Online Appendix

When Do Politicians Appeal Broadly? The Economic Consequences of Electoral Rules in Brazil

Moya Chin

Contents

A. Data Appendix

Table 1 Variables used to construct the equipment index, from the *Censo Escolar*

Variable	2000	2001-2003	2004	2005-2006	2007	2008	2009-2011	2012-2016
VCR/DVD	x	x	x	x	x	x	x	x
TV	x	x	x	x	x	x	x	x
satellite dish	x	x	x	x	x	x	x	x
satellite dish with internet			x	x				
overhead projector	x	x	x	x	x	x	x	x
projector			x	x				x
fax		x	x	x				x
copier		x	x	x	x	x	x	x
stereo/sound system	x	x	x	x				x
camera/camcorder			x	x				x
drinking fountain			x	x				
special needs accom.	x	x	x	x	x	x		
fan		x	x	x				
air conditioning		x	x	x				
computers	x	x	x	x	x	x	x	x
printer	x	x	x	x	x	x	x	x
local network	x	x	x					
internet	x	x	x		x	x	x	x
broadband						x	x	x

Table 2 Variables used to construct the infrastructure index, from the *Censo Escolar*

Variable	1997	1998	1999-2000	2001-2003	2004-2006	2007-2008	2009-2011	2012-2016
principal office	x	x	x	x	x	x	x	x
secretary office	x	x	x	x	x			x
teacher lounge		x	x	x	x	x	x	x
teacher housing								x
library	x	x	x	x	x	x	x	x
reading room				x	x		x	x
video library / room			x	x	x			
toy library					x			
auditorium					x			x
solarium					x			
science lab	x	x	x	x	x	x	x	x
computer lab	x	x	x	x	x	x	x	x
other lab		x	x	x	x			
kitchen	x	x	x	x	x	x	x	x
food pantry	x	x	x	x	x			x
cafeteria	x	x	x	x	x			x
warehouse				x	x			x
schoolyard	x	x	x	x	x			x
green area								x
sports field	x	x	x	x	x	x	x	x
pool				x	x			
gymnasium					x			
playground	x	x	x	x	x	x	x	x
laundry					x			x
sanitation	x	x	x	x	x	x	x	x
special needs accomm.	x	x	x	x	x	x	x	x
classrooms	x	x	x	x	x	x	x	x

B. Additional Figures and Tables

B.1. Identification

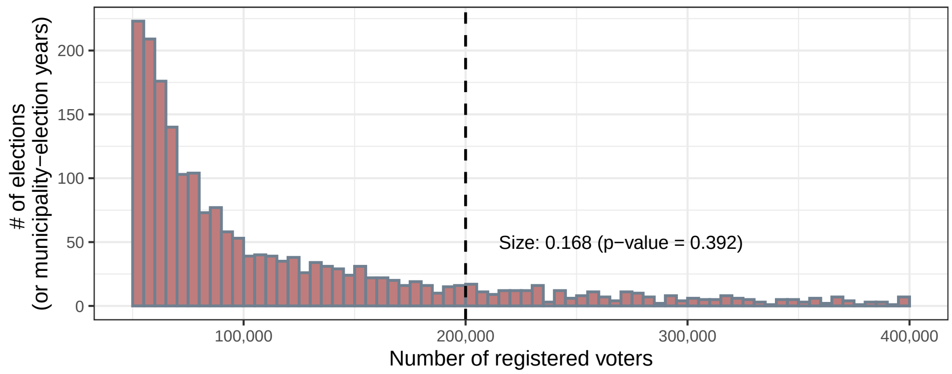


Figure 1 Density of elections around the 200,000 registered voter threshold

Note: Figure plots elections with 50,000−400,000 registered voters (6.0% of the universe of elections) in 5,000 voter bins. Size of the discontinuity is estimated based on McCrary (2008). Due to the skewed right tail of municipality sizes, the size of the discontinuity was estimated using a sample excluding those above the 99.9 percentile of registered voters.

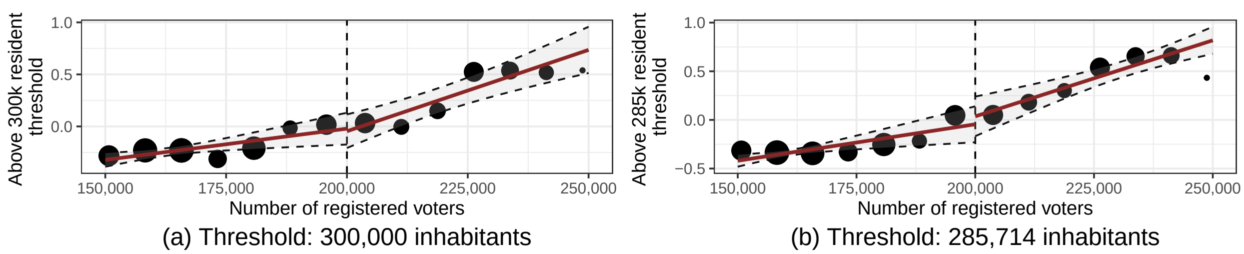


Figure 2 Regression discontinuity plots of the probability of falling above/below other policy thresholds

