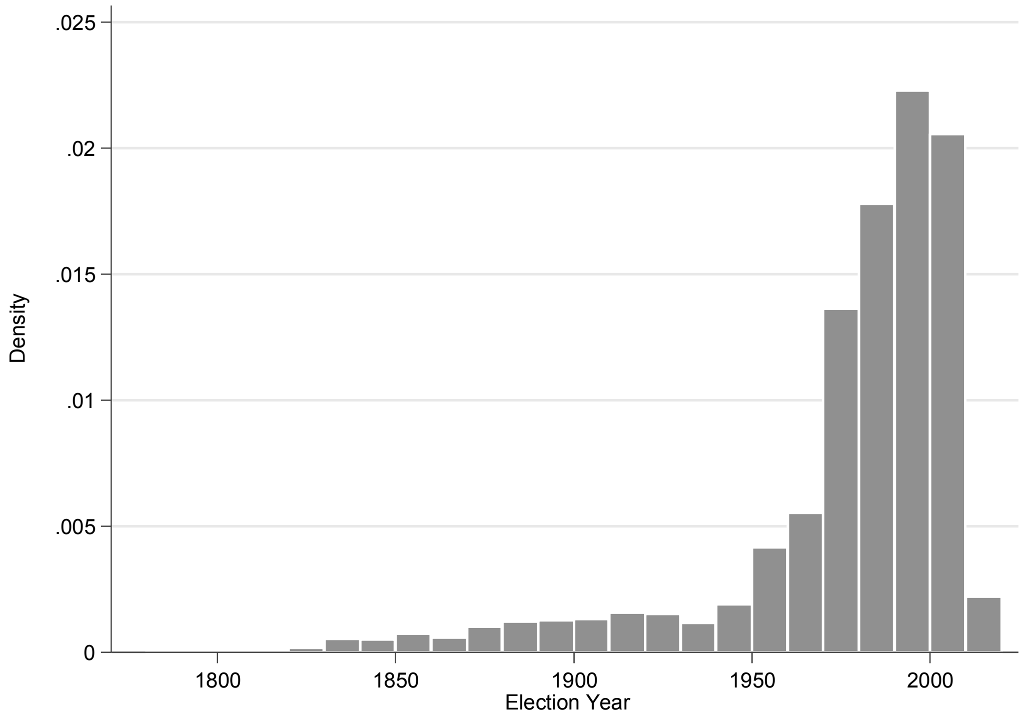
APPENDIX A:

Multi-Level Elections Archive (MLEA)

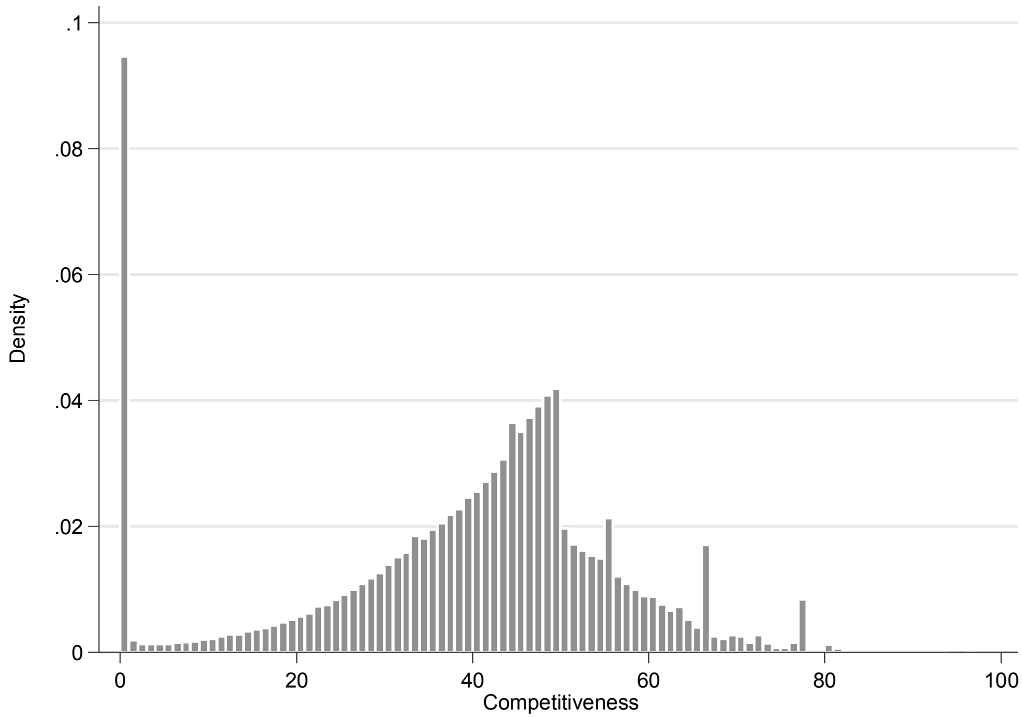
This appendix contains discussion and descriptive statistics corresponding to analyses in Table 1 and other analyses relying on data from MLEA. For present purposes, an "observation" in MLEA signifies that information is available for the principal dependent variable, Competitiveness, though not necessarily for other variables contained in the analyses.

Figure A1:
Distribution of Data through Time (MLEA)



Density plot of observations for Competitiveness (100 – share of largest party) from 1788 to 2013 in the MLEA dataset (*N*=415,095).

Figure A2:
Histogram of Competitiveness (MLEA)

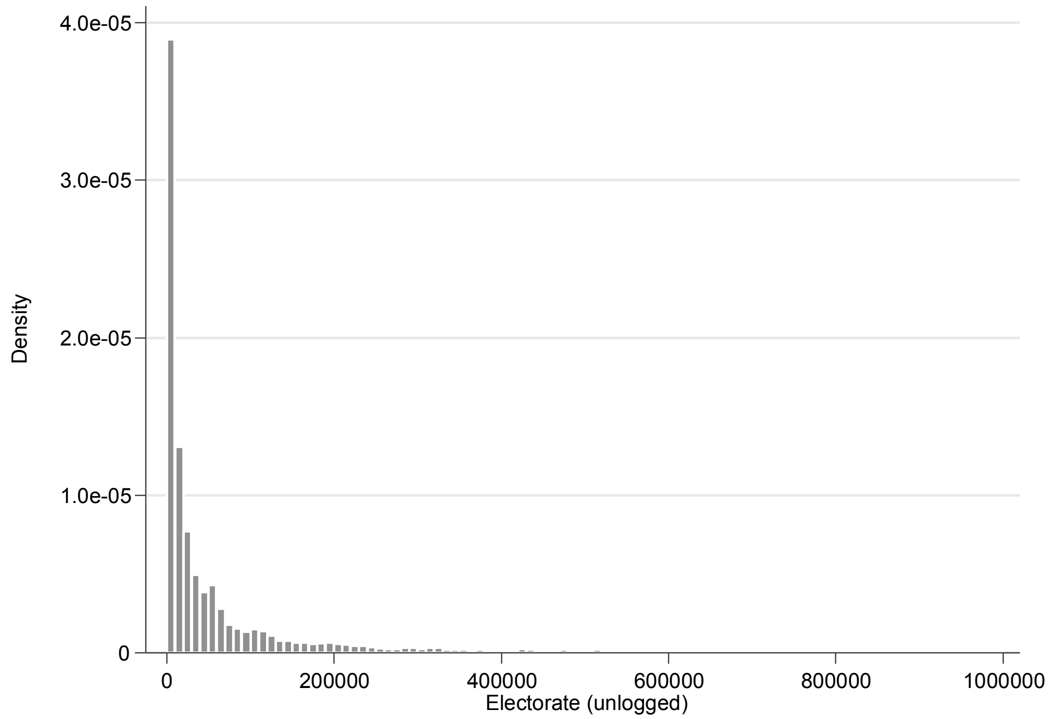


Histogram of Competitiveness (100 – share of largest party) in the MLEA dataset (*N*=415,095).

Figure A3:
Competitiveness through Time (MLEA)

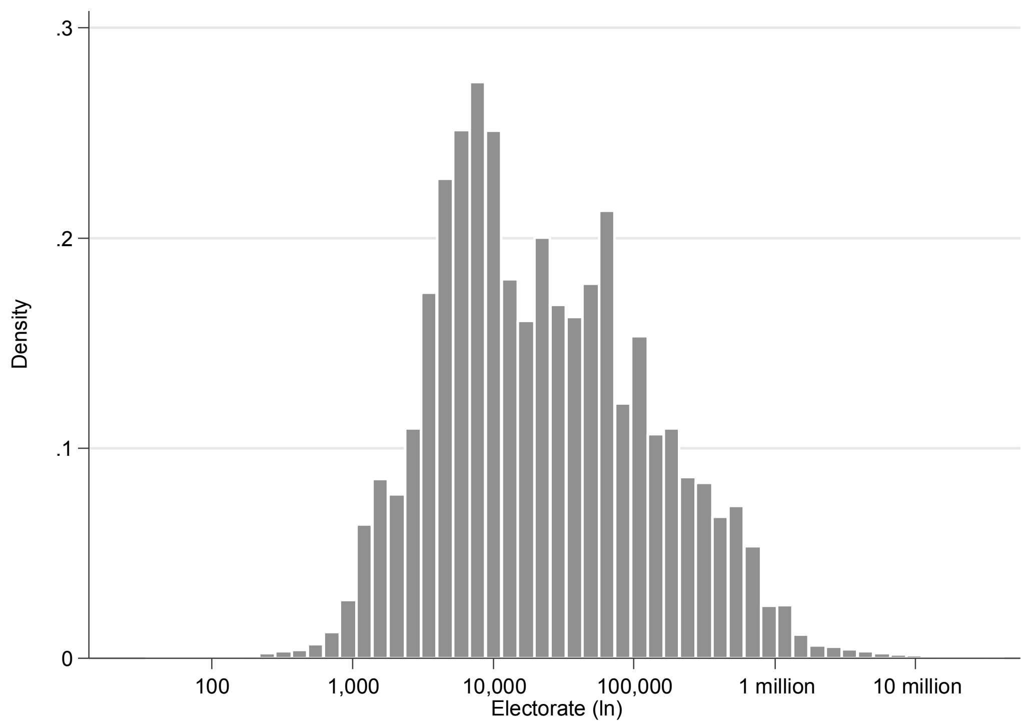
Smoothed graph (10-year running mean) of Competitiveness (100 – share of largest party) from 1792 to 2008 in the MLEA dataset.

Figure A4:
Histogram of Electorate (unlogged) (MLEA)



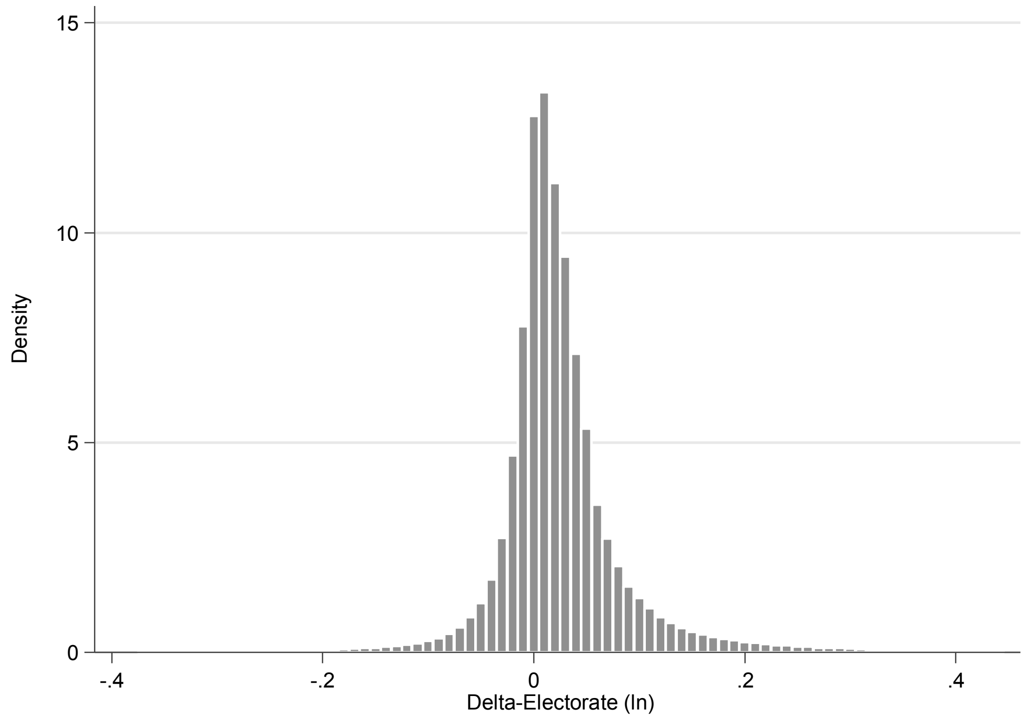
Histogram of Electorate (unlogged) observations in the MLEA dataset (*N*=377,519), excluding extreme outliers (Electorate>1 million; *N*=6,811).

Figure A5:
Histogram of Electorate (ln) (MLEA)



Histogram of Electorate (natural logarithm) observations in the MLEA dataset (*N*=384,330).

Figure A6:
Histogram of ΔElectorate (ln) (MLEA)



Histogram of ΔElectorate (natural logarithm) from election to election in the MLEA dataset (*N*=301,181), excluding extreme outliers (SD > 3.0, *N*=3,788).

Table A1:

Summary by Elective Office (MLEA)

Office	Years	Countries	Elections	Contests	Electorate (1000s)				Comp
					Min	*Max*	*Mean*	*SD*	*Mean*
National									
Upper chamber	1902-2010	2	133	1,598	50.24	42,108	3,820	4,513	40.1
Lower chamber	1788-2013	86	1,359	108,948	0.04	28,038	224	426	41.6
Regional									
Governor	1792-2010	2	489	2,926	21.11	42,108	2,280	3,203	41.6
Upper chamber	1968-2010	1	77	23,839	11.13	1,142	139	141	28.6
Lower chamber	1968-2010	1	94	84,772	7.27	612	62	65	26.6
Local									
Mayor	1853-2012	3	623	46,032	0.06	11,090	47	214	41.7
Council	1912-2010	4	188	146,980	0.08	11,090	11	72	45.3
All									
Total	1788-2013	88	2,344	415,095	0.04	42,108	123	569	35.9

The MLEA dataset, disaggregated by elective office. *Comp:* Competitiveness (100 – share of largest party).

Table A2:
Distribution of Contests across Countries (MLEA)

United States	155,515	Cameroon	123
United Kingdom	143,276	Guyana	102
Brazil	42,840	Russia	102
Mexico	16,727	Saint Lucia	85
Germany	9,291	Armenia	82
Sweden	5,782	Gambia	81
India	5,361	Georgia	73
France	3,736	Togo	61
Australia	3,343	Czech Republic	60
Denmark	3,214	Honduras	54
Japan	2,219	Cape Verde	53
Belgium	1,573	Estonia	46
Korea	1,512	Liechtenstein	42
Italy	1,477	Luxembourg	40
Norway	1,336	Mozambique	37
New Zealand	1,264	Andorra	35
Hungary	1,174	Suriname	30
Netherlands	1,036	Guinea-Bissau	27
Switzerland	1,036	Macedonia	27
Turkey	961	Tunisia	27
Ireland	957	Seychelles	26
Greece	906	Cambodia	24
Iceland	722	Costa Rica	21
Jamaica	716	British Virgin Is	20
Austria	602	Equatorial Guinea	18
Zambia	597	Croatia	17
Bangladesh	573	Vanuatu	17
Pakistan	541	Cayman Is	12
Finland	540	South Africa	9
Lesotho	502	Rwanda	6
Singapore	475	Latvia	5
Solomon Is	464	Gibraltar	4
Barbados	376	Namibia	4
Bolivia	369	San Marino	4
Spain	364		
Kenya	329		
Belize	313		
Bermuda	283		
Argentina	239		
Tanzania	224		
Poland	223		
Malaysia	222		
Bahamas	219		
Albania	200		
Ghana	200		
Vincent/Grenadines	193		
Malawi	185		
Portugal	180		
Botswana	169		
Moldova	162		
Sri Lanka	132		
Romania	126		
Azerbaijan	125		
Dominican Rep	124		

Table A3:
Variable Definitions

Outcomes

Competitiveness. 100 – vote share of the largest party. Source: coded by authors. *comp_largest*

Competitiveness (incumbent). 100 – vote share of the incumbent party (i.e., the party with the largest vote share in the previous election). Source: coded by authors. *comp_incumb*

Herfindahl index. The sum of the squares of each party's vote share. Source: coded by authors. *Herf*

Margin of victory (top two). 100 – the difference the two parties with the top two vote shares. Source: coded by authors. *comp_diff*

Parties contesting (ln). The number of parties contesting the election, logged. Source: coded by authors. *Lnparties*

Turnover. 1 if change in party control, 0 otherwise (applies only to SMDs). Source: coded by authors. *Turnover*

Causal Factors of Theoretical Interest

Electorate. The number of eligible voters in a district. If unavailable, this may be proxied by the population of the district in time-periods when suffrage is universal. Sources: see text. *electorate*

Electorate (ln). Electorate, transformed by the natural logarithm (the benchmark measure). Sources: see text. *electorate_ln*

ΔElectorate (ln). Electorate recoded as first-difference. *pev_change*

□ **Electorate (ln).** ΔElectorate (ln) recoded as 0 if <0. Source: authors. *pev_more*

□ **Electorate (ln).** ΔElectorate (ln) recoded as 0 if >0, then multiplied by -1. Source: authors. *pev_less*

Female suffrage. Coded 0 until universal female adult suffrage is established, 1 thereafter. Coded 0 if female and male suffrage are introduced simultaneously. Source: Przeworski (2013). *female_suffrage*

Youth suffrage. Coded 0 until suffrage is extended to youth (variously defined between the ages of 18-25), 1 thereafter. If there is more than one episode of youth suffrage extension in a country's history, the largest such extension is regarded as the treatment and other episodes are ignored. Source: Przeworski (2013). *youth_suffrage*

Electoral System Dummies

Majoritarian, block ballot. Coded 1 if electoral system is majoritarian with block ballot. Source: Colomer et al. (2006). *maj1*

Majoritarian, cumulative ballot. Coded 1 if electoral system is majoritarian with cumulative ballot. These multi-member districts allow voters to cast multiple votes for one or more candidates. Source: Colomer et al. (2006). *maj2*

Mixed. Coded 1 if electoral system includes parallel SMD and MMD seats (with or without compensation for disproportionality induced by SMD elections) and data sources do not allow us to determine which districts are SMD. Source: Colomer et al. (2006). *Mix*

PR, avg. mag<9. Coded 1 if electoral system is proportional with mean district magnitude less than 9. Source: Colomer et al. (2006). *pr1*

PR, avg. mag>9, closed list. Coded 1 if electoral system is proportional with mean district magnitude greater than 9 and closed lists. Source: Colomer et al. (2006). *pr2*

PR, avg. mag>9, open list. Coded 1 if electoral system is proportional with mean district magnitude greater than 9 and open lists. Source: Colomer et al. (2006). *pr3*

Round. Coded 2 if the second of two rounds, 1 otherwise. Source: coded by authors. *Round*

Secret ballot. Coded 1 if ballot is secret. 0 otherwise. Source: Przeworski (2013). *Secret*

SMD. Coded 1 if electoral system is single member district. Source: Colomer et al. (2006). *maj3*

Additional Variables

District. Identifies each unique office-district in MLEA. If two elective offices have identical districts (e.g., Senate and Gubernatorial elections in the US) they are assigned unique district identifiers. The term "district fixed-effects" thus refers to office-district fixed-effects. This variable is constructed with three pieces of information: country, office type, and district. In order to keep the id constant for districts that have multiple years of election data we use either uniquely identifying district codes or, if those are unavailable, district names from the original data source. Sources: various. *Id*

District magnitude (ln). The number of positions contested, logged. Source: various. *dm_ln*

Year. Dummies for each year in the dataset. *Year*

Note: Includes variables employed in the main analyses as well as those employed in robustness tests.

Table A4:
Descriptive Statistics (MLEA)

Outcomes	**Countries**	**Years**	**Obs**	**Mean**	**SD**	**Min**	**Max**
Competitiveness	88	1788-2013	384330	39.784	17.864	0	96
Competitiveness (incumbent)	51	1790-2013	275430	42.905	22.815	0	100
Herfindahl index	62	1832-2011	71819	0.461	0.183	0	1
Margin of victory (top two)	88	1788-2013	330178	75.389	21.018	0	100
Parties contesting (ln)	62	1832-2011	71819	1.275	0.667	0	6
Turnover	27	1790-2013	275464	0.274	0.446	0	1
Causal factors of theoretical interest							
Electorate	88	1788-2013	385741	122769	568501	41	42108444
Electorate (ln)	88	1788-2013	385741	10.032	1.741	4	18
ΔElectorate (ln)	70	1788-2013	304969	0.031	0.135	-6	6
□ Electorate (ln)	70	1788-2013	304969	0.042	0.118	0	6
□ Electorate (ln)	70	1788-2013	304969	0.011	0.059	0	6
Female suffrage	88	1788-2013	384641	0.706	0.456	0	1
Youth suffrage	88	1788-2013	384587	0.741	0.438	0	1
Electoral system covariates							
Majoritarian, block ballot	88	1788-2013	385741	0.012	0.109	0	1
Majoritarian, cumulative ballot	88	1788-2013	385741	0.000	0.000	0	0
SMD	88	1788-2013	385741	0.865	0.342	0	1
PR, avg. mag <9	88	1788-2013	385741	0.031	0.173	0	1
PR, avg. mag >9, closed list	88	1788-2013	385741	0.016	0.124	0	1
PR, avg. mag >9, open list	88	1788-2013	385741	0.072	0.259	0	1
Mixed system	88	1788-2013	385741	0.004	0.062	0	1
District magnitude (ln)	88	1788-2013	369800	0.206	0.687	0	5
Round	88	1788-2013	385741	1.017	0.128	1	2
Secret ballot	88	1788-2013	385741	0.961	0.194	0	1

Includes variables employed in the main analyses as well as those employed in robustness tests shown in appendices. Limited to observations for which data is available for Electorate. Min and Max values are rounded to nearest integer.

Alternate Measures of Contestation

The concept of interest in this study, contestation, may be measured in a variety of ways. We focus our discussion on methods that are (a) applicable at the district level, (b) broad in coverage, and (c) relevant for considerations of democracy and governance.[1] Six options are summarized in Table A5.

A first set of alternatives arises from altering the *aggregation formula* for competitiveness. Our benchmark measure (1) is 100 minus the vote-share of the largest party. One might also focus on (2) competitiveness relative to the incumbent party (100 minus the vote-share of the party that received the highest vote-share in the previous election) or on (3) the margin of victory between the two largest parties. Because the latter two options rely on information for successive elections and/or for more than one party they reduce the potential sample of observations, and are on this account less desirable.

Option (2) also introduces a risk of measurement error. Note that measuring incumbency requires identifying the incumbent, and this is difficult in polities where parties change names (but not necessarily identities) from election to election. Sometimes, this change is imposed by state authorities, who outlaw an important party but cannot prevent its recrudescence under another name (e.g., Thailand and Turkey in recent decades). Consequently, this measurement strategy may assign a high competitiveness score to a district that features exceptionally fluid party names but is in reality controlled by the same cadre. Measuring competitiveness in this fashion may end up conflating contestation with non-institutionalized party politics.

Another set of alternatives focuses on the composition of the *party system* within a district. One may simply count (4) the number of parties competing in a district, transformed by the natural logarithm (in order to discount the value of each additional party). Alternatively, one may take account of the relative size of each party by constructing (5) a Herfindahl index, understood as the sum of squares of each party's share of the vote. A number close to zero represents a highly fragmented party system and 1.00 represents a district in which a single party wins all the votes. The Herfindahl index is thus inversely correlated with other indicators of competitiveness. Note that the sum-of-squares aggregation procedure gives great weight to the largest party, and helps to account for why the Herfindahl index and our benchmark measure of competitiveness are so highly correlated (Pearson's r = -.961).

Measures of contestation based on overall party system composition invoke an implicit assumption that low entry barriers, and many competitors, enhance system performance. While this is presumably true in market competition (where goods are "private") it is perhaps less true when evaluating the performance of political systems (where goods are "public"). The production of a public good requires that a party be accountable to *all* members of a district (or a polity), or at least to a majority. Insofar as competition is highly fragmented, accountability relationships are likely to be attenuated (because the behavior of elected officials is harder to monitor) and narrowed (to specific constituencies within a district). Arguably, competition among two or three parties has a more positive effect on the quality of democracy and governance than competition across twenty or thirty. This is why we prefer to measure (political) contestation as the size of the largest party rather than the number of parties or the overall dispersion of party votes.