Note: The figures plot the fraction of elections above the 300,000 resident threshold (Panel a) and the 285,714 resident threshold (Panel b). At 300,000 residents, a salary cap for municipal legislators comes into effect. At 285,714 residents, the size of the legislature changes. In each panel, each point plots an average value within a 7,500 voter bin. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

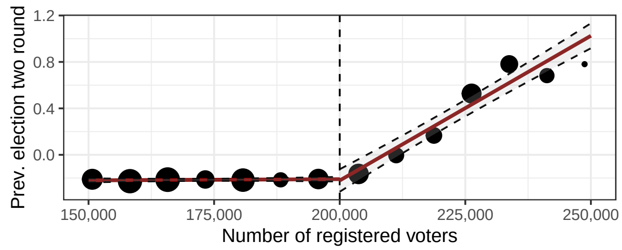
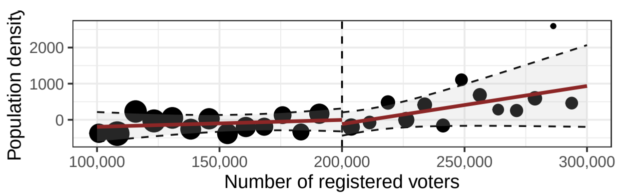
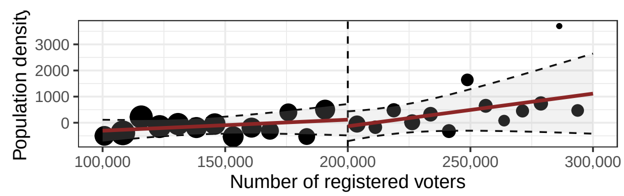


Figure 3 Regression discontinuity plot of the probability of treatment in previous election

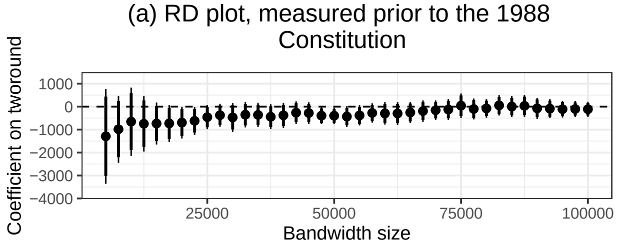
Note: The figure plots the probability that the *previous* election was a two-round election. Each point plots an average value within a 7,500 voter bin. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level. Note that because all observations to the left of the threshold are 0, there are no standard errors. The size of the discontinuity is -0.014 ($p = 0.77$).



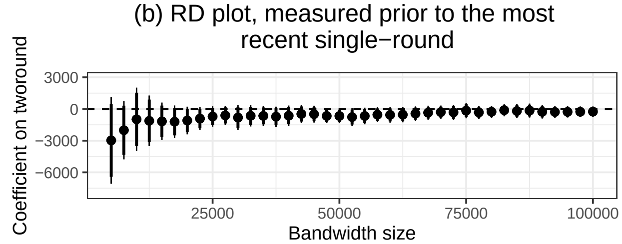
(a) RD plot, measured prior to the 1988 Constitution



(b) RD plot, measured prior to the most recent single−round



(c) RD estimates at different bandwidths, measured prior to the 1988 Constitution



(d) RD estimates at different bandwidths, measured prior to the most recent single−round

Figure 4 Pre-treatment population density

Note: The figures plot population density measured in the 1980 census (Panels a and c) or in the census prior to the most recent year in a single-round system or in the 1991 census (Panels b and d). In panels a and b, each point plots an average value within a 7,500 voter bin. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level. Panels c and d plot the RD coefficients at different bandwidths. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. *Estimation method:* Local linear regression with the specified voter bandwidth and election-year fixed effects. Standard errors clustered at the municipality level.

B.2. Other results on electoral outcomes

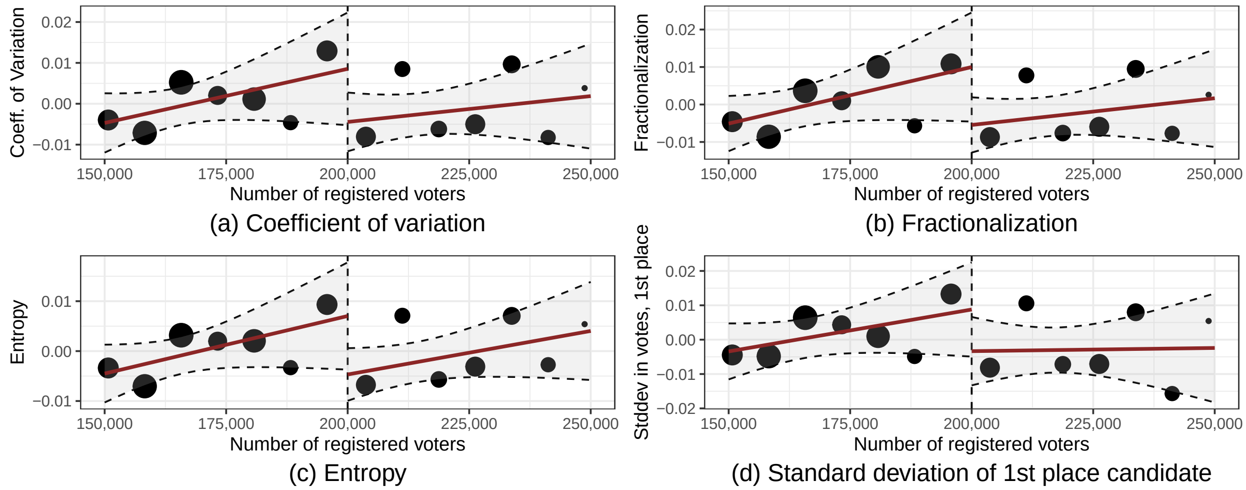


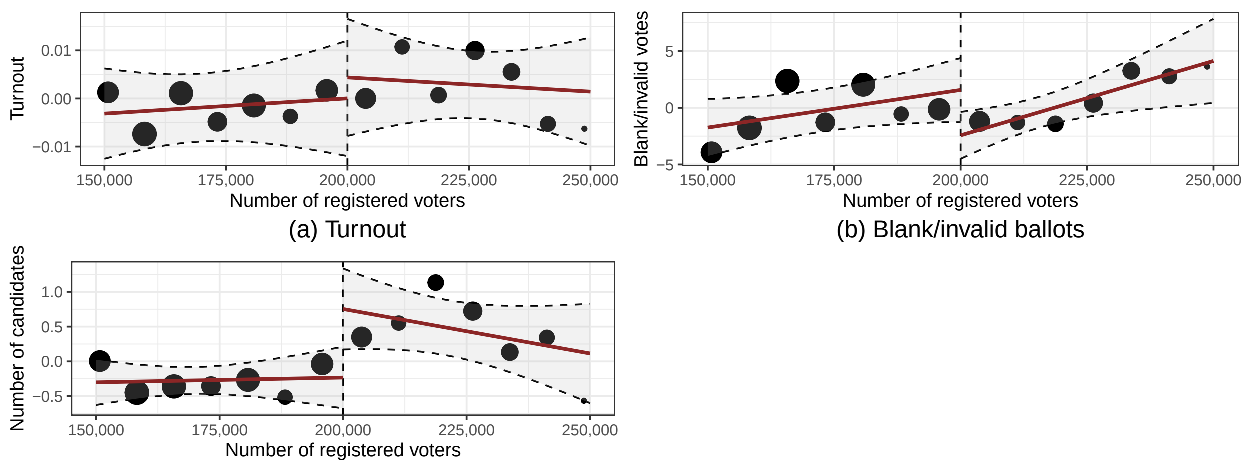
Figure 5 Regression discontinuity plots of geographic concentration of voters, using vote shares from top two candidates only

Note: The figures plot the overall concentration of voters for specific candidates (Panels a-c) and the standard deviation in the 1st place candidate's vote shares across electoral sections (Panel d). All outcomes use vote shares from the top two candidates only. Vote shares are from the first round. In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

Table 1 Regression discontinuity estimates on the geographic concentration of voters, using vote shares from the final round

	Coefficient of variation	Fractionalization	Entropy	Std Dev of 1st place candidate
TwoRound	−0.019	−0.022	−0.016	−0.021
	(0.008)	(0.008)	(0.006)	(0.009)
Single-round mean	0.036	0.038	0.029	0.088
Observations	263	263	263	263
Municipalities	88	88	88	88

Note: The table presents RD estimates on the overall concentration of voters for specific candidates and the standard deviation in the 1st place candidate's vote shares across electoral sections, using vote shares from the final round (1st round results in single-round elections and 2nd round results in two-round elections). Outcomes are calculated using vote shares from the top two candidates only. Observations are at the election level. *Single-round mean* refers to the dependent variable mean for single-round municipalities within the bandwidth. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.