A final approach to measuring contestation focuses on (6) *turnover*, understood as a change in party control of a district. This has a clear and intuitive meaning in SMD contests. Unfortunately, it

[1] For discussion of the conceptualization and measurement of contestation, competition, competitiveness, and related concepts see Altman & Perez-Linan (2002), Bartolini (1999, 2000), Elkins (1974), Strøm (1989). Note that some measures (e.g., Blais & Lago 2009; Grofman & Selb 2009) are designed specifically to test the impact of competitiveness on turnout and are not appropriate in the present context.

does not translate well to MMDs. Of course, one can calculate the turnover of seats possessed by the largest party (perhaps as a ratio of the total seats in that district). However, low turnover cannot be viewed as a lack of electoral democracy; it may simply indicate that voting patterns are stable. Suppose a district features three parties with roughly equal shares of the votes and seats in that district. Under the circumstances, the turnover of seats from election to election is not a sign of greater or lesser democracy. Even if restricted to SMDs, turnover is problematic as an empirical indicator by virtue of its "lumpy" character. Occurring on an irregular basis, many observations are required in order to tease out a signal from the background noise.

Thus, for a variety of reasons we are inclined to regard our benchmark measure of competitiveness as the most useful overall measure of contestation. Reassuringly, alternate measures (with the exception of Turnover) are highly correlated, as shown in the final column of Table A5. Indeed, some measures such as the Herfindahl index are virtually indistinguishable from our benchmark measure of competitiveness.

Not surprisingly, when analyses are replicated with alternative measures the results are generally robust. In Table A6, we replicate all models in Table 1 except Model 4 (by reason of space and its tangential nature). Results are shown for the benchmark measure of contestation (1) and four alternatives (2-5), as described above. (We exclude Turnover, for reasons stated.) Although we cannot directly compare coefficients (because competitiveness is measured in different units), it is clear that the relationship is quite robust. All but three of these seventy-eight tests show the expected relationship between Electorate and Competitiveness (at .10 levels).

Table A5:
Alternate Indicators of Electoral Contestation

	Period	Electoral Systems	Countries	Elections	Districts	Contests	r
1. Competitiveness (benchmark) 100 – largest VS	1788-2013	All	88	2,212	79,658	384,330	--
2. Competitiveness (incumbent) 100 – incumbent VS	1790-2012	All	51	1,914	63,856	275,430	.799
3. Margin of victory (top two) 100 – (largest VS – second largest VS)	1788-2013	All	88	2,164	71,430	330,178	.900
4. Parties contesting Number of parties contesting (ln)	1832-2011	All	62	778	11,084	71,819	.632
5. Herfindahl index Sum of squares of each VS	1832-2011	All	62	778	11,084	71,819	-.961
6. Turnover Change in party control (0/1)	1790-2012	SMD	27	1,514	60,415	275,464	.325

VS: vote-share of a party. *Sample:* MLEA dataset, including only observations that are available for Electorate. *r:* Pearson's r correlation with Competitiveness (benchmark).

Table A6:
Robustness Tests with Alternate Measures of Competitiveness

Models	1	2	3§	5	6	7	8	9	10	11	12	13	14
Outcome	Y	Y	ΔY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Estimator	OLS, FE	OLS, FE	OLS	RE	OLS, FE	OLS, FE	RE	RE	OLS, FE	OLS, FE	OLS, FE	OLS, FE	OLS, FE
Sample	Entire	Entire	Entire	Entire	Entire	Historical	Non-OECD	Semi-Demos	No proxies	Lower chamber	Local	SMD	MMD
1. Competitiveness	3.022***	2.632***	2.884***	2.135***	3.413***	3.994***	2.327***	3.344***	2.920***	2.960***	2.571***	3.554***	3.243***
(benchmark)	[0.247]	[0.249]	[0.248]	[0.053]	[0.274]	[0.466]	[0.060]	[0.151]	[0.255]	[0.318]	[0.370]	[0.293]	[0.471]
2. Competitiveness	1.987***	1.150***	7.578***	1.780***	2.063***	7.433***	2.933***	3.332***	2.206***	2.840***	2.106***	3.013***	2.089**
(incumbent)	[0.333]	[0.298]	[0.465]	[0.074]	[0.372]	[0.633]	[0.115]	[0.224]	[0.362]	[0.418]	[0.559]	[0.386]	[0.817]
3. Margin of	2.092***	0.823**	1.885***	0.740***	3.205***	1.403*	1.301***	4.182***	2.006***	1.152**	3.860***	2.766***	1.893**
victory (top two)	[0.363]	[0.353]	[0.366]	[0.077]	[0.413]	[0.835]	[0.113]	[0.286]	[0.403]	[0.482]	[0.659]	[0.418]	[0.761]
4. Parties	0.093***	0.057***	0.130***	0.085***	0.093***	0.063***	0.192***	0.035***	0.093***	0.093***	n/a	0.118***	0.017
contesting (ln)	[0.010]	[0.009]	[0.009]	[0.007]	[0.010]	[0.013]	[0.017]	[0.007]	[0.010]	[0.010]		[0.011]	[0.017]
5. Herfindahl	-0.026***	-0.016***	-0.040***	-0.019***	-0.025***	-0.036***	0	0.005	-0.026***	-0.026***	n/a	-0.030***	-0.029**
Index	[0.005]	[0.005]	[0.004]	[0.003]	[0.005]	[0.008]	[0.004]	[0.005]	[0.005]	[0.005]		[0.005]	[0.012]
Y_{t-1}		X											
District mag, (ln)					X								
Electoral system (D)	X	X	X	X		X	X	X	X	X	X		X
Office (D)				X			X	X					
District (D)	X	X			X	X			X	X	X	X	X
Country (D)				X			X	X					
Year (D)	X	X	X	X	X	X	X	X	X	X	X	X	X

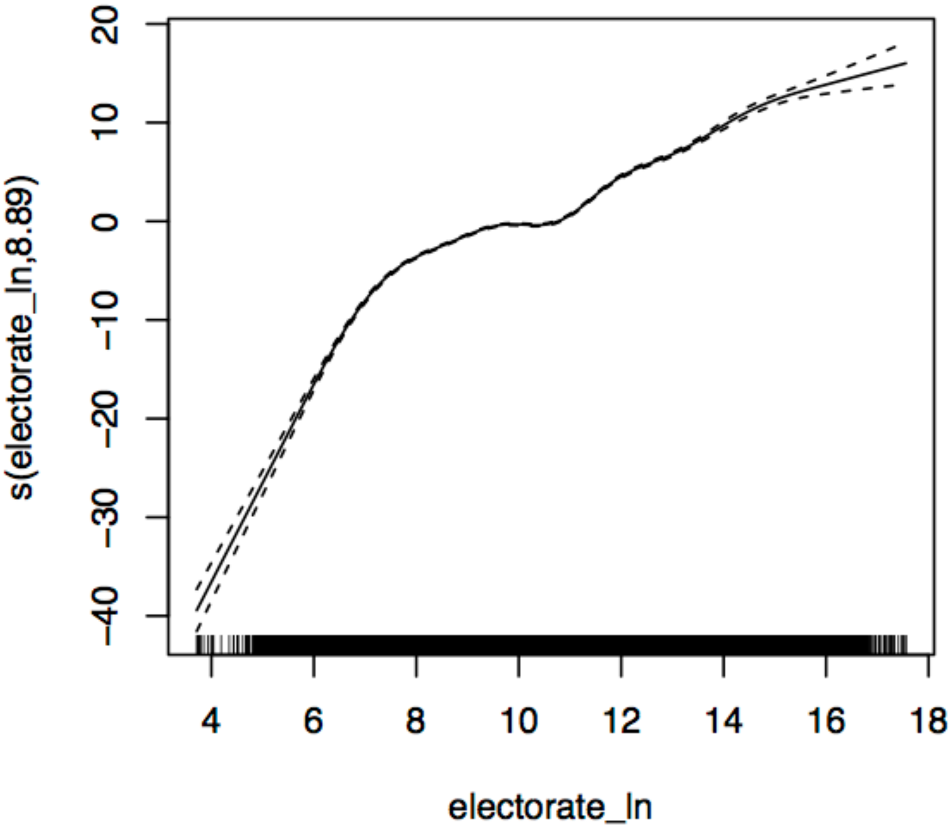
Model numbers correspond to those in Table 1 (Model 4 excluded, as explained in text). §=Δ*X* *Outcome:* Competitiveness (100 – share of largest party). *Samples: Entire* (MLEA dataset, including all non-missing observations), *No proxies* (electorate not proxied by population). *D:* dummies. *Estimators:* OLS, FE (ordinary least squares regression with district fixed effects), RE (random effects), all standard errors clustered by district. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

Table A7:
Modeling the Outcome and the Predictor

	__Variations in *Y*__		______Variations in *X*______		
	1	**2**	**3**	**4**	**5**
Competitiveness (Y)	1-100	0/1	0-100	0-100	0-100
Electorate (ln) (X)	Unrestricted	Unrestricted	E<3k	3k<E<500k	E>500k
Estimator	OLS, FE	Logit, FE	OLS, FE	OLS, FE	OLS, FE
Electorate (ln)	1.785***	0.765***	4.428***	2.093***	3.005**
	[0.191]	[0.055]	[0.889]	[0.301]	[1.273]
Year time trend		X			
Year dummies	X		X	X	X
Years	1788-2013	1788-2013	1832-2012	1788-2013	1805-2013
Countries	88	20	36	84	40
Districts	77,954	12,182	11,833	66,910	3,542
Contests (N)	354,659	69,675	39,045	325,280	20,005
R2 (within)	(0.060)		(0.116)	(0.057)	(0.144)
Log likelihood		-24,846			

Outcome = *Competitiveness*, originally coded as 100 – share of largest party (range = 0-100). Binary coding equals 0 (Competitiveness=0) or 1 (Competitiveness > 0). *E* = Electorate (ln). *Estimators:* Logit, FE (logistic regression with district fixed effects), OLS, FE (ordinary least squares with district fixed effects, standard errors clustered by district). *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

Figure A7:
Generalized Additive Model



Generalized additive model (GAM) of Competitiveness regressed against Electorate (ln) along with fixed effects for Year, Country, and Electoral system.

Table A8: **Cubic Polynomial Tests**

	1	2	5	6	7	8	9	10	11	12	13	14
Outcome	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Estimator	OLS, FE	OLS, FE	RE	OLS, FE	OLS, FE	RE	RE	OLS, FE	OLS, FE	OLS, FE	OLS, FE	OLS, FE
Electorate (ln)	26.725***	28.140***	23.110***	27.713***	14.644*	5.632*	-13.648***	57.153***	53.036***	-19.667*	34.768***	-2.725
	[5.129]	[5.160]	[2.485]	[5.168]	[7.513]	[3.035]	[4.100]	[6.435]	[6.752]	[10.673]	[5.734]	[8.860]
Electorate (ln)^2	-1.865***	-2.002***	-1.693***	-1.834***	-0.306	0.105	1.791***	-5.663***	-5.069***	2.788**	-2.390***	0.364
	[0.477]	[0.476]	[0.240]	[0.482]	[0.757]	[0.288]	[0.411]	[0.664]	[0.675]	[1.132]	[0.533]	[0.833]
Electorate (ln)^3	0.042***	0.045***	0.041***	0.036**	-0.022	-0.017*	-0.061***	0.190***	0.164***	-0.113***	0.049***	-0.005
	[0.014]	[0.014]	[0.008]	[0.014]	[0.025]	[0.009]	[0.013]	[0.022]	[0.022]	[0.040]	[0.016]	[0.025]
Y_{t-1}		X										
District mag (ln)				X								
Elect syst (D)	X	X	X		X	X	X	X	X	X		X
Office (D)			X			X	X					
District (D)	X	X		X	X			X	X	X	X	X
Country (D)			X			X	X					
Year (D)	X	X	X	X	X	X	X	X	X	X	X	X
Sample	Entire	Entire	Entire	Entire	Historical	Non-OECD	Semi-Demos	No proxies	Lower chamber	Local	SMD	MMD
Years	1788-2013	1790-2013	1788-2013	1788-2013	1788-1919	1944-2013	1801-2003	1832-2013	1788-2013	1862-2012	1788-2013	1788-2013
Countries	88	70	88	70	14	66	43	85	86	5	51	60
Districts	79,658	69,063	79,658	77,196	5,817	19,410	6,785	37,906	19,662	39,488	68,896	11,515
Contests (N)	384,330	322,333	384,330	368,389	37,923	77,174	27,590	201,970	108,262	190,754	332,193	52,137
R2 (within)	(0.067)	(0.069)	0.337	(0.061)	(0.150)	0.464	0.430	(0.108)	(0.165)	(0.082)	(0.059)	(0.158)

These tests replicate those in Table 1, with the addition of quadratic and cubic terms. For ease of comparison, model numbers follow those in Table 1. (Models 3-4 are not replicated as they have no obvious interpretation with a cubic model.)

Outcome: Competitiveness (100 – share of largest party). *No proxies:* Electorate is not proxied by population. *D:* dummies. *Estimator:* OLS, FE (ordinary least squares regression with district fixed effects), RE (random effects), all standard errors clustered by district. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

APPENDIX B:

The Mechanical Effect

This appendix contains a proof of the mechanical effect and discussion and descriptive statistics relevant to analyses in Table 2.

The mechanical effect is the result of aggregating single-member districts that favor different political parties into one larger unit. When this happens, if each voter's preferred party remains the same, the competitiveness of the new larger district will be greater than the average of the two individual districts.

Formally, suppose there are two single-member districts, $i = \{1, 2\}$, with v_i voters in each. For simplicity, assume $v_1 = v_2$.[2] Assume there are two parties, $\{L, R\}$, and that each voter must vote for one of the two parties (full voter turnout). Let the subscript i denote the vote share of the party in each district (i.e. L_1 is the vote share of party L in district 1) and $L_i + R_i = 1$. Given that there are only two parties, the competitiveness of each district, c_i, is the vote share of the losing party: $c_i = 1 - max(L_i, R_i) = min(L_i, R_i)$. The average competitiveness of the two districts is $C = (c_1 + c_2)/2$.

Suppose the two districts are combined into one single-member district, and that the individual voters continue to support the same parties that they supported in contests at the original smaller district level. The vote shares of each party will be $L_A = (L_1 + L_2)/2$ and $R_A = (R_1 + R_2)/2$, and the competitiveness of the aggregate district is $c_A = 1 - max(L_A, R_A) = min(L_A, R_A)$.

How is the aggregate competitiveness of the single district different from the average competitiveness of the two smaller districts? First, suppose that $L_1 < R_1$ and $L_2 < R_2$. Here, party R wins both elections at the lower level, as well as the single election in the aggregate district. In this case, $c_A = (L_1 + L_2)/2$ and $C = (L_1 + L_2)/2$, such that $c_A = C$. The competitiveness of the aggregate district is the same as the average competitiveness of the two smaller districts. This is also true of the case where $L_1 > R_1$ and $L_2 > R_2$. If a majority of voters prefer the same party in both of the smaller districts and keep their vote choice constant when the districts are aggregated, there is not any mechanical effect that increases competition in the larger district. Instead, the aggregate competitiveness of the larger district is the exact same as the average competitiveness in the two smaller districts.

However, there is another case where $L_1 < R_1$ and $L_2 > R_2$.[3] Here, each party wins one of the two smaller districts. For simplicity, assume that $L_1 + L_2 > R_1 + R_2$, such that in the aggregate district

---

[2] We assume equal district population for simplicity, but the result holds with districts of different populations as well. Additionally, the mechanical effect holds when more than two districts are aggregated together into one larger district.

[3] We ignore the trivial variations where $L_1 = R_1$ or $L_2 = R_2$.

party L wins. In this case $c_A = (R_1+R_2)/2$ and $C = (L_1+R_2)/2$. Since $L_1<R_1$, this means that $c_A>C$. The same results hold in the equivalent case where $L_1>R_1$ and $L_2<R_2$. Simply by aggregating the two districts, without any voters changing the party they support, the level of electoral competitiveness increases. More broadly, this case applies whenever two (or more) lower-level districts are aggregated together where some districts support one party and some districts support the other party.

Given this mechanical result, we should expect increased electoral competitiveness whenever areas favoring different parties are combined together. On average, we would predict that national elections, aggregating some states that support one party with other states that support a second party, would be more competitive than the average state-level election. At the lower level, we should expect state elections to be more competitive than local elections. This raises a concern for our analysis: how do we prove that our hypothesized mechanisms for increasing competition are actually at work, rather than attributing the entire measured effect to the mechanical effect? To show that the mechanical effect is only one part of the total effect, we use precinct-level election data to isolate the mechanical effect by estimating the within-precinct changes in competition as the overall size of the electoral changes. This approach allows us to look at the party choices of the same groups of voters when they vote for candidates in several different and simultaneous electoral contests (treatments). Because we calculate competitiveness at the precinct level – rather than at the district/treatment level – we are able to eliminate mechanical effects that arise solely from aggregating and disaggregating voting units. Any changes in competition registered in these analyses must arise from other factors, classified as "strategic" in our theoretical discussion.

We use data from the Harvard Election Data Archive (HEDA) and the Record of American Democracy (ROAD) project (Ansolabehere et. al. 2014; King et. al. 1997). These datasets contain election data at the precinct or voting tabulation district (VTD) level – both of which are referred to as "precinct" in the text. HEDA includes election data from 2000-2012 and ROAD from 1984-1990.[4] While neither dataset is fully complete for all states, years, and contests, both include many different electoral contests at the precinct-level for many states and years, including state legislative elections, U.S. House and Senate elections, presidential elections, and state level contests such as governor, attorney general, secretary of state, and other positions depending on the state. To estimate the electorate size for each contest, we used state-level U.S. census data to estimate the population of each state in each election year. We interpolated state populations for non-census years assuming a smooth growth rate in each state between each census. For legislative elections, we divided the state population estimate in each year by the number of congressional districts apportioned to the state or the size of the state legislative chamber. We use these population estimates as proxies for the true size of the electorate (as in Table 1). We weight the observations by the total number of ballots cast in the precinct. When the number of ballots cast is unavailable, we use the largest number of votes cast for any single contest as a proxy for total ballots. In general, this is the number of votes cast for a presidential race, or for a governor or senator in non-presidential years – the race at the "top" of the ballot.

Table B1 summarizes the combined datasets by election type. Table B2 reports the number of years of data and the number of elections by type for each state in the combined dataset. We have data for multiple years for every state except Louisiana, which we exclude due to its jungle primary system. For most states, we have data from both the ROAD and HEDA archives. Table B3 lists the states and years included in our analysis.

The main results of the analysis are reported in Table 2. Table B4 provides the predicted within-precinct competition levels using Model 1 and estimated 2012 population. Table B5 provides the results from robustness checks of Model 1 in which the model is replicated with various sub-

[4] The HEDA dataset also includes data from 1996 and 1998 for Texas and Pennsylvania.

samples, including all even-numbered election years and census regions by decade. All of the results in Table B5 are significant and similar to the results obtained with the complete sample.

Table B1:
HEDA and ROAD Data Summary: Contests by Type

Election Type	N	Mean	SE
US President	7	0.3320	0.0001
US Senate	264	0.3267	0.0001
US House	3,234	0.2746	0.0001
Governor	197	0.3323	0.0002
State Senate (upper chamber)	6,248	0.2482	0.0002
State House (lower chamber)	21,460	0.2383	0.0002
State Attorney General	127	0.3338	0.0002
State Secretary of State	105	0.3380	0.0002
State Treasurer	111	0.3455	0.0002

Competitiveness (100 – share of largest party) based on two-party vote.

Table B2:

HEDA and ROAD Data Summary: Contests by State

State	USP	USS	USH	GOV	STS	STH	ATG	TRE	SOS
Alabama	6	7	77	5	105	300	4	3	5
Alaska	4	4	8	5	38	91	-	-	-
Arizona	5	5	62	4	150	30	3	4	3
Arkansas	6	6	40	6	12	65	4	3	4
California	2	3	269	3	105	404	3	3	3
Colorado	2	3	24	2	72	260	2	2	2
Connecticut	3	4	34	3	143	591	3	3	3
Delaware	5	8	10	5	127	410	4	5	-
Florida	2	3	76	3	70	333	2	3	3
Georgia	5	5	54	4	262	723	-	-	-
Hawaii	5	7	18	4	61	209	-	-	-
Idaho	3	4	12	3	26	26	2	3	-
Illinois	3	3	106	2	122	472	2	2	2
Indiana	2	3	40	2	100	-	2	2	2
Iowa	6	6	53	4	240	878	3	3	4
Kansas	3	3	28	2	82	619	2	2	2
Kentucky	4	5	40	1	76	499	1	1	1
Louisiana	-	-	-	-	-	-	-	-	-
Maine	6	7	22	5	140	729	-	-	-
Maryland	6	7	80	4	188	94	2	-	-
Massachusetts	5	6	93	4	150	619	4	3	3
Michigan	5	7	173	5	76	440	5	1	4
Minnesota	6	8	88	5	469	402	4	2	2
Mississippi	5	5	39	1	-	-	1	1	1
Missouri	4	6	71	4	278	1066	4	4	4
Montana	6	8	15	6	193	751	3	-	3
Nebraska	5	5	21	2	-	-	2	2	2
Nevada	2	3	11	3	41	166	3	3	3
New Hampshire	6	7	22	11	96	698	-	-	-
New Jersey	5	7	108	3	122	201	-	-	-
New Mexico	6	6	15	2	85	280	2	2	2
New York	3	4	136	4	244	598	3	-	-
North Carolina	6	6	133	6	380	783	5	5	6
North Dakota	5	6	10	5	105	203	5	3	5
Ohio	4	5	156	4	85	388	4	4	4
Oklahoma	5	5	45	4	75	293	3	3	-
Oregon	3	5	30	3	69	265	1	4	3
Pennsylvania	7	7	246	6	301	2409	6	6	-
Rhode Island	6	8	20	7	238	475	6	6	7
South Carolina	5	6	49	4	138	734	1	2	4
South Dakota	5	5	9	4	140	35	3	2	3
Tennessee	5	7	90	5	166	990	-	-	-
Texas	7	9	326	6	139	1095	6	1	-
Utah	3	2	12	2	57	292	2	2	-
Vermont	3	2	4	5	39	318	3	3	2
Virginia	4	5	93	2	-	-	2	-	-
Washington	4	4	60	4	100	149	3	3	3
West Virginia	3	4	19	3	85	107	2	3	2
Wisconsin	4	5	76	5	235	891	5	5	5
Wyoming	6	8	11	5	23	79	-	2	3

Offices: USP (President), USS (U.S. Senate), USH (U.S. House), GOV (Governor), STS (State Senate, upper chamber), STH (State House, lower chamber), ATG (State Attorney General), TRE (State Treasurer), SOS (State Secretary of State).

Table B3:
HEDA and ROAD Data Summary: States and Years

	ROAD				HEDA						
State	**1984**	**1986**	**1988**	**1990**	**2000**	**2002**	**2004**	**2006**	**2008**	**2010**	**2012**
Alabama	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Alaska	-	Y	Y	Y	-	Y	Y	Y	Y	Y	Y
Arizona	Y	Y	Y	Y	-	Y	Y	Y	Y	-	Y
Arkansas	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
California	-	-	-	-	-	Y	Y	Y	Y	Y	-
Colorado	Y	Y	Y	Y	-	-	-	-	-	-	-
Connecticut	Y	Y	Y	Y	-	-	-	-	-	Y	Y
Delaware	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y
Florida	Y	Y	Y	Y	-	-	-	-	-	Y	-
Georgia	Y	Y	Y	Y	-	Y	Y	Y	Y	-	Y
Hawaii	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y
Idaho	Y	Y	Y	Y	-	Y	-	-	-	-	Y
Illinois	Y	Y	Y	Y	-	-	-	-	-	-	Y
Indiana	Y	Y	Y	Y	-	-	-	-	-	-	-
Iowa	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Kansas	Y	Y	Y	Y	-	-	-	-	-	Y	Y
Kentucky[1]	Y	Y	Y	Y	-	-	-	-	Y	Y	Y
Louisiana[2]	-	-	-	-	-	-	-	-	-	-	-
Maine	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Maryland	Y	Y	Y	Y	Y	-	Y	Y	Y	Y	Y
Massachusetts	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y
Michigan	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	-
Minnesota	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mississippi[1]	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y
Missouri	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	-
Montana	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Nebraska	Y	Y	Y	Y	-	-	Y	-	Y	-	Y
Nevada	Y	Y	Y	Y	-	-	-	-	-	Y	-
New Hampshire	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
New Jersey[3]	Y	Y	Y	Y	Y	Y	Y	Y	Y	-	-
New Mexico	Y	Y	Y	Y	Y	-	Y	-	Y	-	Y
New York	Y	Y	Y	Y	-	-	-	Y	Y	Y	-
North Carolina	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
North Dakota	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y
Ohio	Y	Y	Y	Y	-	-	Y	Y	Y	Y	-
Oklahoma	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y
Oregon	Y	Y	Y	Y	-	-	-	-	Y	Y	-
Pennsylvania[4]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Rhode Island	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
South Carolina	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y
South Dakota	Y	Y	Y	Y	-	-	Y	Y	Y	Y	Y
Tennessee	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	Y
Texas[4]	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Utah	Y	Y	Y	Y	-	-	-	-	Y	-	-
Vermont	Y	Y	Y	Y	-	-	-	-	-	-	Y
Virginia[5]	Y	Y	Y	Y	-	-	-	Y	Y	Y	Y
Washington	Y	Y	Y	Y	-	-	-	-	Y	Y	Y
West Virginia	Y	Y	Y	Y	-	-	-	-	-	-	Y
Wisconsin	Y	Y	Y	Y	-	Y	Y	Y	Y	Y	-
Wyoming	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

1. Data for Kentucky and Mississippi also includes 2007.
2. Data for Louisiana is included in both datasets, but excluded in our analysis due to its jungle primary system.
3. Data for New Jersey also includes 2001, 2003, 2005, 2007, and 2009.
4. Data for Pennsylvania and Texas also includes 1996 and 1998.
5. Data for Virginia also includes 2001 and 2005.