(a) Turnout

(b) Blank/invalid ballots

(c) Number of candidates

Figure 6 Regression discontinuity plots of other electoral outcomes

Note: The figures plot other electoral outcomes. *Turnout* is the fraction of eligible voters who cast a ballot in the election. *Blank/invalid ballots* is the number of ballots (in thousands) that were either blank or voided. In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

Table 2 Regression discontinuity estimates on the geographic concentration of voters, with number of candidates as a control

Panel A: Concentration indices of voters for specific candidates				
	Coefficient of variation	Fractionalization	Entropy	
TwoRound	−0.005	−0.010	−0.009	
	(0.003)	(0.005)	(0.005)	
Single-round mean	0.019	0.027	0.030	
Observations	264	264	264	
Municipalities	89	89	89	
Panel B: Standard deviation in vote shares for each candidate				
	1st place candidate	2nd place candidate	3rd place candidate	4th place candidate
TwoRound	−0.016	−0.012	−0.010	−0.002
	(0.007)	(0.008)	(0.007)	(0.004)
Single-round mean	0.080	0.075	0.042	0.023
Observations	264	264	251	216
Municipalities	89	89	89	84

Note: The table presents RD estimates on the overall concentration of voters for specific candidates (Panel A) and the standard deviation in a candidate's vote shares for the 1st-4th place candidate across electoral sections (Panel B). Vote shares are from the first round. Observations are at the election level. *Single-round mean* refers to the dependent variable mean for single-round municipalities within the bandwidth. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Number of candidates included as a control. Standard errors clustered at the municipality level.

B.3. Other results on education resources

Table 3 Regression discontinuity estimates on resources in municipal schools, using z-scores

	Mean level of resources		*Standard deviation in resources*	
	Equipment	Infrastructure	Equipment	Infrastructure
TwoRound	0.079	0.069	−0.014	−0.007
	(0.033)	(0.037)	(0.009)	(0.017)
Single-round mean	0.724	0.739	0.120	0.146
Observations	820	912	820	912
Municipalities	79	79	79	79

Note: The table presents RD estimates on the mean level (first two columns) and standard deviation (last two columns) in resources in municipal schools. *Equipment* and *Infrastructure* are indices constructed by taking the z-score of a school's equipment and infrastructure elements, calculating the school's percentile in the national distribution, then averaging across schools in the municipality. Observations are at the year level. *Single-round mean* refers to the dependent variable mean for single-round municipalities within the bandwidth. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