Table B4:
Predicted Competition by Election Type

Office	Mean electorate	Mean electorate (ln)	Predicted competitiveness
State House	57,615	10.9615	0.2583
State Senate	159,028	11.9768	0.2712
U.S. House	720,926	13.4883	0.2905
U.S. Senate	6,272,052	15.6516	0.3180
U.S. President	314,000,000	19.5636	0.3678

Competitiveness (100 – share of largest party) based on two-party vote. Population based on average district sizes in 2012 using predicted state population based on 2000 and 2010 censuses. *Predicted Competitiveness* calculated using Model 1 in Table 2.

Table B5:
Precinct-Level Competitiveness by Year and Census Region

Years	**Region**	**Electorate (ln)** □ *(SE)*	**Obs** **(*N*)**	**States** **(*N*)**
1984		0.0099 (0.0023)	607,090	47
1986		0.0226 (0.0022)	580,866	48
1988		0.0152 (0.0016)	629,644	48
1990		0.0195 (0.0021)	560,309	48
1996		0.0117 (0.0012)	54,680	2
1998		0.0167 (0.0046)	55,784	2
2000		0.0130 (0.0020)	159,897	16
2002		0.0225 (0.0027)	309,529	24
2004		0.0127 (0.0016)	428,463	33
2006		0.0147 (0.0029)	449,537	33
2008		0.0086 (0.0012)	473,787	39
2010		0.0120 (0.0025)	516,270	37
2012		0.0114 (0.0014)	252,540	36
1984-1990	Midwest	0.0119 (0.0022)	934,996	12
2002-2012	Midwest	0.0111 (0.0012)	585,157	11
1984-1990	Northeast	0.0151 (0.0021)	546,837	9
2002-2012	Northeast	0.0198 (0.0021)	427,810	9
1984-1990	South	0.0172 (0.0030)	571,277	15
2002-2012	South	0.0145 (0.0018)	623,709	15
1984-1990	West	0.0127 (0.0018)	324,799	12
2002-2012	West	0.0080 (0.0011)	851,805	12

Outcome: Competitiveness (100 – share of largest party) based on two-party vote. *Electorate (ln)* defined by the population of the district corresponding to a particular contest – U.S. President, U.S. Senate, U.S. House, State Governor, State Senate (upper chamber), and State House (lower chamber). *Observations:* Precinct-level election returns for a particular contest, weighted by total ballots cast in that contest. *Estimator:* Ordinary least squares regression with precinct-year fixed effects, standard errors clustered by electoral contest. (Replicates Model 1, Table 2.)

APPENDIX C:

United Kingdom Council Elections

Data for local elections in the United Kingdom over the past century are drawn from Rallings, Thrasher & Ware (2006). They are included in full sample analyses in Table 1. They also form the centerpiece for the analysis shown in Model 1, Table 3, for which this appendix serves as a supplement.

In Table 3, we consider only council elections (not mayoral elections), and only those fought with SMD, first-past-the-post rules. Table C1 contains descriptive statistics for key variables, and Figure C1 depicts election-to-election changes in Electorate size over the observed period.

Table C2 contains results from a series of robustness tests in which Competitiveness is regressed against Electorate (ln) along with a series of covariates. Standard errors are clustered by ward, the smallest electoral unit.

Model 1 includes only county/borough and year fixed effects, replicating Model 1 in Table 3. Subsequent models introduce variations to this benchmark. Model 2 replaces county fixed effects with ward (or borough) fixed-effects. Model 3 introduces a lagged dependent variable. Model 4 adopts a first-difference model, in which changes in Competitiveness are regressed against changes in Electorate. Estimates of the causal effect are strikingly similar across all tests.

Model 5 differentiates between increases and decreases in electorate size. This analysis suggests that increases in electorate size have a stronger impact on competitiveness than decreases in electorate size. However, when the sample is limited to large changes in electorate size (greater than one standard deviation from the mean), the impact of decreases is larger than of increases, as shown in Model 6. Although limited to a smaller sample, Model 6 may offer a more appropriate test of our hypothesis since large changes in district size are generated by large and sudden changes in district composition. This sort of treatment effect, measured in a first-difference model, is less likely to be confounded by natural demographic changes or other incremental threats to inference.

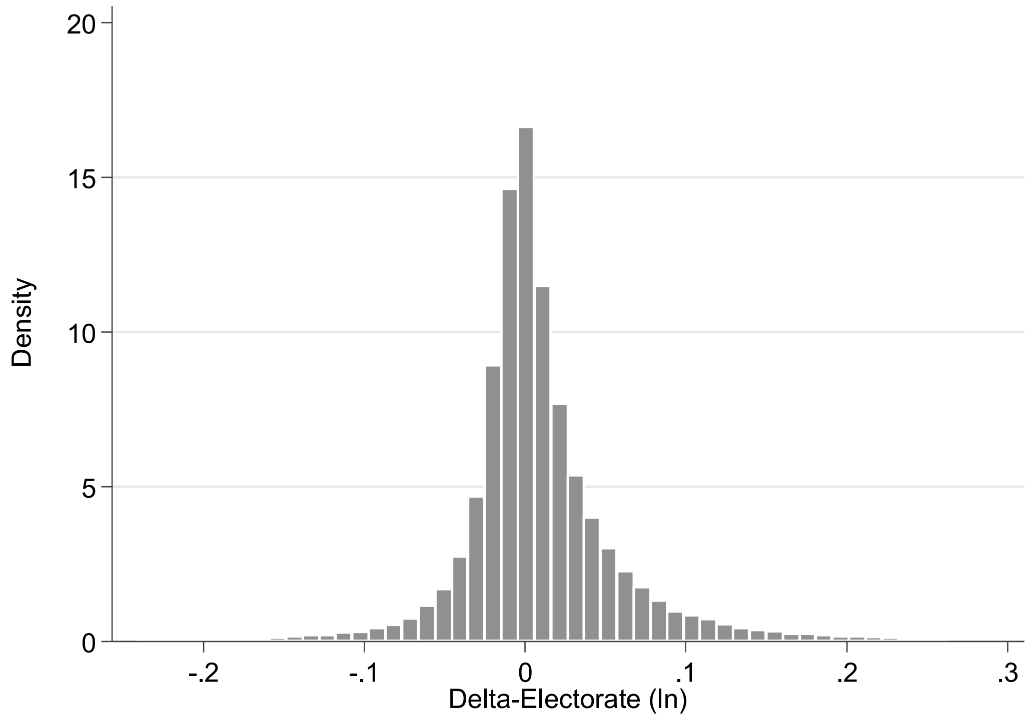
Likewise, we regard the bi-directional nature of the relationship as a strong placebo test of our hypothesis. Note that slow-moving changes such as modernization and the ongoing politicization of local government (Gyford et al. 1989) would likely be correlated with increases, but not decreases, in electorate size and with longer-term developments rather than election-to-election changes in competitiveness.

Table C1:

Descriptive Statistics (UK Council Elections)

	Obs	Mean	SD	Min	Max
Competitiveness	121,378	41.80	11.92	0.00	79.20
Electorate (ln)	121,378	8.53	0.76	4.43	11.74
ΔElectorate (ln)	96,555	0.01	0.08	-2.36	2.83

Figure C1:
Ward-level Changes in Electorate Size (UK Council Elections)



Histogram of ΔElectorate (natural logarithm) from election to election in British council elections (*N*=94,805), excluding extreme outliers (SD > 3.0, *N*=1,750).

Table C2:
Full Results (UK Council Elections)

	1	2	3	4	5	6
Outcome	Y	Y	Y	ΔY	ΔY	ΔY
Estimator	RE	OLS, FE	OLS, FE	OLS	OLS	OLS
Electorate (ln)	2.843*** [0.125]	2.874*** [0.428]	2.315*** [0.420]			
ΔElectorate (ln)				3.120*** [0.511]		
□ **Electorate (ln)**					4.252*** [0.669]	2.928*** [0.956]
□ **Electorate (ln)**					-1.364 [0.871]	-4.186*** [1.081]
Y_{t-1}			0.175*** [0.005]			
County/borough (D)	X					
Ward/district (D)		X	X			
Year (D)	X	X	X	X	X	X
Sample	Entire	Entire	Entire	Entire	Entire	Strong treatment
Years	1912-2003	1912-2003	1913-2003	1913-2003	1913-2003	1919-2003
Wards/municipalities	24,823	24,823	19,247	19,247	19,247	6,336
Contests (N)	121,378	121,378	96,555	96,555	96,555	9,950
R2 (within)	0.097	(0.040)	(0.075)	0.035	0.036	0.036

Outcome: Competitiveness (100 – share of largest party). *Strong treatment:* ΔElectorate (ln) > 1 standard deviation from the mean. *D:* dummies. *Estimators:* OLS, FE (ordinary least squares regression with district fixed effects), RE (random effects), standard errors clustered by district (ward). *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests). (Model 1 replicates Model 1, Table 3.)

APPENDIX D:

Brazil

Brazil possesses an enormous variety of elective offices and tremendous variation in district size across contests for those offices. Electoral data is drawn from the Superior Electoral Court (Tribunal Superior Eleitoral), the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatistica), and Brambor & Ceneviva (2012). This data is included in the full sample analyses in Table 1. Data on council elections form the centerpiece for the analysis shown in Model 2, Table 3. This appendix offers robustness tests that are inclusive (including all available Brazilian elections) as well as tests that focus on council elections.

The distribution of data across all available Brazilian elections is provided in Table D1, variable definitions in Table D2, descriptive statistics in Table D3, and estimations in Table D4.

Most of Brazil's legislative elections feature MMDs with open lists, allowing for preferential voting within a list. This is measured in our analysis as a residual category, i.e., all non-SMD elections. Some mayoral elections (those where the municipal population surpasses 200,000) employ a two-round voting system, a feature that we also control for in the analyses.

Following the template of our benchmark tests, we regress Competitiveness against Electorate with a series of covariate controls. Models 1-3 in Table D4 incorporate data from all available elective offices from 1945-2010. This includes the period of military rule. However, most of the observations are drawn from the past several decades, after Brazil's democratic transition. (Note that the scope-conditions of our theory presume that multi-party competition is allowed but do not require that elections be entirely free and fair.) Model 1 includes only electoral system variables, along with state and year fixed-effects. Model 2 introduces a series of additional controls measuring urbanization, income, and literacy. Model 3 adds Land area (logged).

Models 4-6 focus on council elections, and thus serve as robustness tests for Model 2 in Table 3. These analyses follow the same format as the previous analyses but without electoral system controls (which are redundant).

Models 7-9, focused on mayoral elections, follow the same format with a control variable that measures the second round of a two-round election.

All tests indicate that district size has a modest, but statistically significant, impact on competitiveness. Covariates behave as expected, with the exception of income, which is negatively signed. We have no theoretical prior about the behavior of Land area.

Table D1:
Data Description (Brazil)

	Electorate (1000s)			**Competitiveness**			**Coverage**		
	Min	**Max**	**Mean**	**Min**	**Max**	**Mean**	**Years**	**Districts**	**Contests**
Offices									
Mayoral	1	11,090	31	0	77.3	44.6	1996-2008	5,560	22,003
Council	1	11,090	31	0	90.5	63.9	1996-2010	5,563	20,359
Governor	372	42,108	7,053	17.2	49.9	42.3	2006-2010	27	54
National lower	2	28,037	2,283	0	87.5	60.4	1945-2006	34	317
National upper	336	42,108	6,720	11.2	80.1	55.1	1998-2010	27	108
Eras									
1945-1990				0	85.7	54.1			209
1991-2000				0	89.5	52.6			20,759
2001-2010				0	90.5	55.1			21,872
Total	1	42,108	74	0	90.5	53.9	1945-2010	11,211	42,841

Empty cells = data not relevant.