B.4. Other results on downstream outcomes

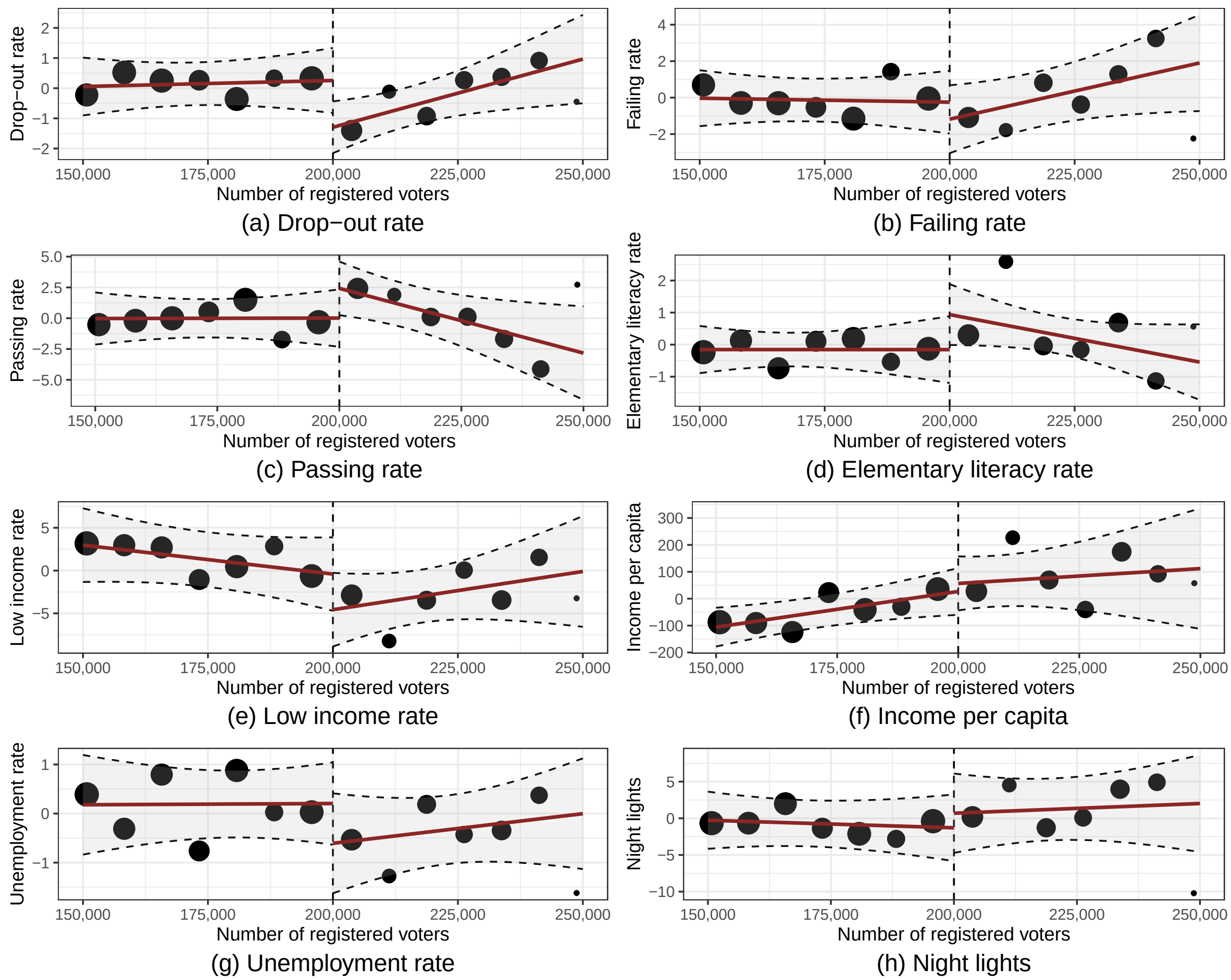


Figure 7 Regression discontinuity plots of downstream municipal outcomes

Note: The figures plot education outcomes (Panels a-d) and economic outcomes (Panels e-h). *Drop-out rate*, *Failing rate*, and *Passing rate* are the mean rates across schools in the municipality, from the School Census. *Elementary literacy rate* is the literacy rate of cohorts who are of elementary school age during the mayoral term, from the 2000 and 2010 Demographic Census. *Low income rate*, *Income per capita*, and *Unemployment rate* are from the 2000 and 2010 Demographic Census. *Low income rate* is the fraction of households earning between 0 and 50% of the minimum wage. *Night lights* is the mean night lights level in the municipality, from the 1997-2013 NOAA night lights series. In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

B.5. Robustness tests on RDD design

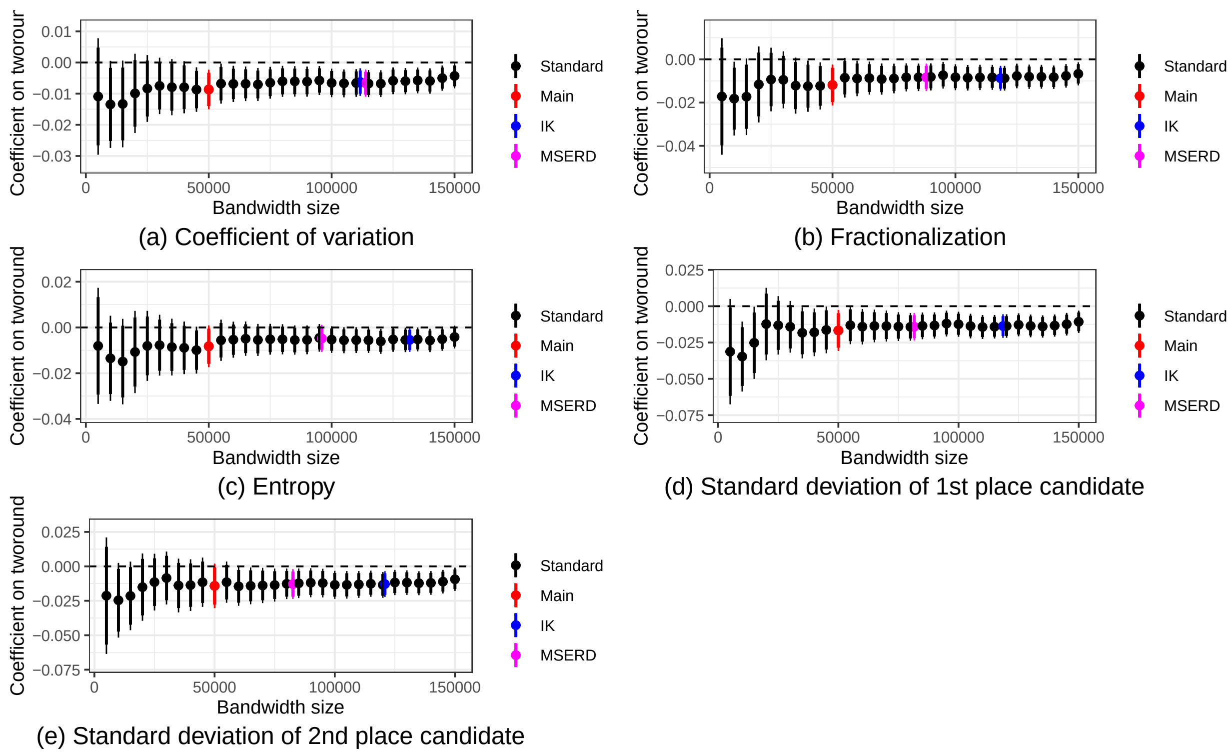


Figure 8 Regression discontinuity coefficients on geographic concentration of voters at different bandwidths