Table D2:
Variable Definitions (Brazil)

District magnitude (ln). Number of seats per distract, natural logarithm. Source: Brazilian Superior Electoral Court (Tribunal Superior Eleitoral) *dm_ln*

Income per cap (ln). Personal income per capita, natural logarithm. Source: Brambor and Ceneviva (2012), Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatistica). *incomepc_ln*

Literacy. Percent of people above 15 who are literate. Source: Brambor and Ceneviva (2012), Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatistica). *Literacy*

Urban. Urban population as percent of total. Source: Brambor and Ceneviva (2012), Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatistica). *urban_perc*

Includes only variables specific to Brazil. For other variable definitions see Table A3.

Table D3:

Descriptive Statistics (Brazil)

Variable	Obs	Mean	SD	Min	Max
Electorate (ln)	42541	9.428	1.230	7	18
SMD	42841	0.516	0.500	0	1
2d round	42841	1.004	0.060	1	2
Income per cap (ln)	42057	6.203	1.278	1	12
Land area (ln)	42305	8.821	0.669	6	12
Literacy	42305	78.976	12.549	27	99
Urban	42305	59.819	23.048	0	100
District magnitude (ln)	42524	1.096	1.150	0	4

Table D4:
Full Results (Brazil)

Offices	All			Council			Mayor		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Electorate (ln)	2.799***	2.111***	1.954***	3.645***	2.063***	1.872***	1.970***	1.698***	1.516***
	[0.075]	[0.094]	[0.107]	[0.096]	[0.143]	[0.156]	[0.102]	[0.120]	[0.138]
SMD	-8.377***	-2.807	-7.867***						
	[2.520]	[3.068]	[1.195]						
2d round	-7.583***	-6.797***	-7.705***				-7.109***	-6.939***	-6.610***
	[0.821]	[0.798]	[0.796]				[0.715]	[0.718]	[0.723]
District magnitude		4.878***	4.896***		9.126***	9.187***			
(ln)		[0.512]	[0.514]		[0.697]	[0.693]			
Urban		0.033***	0.037***		0.039***	0.043***		0.026***	0.030***
		[0.005]	[0.005]		[0.007]	[0.007]		[0.006]	[0.006]
Income (ln)		-0.943***	-0.917**		-1.199**	-1.272**		-0.932*	-0.977**
		[0.347]	[0.376]		[.529]	[0.530]		[0.498]	[0.498]
Literacy		0.092***	0.099***		0.142***	0.152***		0.061***	0.070***
		[0.015]	[0.016]		[0.023]	[0.024]		[0.020]	[0.020]
Land area (ln)			0.357***			0.399***			0.358***
			[0.086]			[0.121]			[0.114]
Office (D)	X	X	X						
State (D)	X	X	X	X	X	X	X	X	X
Year (D)	X	X	X	X	X	X	X	X	X
Years		1945-2010			1996-2010			1996-2008	
Districts	11,099	11,068	11,009	5,510	5,505	5,505	5,504	5,504	5,504
Contests (N)	42,540	42,223	42,056	20,219	20,219	20,214	21,842	21,842	21,842
R2	0.445	0.447	0.448	0.259	0.270	0.270	0.048	0.050	0.051

Outcome: Competitiveness (100 – vote-share of largest party). *Offices:* lower house, upper house, governor, council, mayor. *D:* dummies. *Estimator:* random effects, standard errors clustered by district. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests). (Model 5 replicates Model 2, Table 3.)

APPENDIX E:

Swedish Council Elections

Electoral data from Sweden is included in the full-sample tests in Table 1. Data for council elections form the basis for Model 3 in Table 3. This appendix serves a supplementary role for the latter.

The historical background to Swedish council elections and their consolidation in the 1960s and 1970s is laid out in Wallin (1973). Recall that the larger, unified districts generated by the series of mergers serve as our units of analysis. This means that the competitiveness of pre-merger units is combined – using population-weighted means – so as to produce a balanced panel with consistent boundaries across three elections: 1966, 1970, and 1973.

Model 1 in Table E2 replicates Model 3 in Table 3, with district and year fixed effects and a control for district magnitude.[5] Subsequent tests explore variations in this benchmark model. Model 2 removes the district magnitude control. Model 3 adds a lagged dependent variable. Model 4 adopts a first-difference estimator. In all instances, the effect of electorate size is positive and statistically significant, though the magnitude of the effect is attenuated relative to estimates for UK local elections (see Model 1, Table 3) and our global sample (see Table 1).

One possible explanation for this attenuation has to do with the unusual dominance of the Social Democratic party during this period of Swedish history.[6] The Social Democrats were the largest party in 65% of the municipalities in 1966, 78% in 1970, and 80% in 1973. As a result, and because Social Democratic strongholds tended to be in neighboring districts, the aggregation of municipalities did not generally produce a mechanical effect. (When joining two districts in which the same party enjoys the pole position, any increase in competitiveness must be the product of changes in voter behavior, as discussed.) To account for this, Model 5 introduces a dummy variable coded 1 for all districts in which the Social Democrats were the largest party in the previous election, along with the interaction effect of this variable and Electorate. This model demonstrates that the impact of Electorate on Competitiveness is substantially increased in districts where the Social Democrats were not dominant (dummy=0), corroborating our expectation and confirming the importance of mechanical effects in a MMD context. (Previous tests of the mechanical effect, contained in Table 2 and Appendix B, focus on SMD contests in the United States.)

---

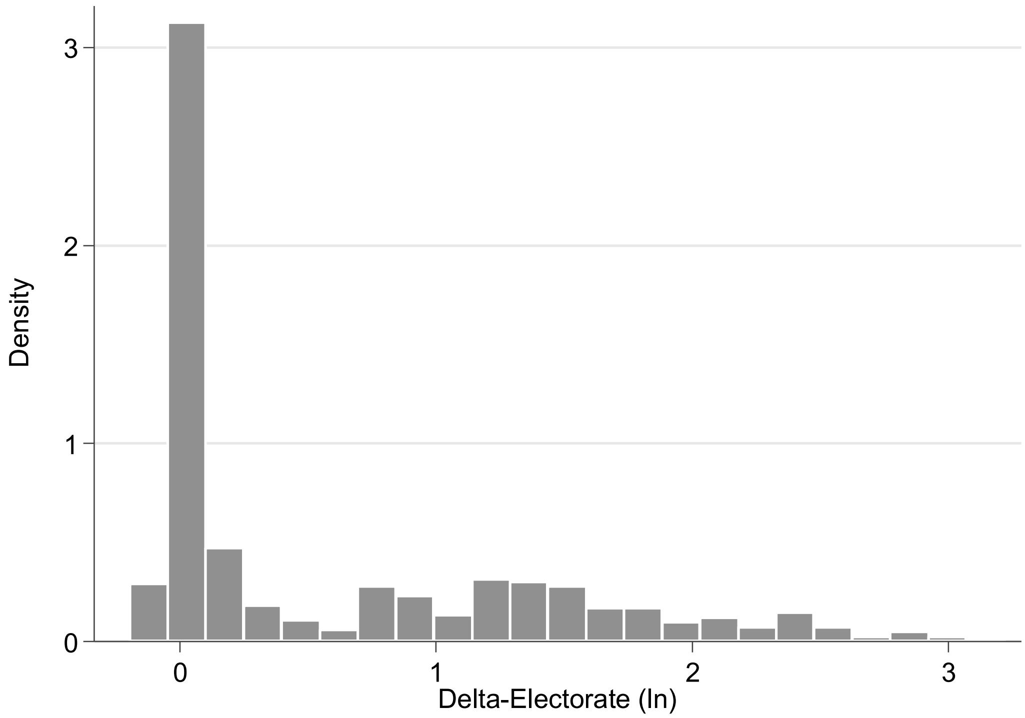
[5] Strictly speaking, the control variable is assembly size, not district magnitude. The absolute majority of municipalities before the merger however held their elections in a single district; in 1970 this occured in 87 % of the municipalitites, in 1973 in 78 % (Wångmar 2006: 81-84).

[6] It is rare for a single party to dominate party competition in free and fair elections for such an extended period of time (Pempel 1990).

Table E1:
Descriptive Statistics (Swedish Council Elections)

	Obs	Mean	SD	Min	Max
Competitiveness	834	54.27	7.910	23.53	72.890
Electorate (ln)	834	8.922	0.965	7.213	13.250
ΔElectorate (ln)	556	0.601	0.802	-0.197	3.220

Figure E1:
Municipality-level Changes in Electorate Size (Swedish Council Elections)



Histogram of ΔElectorate (natural logarithm) from election to election in Swedish council elections from 1966 to 1973 (*N*=556).

Table E2:
Full Results (Swedish Council Elections)

	1	2	3	4	5
Outcome	Y	Y	Y	ΔY	Y
Estimator	OLS, FE	OLS, FE	OLS, FE	OLS	OLS, FE
Electorate (ln)	2.212*** [0.787]	1.227*** [0.332]	1.098** [0.488]		2.372*** [0.773]
ΔElectorate (ln)				1.266*** [0.300]	
Y_{t-1}			-0.187*** [0.057]		
District magnitude (ln)	-4.050 [2.745]				
Social Dems largest party, t-1					17.521** [7.297]
Electorate(ln)* Social Dems					-1.822** [0.771]
District (D)	X	X	X		X
Year (D)	X	X	X	X	X
Years	1966-1973	1966-1973	1970-1973	1970-1973	1970-1973
Wards/municipalities	278	278	278	278	278
Contests (N)	834	834	556	556	556
R2 (within)	(0.118)	(0.112)	(0.269)	(0.140)	(0.245)

Outcome: Competitiveness (100 – share of largest party). *D:* dummies. *Estimators:* OLS, FE (ordinary least squares with district fixed effects), RE (random effects), standard errors clustered by district. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests). (Model 1 replicates Model 3, Table 3.)

APPENDIX F:

United States

The United States offers the greatest number of election types and the greatest variation in district size of any multi-party democracy. Our sample brings together election contests from Senate, gubernatorial, upper state house, lower state house, mayoral, and local council elections – a total of 6,809 districts and 90,957 district-level contests from 1792 to 2010. This data is included in full sample analyses in Table 1 as well as US-focused analyses in Table F4.

Electoral data used in calculating Competitiveness is drawn from the Inter-university Consortium for Political and Social Research (1994) as well as a variety of other sources depending upon the office: (a) US Senate (the Office of the Clerk Election Statistics[7]), (b) state upper and lower house (Klarner et al. 2013), (c) Governor (Parker 2010), and (d) municipal (Trounstine 2008; Ferreira & Gyourko 2009; Gerber & Hopkins 2011). Population data is drawn from decennial US Census reports,[8] with missing values within a time-series interpolated.

Regression analyses employ several covariates judged to be important and exogenous influences on the outcomes of interest. This includes *Urban* (percent living in urban areas), *Income per capita* (natural logarithm), *High school* (percent above age 25 with a high school degree), and *College* (percent above age 25 completing college). Some of these covariates are treated as constants since they change little over the observed period; these are taken from the 2000 US Census records. Historical data (at decadal intervals) is available for states from the US Census[9] and the Bureau of Economic Analysis.[10]

The distribution of the data across elections is provided in Table F1, variable definitions in Table F2, and descriptive statistics in Table F3.

Table F4 introduces a set of tests where Competitiveness is regressed against Electorate along with year fixed effects and other covariates. We adopt an ordinary least squares estimator wherever district fixed effects are employed and a random effects estimator in other instances, consistent with our practice elsewhere.

The sample for Model 1 includes elections to all available offices, as listed above, though the vast majority are (upper and lower house) state legislative offices. The specification includes district

---

[7] http://clerk.house.gov/member_info/electionInfo/index.html
[8] http://www.census.gov/main/www/cen2000.html
[9] Ethnicity data from www.census.gov/population/www/documentation/twps0029/tab08.html, education data from www.census.gov/hhes/socdemo/education/data/census/half-century/tables.html (Tables 5 and 6, both sexes), and other data from selected US Census documents.
[10] For income data, see www.bea.gov/national/nipaweb/SelectTable.asp?Selected=N, Table 1.1.

and year fixed-effects. The estimated causal effect of Electorate on Competitiveness in this model is positive but not significant. This is probably because district identifiers (the basis of unit fixed effects) for US House and state legislative districts are recalibrated every 10 years, in accordance with decennial reapportionment. This means that there is very little year-to-year variation in population per unit – especially for state legislative districts, which (as noted) comprise the vast majority of observations.

Model 2 replaces district fixed-effects with state fixed-effects. The estimated coefficient for Electorate is now significant and similar to our main results in Table 1. Model 3 adds a set of additional covariates that may serve as confounders including Urban, Income, High school, and College, as described above. Model 4 includes the same covariates as Model 3, but restricts the time period to elections since 1970. In this time period district population, which we use as a proxy for electorate size, is a more accurate proxy because it reflects suffrage practices in the United States following the implementation of the Voting Rights Act of 1965.

Models 5-7 are limited to elections for statewide offices, i.e., Senate and Governor. Model 5 includes only state and year fixed-effects, along with the variables of interest. (Note that in this context districts are equivalent to states.) Model 6 adds several additional covariates – Urban, Income, High school, and College – which also limits the temporal range of the analysis due to limited data coverage for the additional covariates. We find the estimated causal effect of Electorate on Competitiveness is enhanced, rather than attenuated, when these covariates are added. Model 7 restricts the sample to the post-1970 period and adds Land area (square miles, logged) as a control, removing state fixed effects (to avoid perfect collinearity). The coefficient for Electorate remains positive, though the effect is diminished relative to Model 6. Land area shows no apparent relationship to the outcome.

Model 8 is limited to elections for the U.S. House of Representatives and uses year and state fixed effects. The coefficient on electorate is significant and similar to our benchmark results in Table 1.

Overall, the results for the United States support our thesis, though results vary across samples, offices, and model specifications.

Table F1:
Data Description (US)

	Electorate (1000s)			Competitiveness			Coverage		
	Min	Max	Mean	Min	Max	Mean	Years	Districts	Contests
Offices									
Local council	16	744	367	0	68.3	33.7	1914-1999	39	473
Mayoral	.08	8,077	113	0	88.7	35.9	1853-2007	959	7,303
Lower state house	7	612	62	0	80.2	26.6	1968-2010	7,059	84,772
Upper state house	11	1,142	139	0	67.1	28.6	1968-2010	2,359	23,839
Lower nat'l house	25	999	314	0	92.5	34.9	1788-2008	586	33,364
Governor	21	33,100	2,233	0	80.1	41.8	1792-2000	50	2,770
Senate	50	33,871	3,610	0	69.9	39.0	1902-2000	100	1,490
Eras									
1788-1899				0	80.1	42.9			1,157
1900-1947				0	69.9	39.3			1,643
1948-69				0	88.7	33.2			6,071
1970-79				0	88.2	29.9			24,284
1980-89				0	82.8	26.9			26,553
1990-1999				0	85.2	27.3			28,579
2000-2008				0	83.3	26.3			32,374
Total	.08	33,871	235	0	92.5	29.6	1788-2010	11,152	154,011

Empty cells = not relevant. Note: describes all observations with competitiveness data.

Table F2:
Variable Definitions (US)

College.	Percent of electorate at least 25 with bachelor's degree. Source: US Census. *college*
High school.	Percent of electorate at least 25 with high school degree. Source: US Census. *highschool*
Income per cap (ln).	Personal income per capita, natural logarithm. Source: Bureau of Economic Analysis. *incomepc_ln*
Land Area (ln).	State land area in square miles, logged. Source: US National Atlas (http://www.nationalatlas.gov/articles/mapping/a_general.html). *area_sqm_ln*
Urban.	Urban population as percent of total. Source: US Census. *urban_perc*

Includes only variables specific to the US. For other variable definitions see Table A3.

Table F3:

Descriptive Statistics (US)

Variable	Obs	Mean	SD	Min	Max
College	33,753	23.174	12.834	1.9	84.0
High school	33,753	78.870	14.852	7.8	99.7
Income per cap (ln)	47,292	9.325	1.277	4.8	11.7
Urban	36,047	69.532	30.451	0.0	100.0
Land area (ln) (for states only)	4,281	25.266	1.208	21.8	28.0

Table F4:
Complete Results (US)

Offices	All				Senate & Governor			House
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Estimator	OLS, FE	RE	RE	RE	OLS, FE	OLS, FE	RE	RE
Electorate (ln)	0.853 [0.551]	2.040*** [0.170]	4.720*** [0.390]	4.731*** [0.386]	0.386 [0.888]	5.370** [2.330]	1.524** [0.674]	2.718** [1.314]
Urban			-0.033*** [0.005]	-0.036*** [0.005]		0.174** [0.084]	-0.064 [0.044]	
Income			9.862*** [1.019]	7.949*** [0.960]		33.491*** [5.209]	9.803** [4.858]	
High School			0.275*** [0.023]	0.299*** [0.023]		-0.147 [0.145]	0.329*** [0.120]	
College			-0.211*** [0.022]	-0.183*** [0.022]		-0.328 [0.274]	-0.283* [0.161]	
Land area							0.070 [0.564]	
Year (D)	X	X	X	X	X	X	X	X
Office (D)		X	X	X			X	
State (D)		X	X	X				X
District (D)	X				X	X		
Years	1788-2010	1788-2010	1940-2010	1970-2010	1792-1999	1940-1999	1970-1999	1788-2008
Districts	28,309	28,309	11,299	11,299	100	100	100	7,229
Contests (N)	124,207	124,207	33,753	32,734	4,260	1,987	968	33,250
R2 (within)	(0.022)	0.234	0.210	0.200	(0.151)	(0.206)	0.110	0.350

Outcome: Competitiveness (100 – vote share of largest party). *Offices:* House of Representatives, Senate, Governor, state lower house, state upper house, mayor, and city council. *D:* dummies. *Estimators:* OLS, FE (ordinary least squares regression with district fixed effects), RE (random effects), standard errors clustered by district. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

APPENDIX G:

Miscellaneous Threats to Inference

Threats to inference in the foregoing analyses are posed by several potential confounders. These may be classified broadly as (a) time trends, (b) office power, (c) modernization, and (d) districting. To guard against threats to inference from these confounders – and, where possible, to estimate their independent effect on Competitiveness – we employ a variety of tests, summarized in Table G1. Each will be discussed at some length in this appendix.

Trends

Time trends are a persistent concern in panel analysis (Wooldridge 2010), especially where left- and right-side variables share common trends. While the sample mean of Competitiveness is constant over time (Table A3), this is largely a product of our changing sample, which adds countries and elections over time. Fixed-effect analysis reveals an upward trend in Competitiveness. Likewise, the size of electorates tends to grow over time due to demographic factors (see Figure A6).

We take several approaches (sometimes in tandem) to deal with the potentially confounding effects of common time-trends. First, we impose annual fixed effects, a unique intercept for each year in the global analysis (Table 1) and in country-specific analyses (elsewhere). Second, we include a lagged dependent variable in the global analysis (Model 2, Table 1) and in country-specific analyses for the United Kingdom (Model 3, Table C2) and Sweden (Model 3, Table E2). Third, we employ difference-in-difference models in the global analysis (Model 3, Table 1) and in country-specific analyses focused on the United Kingdom (Model 4, Table C2) and Sweden (Model 4, Table E2). Fourth, we employ a bi-directional causal test, in which we decompose Electorate into increases and decreases in the global analysis (Model 4, Table 1) and the country-specific analysis focused on the United Kingdom (Models 5-6, Table C2). Fifth, we employ precinct and year fixed-effects in the analysis of the mechanical effect in the United States (Table 2). Finally, we analyze the short-term effect of historical suffrage extensions (Table 4). Whatever trend effects are present in the data generating process should be controlled in these various tests, all of which show a significant relationship between Electorate and Competitiveness.

Office power

Consider the possibility that differences in competitiveness across large and small districts are driven by the fact that more powerful offices often have larger districts. Note that elections to a powerful office are likely to attract greater attention from high-quality challengers, party organizations, and the media. These factors, in turn, may be responsible for greater competitiveness, generating a spurious association between size and competitiveness.

However, this sort of confounding does not affect (a) comparisons across districts for the same office or (b) comparisons through time for the same district (e.g., models with district fixed effects). Since these formats predominate among results presented in previous tables, we can be fairly certain that our finding is not driven by differences in power across offices.

It is interesting nonetheless to explore the hypothesis that variation in political power across offices might impact the electoral competitiveness to those offices. One approach, adopted in Table G2, focuses on estimated coefficients for each office type in the pooled sample – governor, upper chamber of national legislature, lower chamber of national legislature, upper chamber of state legislature, lower chamber of state legislature, mayor, and council. Each office is represented by a dummy variable, with council elections as the excluded category, and electoral system, country, and year fixed effects. We employ a cross-sectional format – ordinary least squares without district fixed effects – since there is no variation through time in the regressors of interest. If the office power hypothesis is correct, more powerful offices should be characterized by greater competitiveness.

The analysis, shown in Model 1, Table G2, provides some support for the idea. For example, we find that national legislative races are more competitive than state legislative races, and gubernatorial races are more competitive than mayoral races. However, we also find that city council elections – the omitted category – are more competitive than upper chamber national legislative races, both upper and lower state legislative races, and mayoral races. This somewhat peculiar result may be a product of our sample, where non-parliamentary elections are drawn from a small handful countries (e.g., the US, UK, and Brazil). Likewise, there are many potential confounders. For these reasons, we do not regard the patterns evident in Table G2 as conclusive.

A better approach to the power hypothesis focuses on settings where members elected from the same district (during the same election) wield asymmetric powers. In this fashion, we may compare US Senate and House elections in states that were (for some period of the 20^{th} century) allocated only one House seat. This analysis (along the lines of Nice 1984) is conducted by matching exactly on state and year – hence, comparing the same election within the same state. Results, shown in Model 1 of Table G3, show no (statistically significant) relationship between the power of an office and its level of competitiveness. Specifically, elections to the more powerful position (Senator) are no more competitive than elections to the less powerful position (Representative) when both elections occur simultaneously in the same state.

Modernization

A third potential confounder stems from factors associated with modernization such as income, education, and urbanization. These factors are likely to be correlated with increases in district size through time and may also have direct causal effects on contestation. Likewise, rural-to-urban migration tends to create large districts with "modern" characteristics and small districts with "premodern" characteristics. This suggests that modernization may serve as an unmeasured confounder in longitudinal as well as cross-sectional analyses.

Analyses of proximal causal effects arising from suffrage extensions (see Table 4) are not subject to confounding from sluggish factors associated with modernization. This suggests that modernization cannot account for all of the effects picked up by our global analyses (in Table 1).

Regression models that include unit and time fixed-effects or matching analyses that match on districts and time should mitigate modernization effects. However, they do not entirely eliminate them, so it is worth testing the thesis directly wherever possible. Our analysis of data from Brazilian elections (Model 2, Table 3, and Table D4) suggests that measures of urbanization and literacy – but not income – exert a slight impact on competitiveness and no impact on the estimated coefficient for Electorate. Analyses focused on the United States (Table F4) suggest that income – but not urbanization or education – play a strong role in conditioning levels of competitiveness across districts. However, they do not appear to serve as confounders in the analysis, as the coefficient for Electorate is enhanced when modernization factors are included in various specification tests.

Although we are unable to test these factors for other countries we see no reason to suppose that Brazil and the United States are unrepresentative in this regard.

Another opportunity to test this confounder arises where similar legislative bodies have overlapping constituencies. In this setting, voters are subject to treatment and control conditions sequentially, as they move down a ballot; that is, the same elector casts a vote in a large district as well as in a smaller district. This is similar to the within-subjects design employed in Table 2. If all other factors across the two electoral choices are the same, and these treatments do not interfere with each other, causal inference is possible and a matching algorithm is appropriate to estimate the effect.

Political settings of this sort are rare. Bicameral legislatures are common but they usually possess unequal powers or operate under different electoral rules. Fortuitously, both the United Kingdom and the United States offer opportunities for simultaneous treatment/control tests, as displayed in Table G4.

Since the 1972 reorganization of local government in the UK, voters in many parts of the country have elected two overlapping tiers of local government, each represented by a council whose members are drawn from single-member districts (wards). We exclude mayoral elections, multimember districts, and elections that do not utilize first-past-the-post rules. We also exclude elections occurring within two months of a national election (for the House of Commons), as the latter may exert a confounding influence on the results. Upper tier councils employ wards that are larger – on average, almost twice as large – as the lower tier councils. An upper tier election is therefore regarded as the treatment and the lower tier as the control condition. Importantly, the two levels of government operate independently of one another; in no sense is the lower tier subordinate to the upper tier. Their tasks, while different and sometimes in conflict, are complementary (Alexander 1982: 12).

Our first analysis matches exactly on two attributes of the ward-level contest: (a) the borough or district and (b) the five-year period within which a set of elections take place. These exact matches are used to provide a composite control with the CEM algorithm (Blackwell et al. 2010). This is followed by OLS analysis on the matched units. Our second analysis follows the first except that the timing of the election is understood as a continuous covariate, modeled with the nearest-neighbor algorithm (Abadie et al. 2001). Coefficients are understood as sample average treatment effects (SATE).[11]

Both analyses show a positive and statistically significant treatment effect, though the estimated effect is somewhat larger in the second analyses. For our purposes, the important result is that elections to upper-tier councils are more competitive than elections to lower-tier councils that take place within the same borough (or district) and in the same time-period.