Note: The figures plot RD coefficients at different bandwidths for the overall concentration of voters for specific candidates (Panels a-c) and the standard deviation in a candidate's vote shares across electoral sections (Panels d-e). Vote shares are from the first round. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and with the specified voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

Table 4 Regression discontinuity estimates on concentration indices of voters for specific candidates, with different specifications

	Coefficient of variation	Fractionalization	Entropy
Panel A: No controls			
TwoRound	−0.008	−0.010	−0.007
	(0.003)	(0.004)	(0.005)
Observations	264	264	264
Municipalities	89	89	89
Panel B: With election-year fixed effects only			
TwoRound	−0.008	−0.011	−0.009
	(0.003)	(0.005)	(0.005)
Observations	264	264	264
Municipalities	89	89	89
Panel C: Baseline specification with controls			
TwoRound	−0.006	−0.006	−0.005
	(0.003)	(0.004)	(0.004)
Observations	230	230	230
Municipalities	74	74	74
Panel D: Local quadratic specification			
TwoRound	−0.009	−0.012	−0.008
	(0.003)	(0.005)	(0.005)
Single-round mean	0.019	0.027	0.030
Observations	264	264	264
Municipalities	89	89	89

Note: The table presents RD estimates on the overall concentration of voters for specific candidates. Vote shares are from the first round. Observations are at the election level. *Panel A:* Local linear regression with a 50,000 voter bandwidth. *Panel B:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. *Panel C:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Pre-treatment characteristics (municipal area change, population growth, population 0-15 years, income segregation, demographic segregation, literacy rate, income per capita, low income rate, unemployment rate, and Gini coefficient) measured prior to the most recent single-round election included as controls. Population density included as a control separately across the cutoff. *Panel D:* Local quadratic regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. All standard errors clustered at the municipality level.

Table 5 Regression discontinuity estimates on standard deviation in vote shares for each candidate, with different specifications

	1st place candidate	2nd place candidate	3rd place candidate	4th place candidate
Panel A: No controls				
TwoRound	−0.012	−0.011	−0.003	0.005
	(0.006)	(0.008)	(0.006)	(0.005)
Observations	264	264	251	216
Municipalities	89	89	89	84
Panel B: With election-year fixed effects only				
TwoRound	−0.014	−0.011	−0.004	0.003
	(0.007)	(0.008)	(0.006)	(0.004)
Observations	264	264	251	216
Municipalities	89	89	89	84
Panel C: Baseline specification with controls				
TwoRound	−0.008	−0.003	−0.003	0.004
	(0.007)	(0.008)	(0.006)	(0.006)
Observations	230	230	217	185
Municipalities	74	74	74	71
Panel D: Local quadratic specification				
TwoRound	−0.017	−0.014	−0.005	0.004
	(0.007)	(0.008)	(0.007)	(0.004)
Single-round mean	0.080	0.075	0.042	0.023
Observations	264	264	251	216
Municipalities	89	89	89	84

Note: The table presents RD estimates on the standard deviation in a candidate's vote shares across electoral sections. Vote shares are from the first round. Observations are at the election level. *Panel A:* Local linear regression with a 50,000 voter bandwidth. *Panel B:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. *Panel C:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Pre-treatment characteristics (municipal area change, population growth, population 0-15 years, income segregation, demographic segregation, literacy rate, income per capita, low income rate, unemployment rate, and Gini coefficient) measured prior to the most recent single-round election included as controls. Population density included as a control separately across the cutoff. *Panel D:* Local quadratic regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. All standard errors clustered at the municipality level.

Table 6 Regression discontinuity estimates on geographic concentration of voters using vote shares from top two candidates only, with different specifications

	Coefficient of variation	Fractionaliza-tion	Entropy	Std Dev of 1st place candidate
Panel A: No controls				
TwoRound	−0.011	−0.015	−0.011	−0.009
	(0.008)	(0.008)	(0.006)	(0.008)
Observations	264	264	264	264
Municipalities	89	89	89	89
Panel B: With election-year fixed effects only				
TwoRound	−0.014	−0.017	−0.013	−0.012
	(0.009)	(0.009)	(0.007)	(0.009)
Observations	264	264	264	264
Municipalities	89	89	89	89
Panel C: Baseline specification with controls				
TwoRound	−0.007	−0.010	−0.008	−0.004
	(0.008)	(0.008)	(0.006)	(0.009)
Observations	230	230	230	230
Municipalities	74	74	74	74
Panel D: Local quadratic specification				
TwoRound	−0.015	−0.018	−0.013	−0.015
	(0.008)	(0.009)	(0.007)	(0.009)
Single-round mean	0.036	0.038	0.029	0.088
Observations	264	264	264	264
Municipalities	89	89	89	89

Note: The table presents RD estimates on the overall concentration of voters for specific candidates and the standard deviation in the 1st place candidate's vote shares across electoral sections, calculated using vote shares from the top two candidates only. Vote shares are from the first round. Observations are at the election level. *Panel A:* Local linear regression with a 50,000 voter bandwidth. *Panel B:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. *Panel C:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Pre-treatment characteristics (municipal area change, population growth, population 0-15 years, income segregation, demographic segregation, literacy rate, income per capita, low income rate, unemployment rate, and Gini coefficient) measured prior to the most recent single-round election included as controls. Population density included as a control separately across the cutoff. *Panel D:* Local quadratic regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. All standard errors clustered at the municipality level.

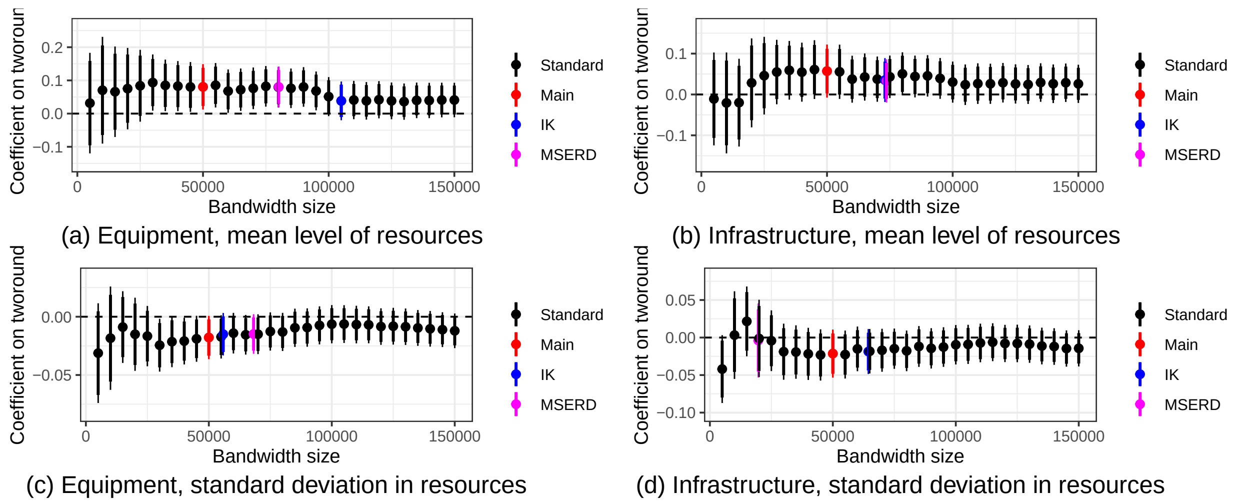


Figure 9 Regression discontinuity coefficients on resources in municipal schools at different bandwidths

Note: The figures plot RD coefficients at different bandwidths for the mean level of resources in schools (Panels a-b) and the standard deviation in resources across schools (Panels c-d). *Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school's equipment and infrastructure elements, calculating the school's percentile in the national distribution, then averaging across schools in the municipality. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and with the specified voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

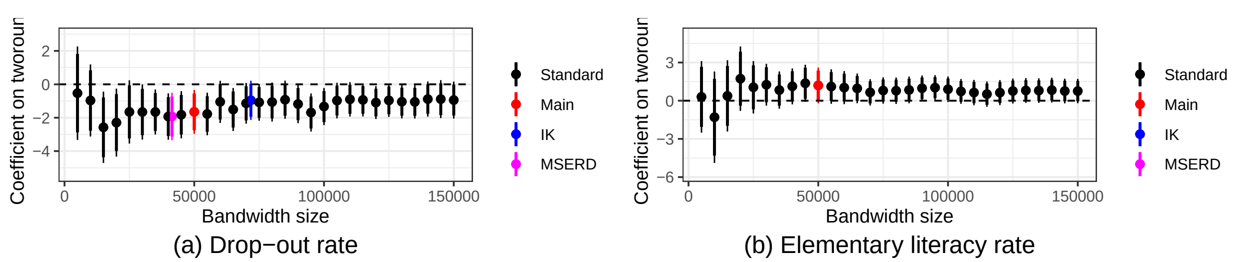


Figure 10 Regression discontinuity coefficients on municipal education outcomes at different bandwidths