In the United States, every state except Nebraska is governed by a bicameral legislature. These chambers are roughly equal in power and are usually elected simultaneously from single-member districts (though many senators serve longer terms and thus are elected in a staggered fashion). (Thirteen states using MMDs in lower house elections are excluded from the following analyses.) Since Baker v. Carr (1962), these districts are constrained to be roughly equal in size; that is, the lower (or upper) house districts within a state have approximately the same number of constituents, subject to decennial reapportionments. The main difference between state senate and house elections is that upper houses are smaller and thus members are drawn from districts that are larger than those employed for lower house seats. The difference in size, as in the UK analyses, is slightly more than two to one.

---

[11] Population average treatment effects are almost identical. Likewise, when the number of minimum matches is increased there is only a slight change in estimated coefficients and standard errors.

To analyze the impact of size on contestation, we repeat the two matching analyses described above. Here, the treatment condition is the senate election and the control condition is the house election. First, we match elections exactly on year and state, providing a composite control using the CEM algorithm (Blackwell et al. 2010). This is followed by OLS analysis on the matched units, which in this case constitute 100% of the sample. A second analysis applies exact matching to state and nearest-neighbor matches to other covariates – Year, Urban, Income, High School, and College – with a minimum of a single match (Abadie et al. 2001).

In both analyses, larger districts show a (statistically significant) relationship to Competitiveness. The estimated treatment effect is virtually identical in both analyses.

The strength of these analyses rests on the high degree of equivalence across treatment and control conditions. Subjects (voters) are exposed, simultaneously, to both conditions. There is no need for randomization so long as we can assume noninterference across treatment and control conditions – a fairly safe assumption in this instance, we think. Not much distinguishes upper and lower tier elections in British municipalities or American states except the size of the districts. It is not the case that upper houses are uniformly more powerful than lower houses, for example. Naturally, the power of an *individual* legislator is greater in a smaller body, an issue discussed in the previous section.

Districting

A final potential confounder concerns the way in which districts are created. If smaller districts are more amenable to gerrymandering, and if gerrymandering is generally employed to minimize competition, our results may reflect strategic line-drawing. Let us explore these possibilities.

In some instances, such as the US Senate, units are fixed by constitutional fiat and cannot be adjusted. Districts of the US House of Representatives, by contrast, are allowed to vary (and indeed, must be adjusted after every census). If constitutionally fixed districts are generally larger than "varying" districts, the estimated relationship between district size and competitiveness may be a product of bias introduced by gerrymandering. However, only a small portion of our global sample is fixed in this special sense so it is unlikely to be driving results shown in Table 1.

Importantly, fixed-effect analyses that focus on longitudinal change within a district (our benchmark model) are not subject to this confounder. Likewise, analyses focused on MMD contests (e.g., Model 14, Table 1), are not subject to this confounder for the simple reason that gerrymandering is extremely rare in multimember districts, for strategic and customary reasons (Taagepera & Shugart 1989). These analyses offer what is perhaps the strongest evidence that our results are not driven by strategic districting.

The only sort of analysis that is subject to the selective gerrymandering confounder is one that compares competitiveness across different sorts of elective districts within a country (e.g., Models 5, 8, & 9, Table 1). Here, one must also consider the possibility that strategic districting serves partisan purposes rather than incumbency protection (Ansolabehere & Snyder 2012, Gelman & King 1994). Insofar as this is true, gerrymandering may *increase* competitiveness overall (as measured by our preferred indicator), rather than decrease it. In any case, recent studies suggest that the problem of uncompetitive districts – at least in contemporary American politics – owes more to demographic sorting than to redistricting (Chen & Rodden 2013).

Table G1:

Potential Confounders and Identification Strategies

CONFOUNDERS AND STRATEGIES	IMPLEMENTATION
Trends	
• Year fixed effects	Virtually all models
• Lagged DV	Model 2/Table 1; Model 3/Table C2; Model 3/Table E2
• Difference-in-difference	Model 3/Table 1; Model 4/Table C2; Model 4/Table E2
• Bi-directional causal test □ & □ Electorate	Model 4/Table 1; Models 5-6/Table C2
• Precinct and year fixed-effects	Table 2
• Suffrage extensions	Table 4
Power	
• District fixed effects	Most models
• Country & Office districts	Model 5/Table 1
• Analyses focused on the same offices	Models 11-12/Table 1; Table 3; Table 4; Table C2; Models 3-6/Table D4; Table E2; Models 8-14/Table F4
• Compare electoral results from districts that simultaneously elect differentially powerful offices	Model 1/Table G3
Modernization	
• District and year fixed-effects	Model 1/Table 1 and elsewhere
• Bi-directional causal test □ & □ Electorate	Model 4/Table 1; Models 5-6/Table C2
• Precinct and year fixed-effects	Table 2
• Condition on covariates	Model 2/Table 3; Models 2, 4, 6/Table D4; Models 3-4, 6-7, 9, 11/Table F4
• As-if random district consolidation	Model 3/Table 3
• Suffrage extensions	Table 4
• Simultaneous treatment and control tests	Table G4
Districting	
• District fixed effects	Benchmark models
• MMD elections only	Model 14/Table 1
• Precinct and year fixed effects	Table 2

Table G2:
Offices

	1
Governor	7.652***
	[1.045]
Upper chamber (national)	4.961***
	[1.174]
Lower chamber (national)	0.682***
	[0.219]
Upper chamber (state)	-7.286***
	[0.313]
Lower chamber (state)	-9.888***
	[0.283]
Mayor	-0.322
	[0.379]
Council	[omitted]
Electoral system (D)	X
Year (D)	X
Country (D)	X
Years	1788-2013
Countries	88
Districts	89,917
Contests (N)	415,095
R2 overall	0.320

Outcome: Competitiveness (100 – vote share of largest party). All right-side variables are dummies. *D:* dummies. *Sample:* full MLEA dataset. *Estimator:* ordinary least squares, standard errors clustered by district. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

Table G3:
Power

Treatment condition	Senate
Control condition	House
Matching covariates	State, Year
Exact matches	100%
Treatment effect	1.415
(SATT)	[0.890]

Outcome: Competitiveness (100 – share of largest party). *Units of analysis:* US Senate and House elections in states with one House member. *States:* 20. *Years:* 1902-2000. *Contests (N):* 400. *Analysis:* Exact matching (1:1) on district and year using CEM (Blackwell et al. 2010), followed by OLS analysis of matched observations. *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

Table G4:
Simultaneous Treatment/Control Tests

	1	**2**
Sample of elections	**UK local council**	**US state house**
Years	1973-2003	1972-2010
Contests (N)	77,219	78,482
Treatment group	**Upper council**	**Upper house**
Contests	30,095	17,684
Mean electorate	8,149	138,530
Control group	**Lower council**	**Lower house**
Contests	47,124	60,798
Mean electorate	3,501	59,985
Exact matching		
Treatment effect (SATT)	2.743 [0.099] ***	2.239 [0.180] ***
Covariates (exact)	District, 5-year intervals	State, Year
Exact matches	87%	85%
Nearest-neighbor matching		
Treatment effect (SATE)	3.866 [0.103] ***	2.171 [0.299] ***
Covariates (exact)	District	State
Exact matches	89%	94%
Covariates (NN)	Year	Year, Urban, Income, High school, College

Outcome: Competitiveness (100 – vote share of largest party). *Estimators:* exact matching using CEM (Blackwell et al. 2010) followed by OLS analysis of matched observations; nearest neighbor (NN) matching (Abadie et al. 2001). *** p<0.01, ** p<0.05, * p<0.1 (two-tailed tests).

REFERENCES

Abadie, Alberto, David Drukker, Jane Leber Herr, Guido W. Imbens. 2001. "Implementing Matching Estimators for Average Treatment Effects in Stata." *The Stata Journal* 1:1, 1-18.

Alexander, Alan. 1982. *Local Government in Britain Since Reorganisation*. London: Allen & Unwin.

Altman, David, Anibal Perez-Linan. 2002. "Assessing the Quality of Democracy: Freedom, Competitiveness and Participation in Eighteen Latin American Countries." *Democratization* 9:2 (Summer) 85-100.

Ansolabehere, Stephen, James M. Snyder, Jr. 2012. "The Effects of Redistricting on Incumbents." *Election Law Journal* 11:4, 490-502.

Ansolabehere, Stephen, Maxwell Palmer, Amanda Lee. 2014. "Precinct-Level Election Data." http://hdl.handle.net/1902.1/21919. Harvard Election Data Archive [Distributor] V1 [Version]

Bartolini, Stefano. 1999. "Collusion, Competition and Democracy, Part I." *Journal of Theoretical Politics* 11:4, 435-470.

Bartolini, Stefano. 2000. "Collusion, Competition and Democracy: Part II." *Journal of Theoretical Politics*, 12:1, 33-65.

Blackwell, Matthew, Stefano Iacus, Gary King, Giuseppe Porro. 2010. "cem: Coarsened Exact Matching in Stata." *Stata Journal* 9:4, 524-46.

Blais, Andre, Ignacio Lago. 2009. "A General Measure of District Competitiveness." *Electoral Studies* 28:1 (March) 94-100.

Brambor, Thomas and Ricardo Ceneviva. 2012. "Reeleição e Continuísmo nos Municípios Brasileiros." *Novos Estudos - CEBRAP*, no. 93: 09–21.

Chen, Jowei, Jonathan Rodden. 2013. "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures." *Quarterly Journal of Political Science* 8, 239-69.

Colomer, Josep M. (ed). 2004. *Handbook of Electoral System Choice*. Palgrave Macmillan.

Elkins, David J. 1974. "The Measurement of Party Competition." *American Political Science Review* 68:2 (June) 682-700.

Ferreira, Fernando, Joseph Gyourko. 2009. "Do Political Parties Matter? Evidence from U.S. Cities." *The Quarterly Journal of Economics* 124(1): 399-422.

Gelman, Andrew, Gary King. 1994. "Enhancing Democracy Through Legislative Redistricting." *American Political Science Review* 88:3 (September) 541-59.

Gerber, E. R., Daniel J. Hopkins. 2011. "When Mayors Matter: Estimating the Impact of Mayoral Partisanship on City Policy." *American Journal of Political Science* 55: 326–339.

Grofman, Bernard, Peter Selb. 2009. "A Fully General Index of Political Competition." *Electoral Studies* 28:2, 291-96.

Gyford, John, Steve Leach, Chris Game. 1989. *The Changing Politics of Local Government*. London: Unwin Hyman.

Inter-university Consortium for Political and Social Research. "Candidate and Constituency Statistics of Elections in the United States, 1788-1990." ICPSR07757-v5. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1994. http://doi.org/10.3886/ICPSR07757.v5

King, Gary, Bradley Palmquist, Greg Adams, Micah Altman, Kenneth Benoit, Claudine Gay, Jeffrey B. Lewis, Russ Mayer, Eric Reinhardt. 1997. "The Record of American Democracy, 1984-1990." Harvard University, Cambridge, MA [producer], Ann Arbor, MI: ICPSR [distributor].

Klarner, Carl, William Berry, Thomas Carsey, Malcolm Jewell, Richard Niemi, Lynda Powell, and James Snyder. State Legislative Election Returns (1967-2010). ICPSR34297-v1. Ann Arbor, MI:

Inter-university Consortium for Political and Social Research [distributor], 2013-01-11. doi:10.3886/ICPSR34297.v1.
Nice, David. 1984. "Competitiveness in House and Senate Elections with Identical Constituencies." *Political Behavior* 6:1, 95-102.
Parker, Randy. 2010. *Our Campaigns* [web portal] www.ourcampaigns.com/StateOfficeList.html
Pempel, T.J. (ed). 1990. *Uncommon Democracies: The One-Party Dominant Regimes*. Ithaca, NY: Cornell University Press.
Przeworski, Adam et al. 2013. Political Institutions and Political Events (PIPE) Data Set. Department of Politics, New York University.
Rallings, Colin S., Michael Thrasher, L. Ware. 2006. *British Local Election Database, 1889-2003*. University of Plymouth.
Strøm, Kaare. 1989. "Inter-party Competition in Advanced Democracies." *Journal of Theoretical Politics* 1:277.
Taagepera, Rein, Matthew S. Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven: Yale University Press.
Trounstine, Jessica. 2008. *Political Monopolies in American Cities: The Rise and Fall of Bosses and Reformers*. Chicago: University of Chicago Press.
Wallin, Gunnar. 1973. "Småkommuner på avskrivning." In Wallin, G. H. Andersson, & N. Andrén, eds. *Kommunerna i förvandling*. Stockholm: Almqvist & Wiksell.
Wångmar, Erik. 2006. *Samlingsstyre – blockstyre – mångstyre: Kommmunalpolitiska styrelseformer 1952-2002*. Stockholm. Stads- och kommunhistoriska institutet.
Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.