Note: The figures plot RD coefficients at different bandwidths for municipal education outcomes. *Drop-out rate* is the mean rate across schools in the municipality, from the School Census. *Elementary literacy rate* is the literacy rate of cohorts who are of elementary school age during the mayoral term, from the 2000 and 2010 Demographic Census. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. IK and MSERD bandwidths not shown for *Elementary literacy rate*, as the bandwidth chosen was larger than the support. *Estimation method:* Local linear regression with election-year fixed effects and with the specified voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

Table 7 Regression discontinuity estimates on resources in municipal schools, with different specifications

	Mean level of resources		*Standard deviation in resources*	
	Equipment	Infrastructure	Equipment	Infrastructure
Panel A: No controls				
TwoRound	0.068	0.036	−0.019	−0.014
	(0.035)	(0.029)	(0.011)	(0.015)
Observations	821	916	821	916
Municipalities	79	79	79	79
Panel B: With election-year fixed effects only				
TwoRound	0.062	0.036	−0.015	−0.013
	(0.033)	(0.029)	(0.009)	(0.015)
Observations	821	916	821	916
Municipalities	79	79	79	79
Panel C: Baseline specification with controls				
TwoRound	0.070	0.047	−0.018	−0.032
	(0.027)	(0.020)	(0.011)	(0.015)
Observations	677	677	677	677
Municipalities	62	62	62	62
Panel D: Local quadratic specification				
TwoRound	0.081	0.057	−0.018	−0.021
	(0.035)	(0.033)	(0.009)	(0.016)
Single-round mean	0.738	0.731	0.121	0.157
Observations	820	912	820	912
Municipalities	79	79	79	79

Note: The table presents RD estimates on the mean level (first two columns) and standard deviation (last two columns) in resources in municipal schools. *Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school's equipment and infrastructure elements, calculating the school's percentile in the national distribution, then averaging across schools in the municipality. Observations are at the year level. *Panel A:* Local linear regression with a 50,000 voter bandwidth. *Panel B:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. *Panel C:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Pre-treatment characteristics (municipal area change, population growth, population 0-15 years, income segregation, demographic segregation, literacy rate, income per capita, low income rate, unemployment rate, and Gini coefficient) measured prior to the most recent single-round election included as controls. Population density included as a control separately across the cutoff. *Panel D:* Local quadratic regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. All standard errors clustered at the municipality level.

Table 8 Regression discontinuity estimates on municipal education outcomes, without controls

	Drop-out rate	Failing rate	Passing rate	Elem. literacy rate
Panel A: No controls				
TwoRound	−1.340	−0.088	1.291	2.918
	(0.755)	(1.163)	(1.686)	(2.030)
Observations	913	913	913	178
Municipalities	79	79	79	71
Panel B: With election-year fixed effects only				
TwoRound	−1.331	−0.138	1.395	1.173
	(0.647)	(1.046)	(1.412)	(0.654)
Observations	913	913	913	178
Municipalities	79	79	79	71
Panel C: Baseline specification with controls				
TwoRound	−0.398	−1.206	1.604	1.238
	(0.492)	(1.284)	(1.429)	(0.662)
Observations	677	677	677	116
Municipalities	62	62	62	53
Panel D: Local quadratic specification				
TwoRound	−1.649	−0.758	2.330	1.199
	(0.667)	(1.114)	(1.459)	(0.710)
Single-round mean	3.211	8.645	88.283	91.445
Observations	909	909	909	177
Municipalities	79	79	79	71

Note: The table presents RD estimates on municipal education outcomes. *Drop-out rate*, *Failing rate*, and *Passing rate* are the mean rates across schools in the municipality, from the School Census. *Elementary literacy rate* is the literacy rate of cohorts who are of elementary school age during the mayoral term, from the 2000 and 2010 Demographic Census. Observations for drop-out rate, failing rate, and passing rate are at the year level. Observations for elementary literacy rate are at the election level. *Panel A:* Local linear regression with a 50,000 voter bandwidth. *Panel B:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. *Panel C:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Pre-treatment characteristics (municipal area change, population growth, population 0-15 years, income segregation, demographic segregation, literacy rate, income per capita, low income rate, unemployment rate, and Gini coefficient) measured prior to the most recent single-round election included as controls. Population density included as a control separately across the cutoff. *Panel D:* Local quadratic regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. All standard errors clustered at the municipality level.

B.6. Placebo tests

Table 9 Placebo regression discontinuity estimates on the geographic concentration of voters, at 285,714 inhabitant threshold

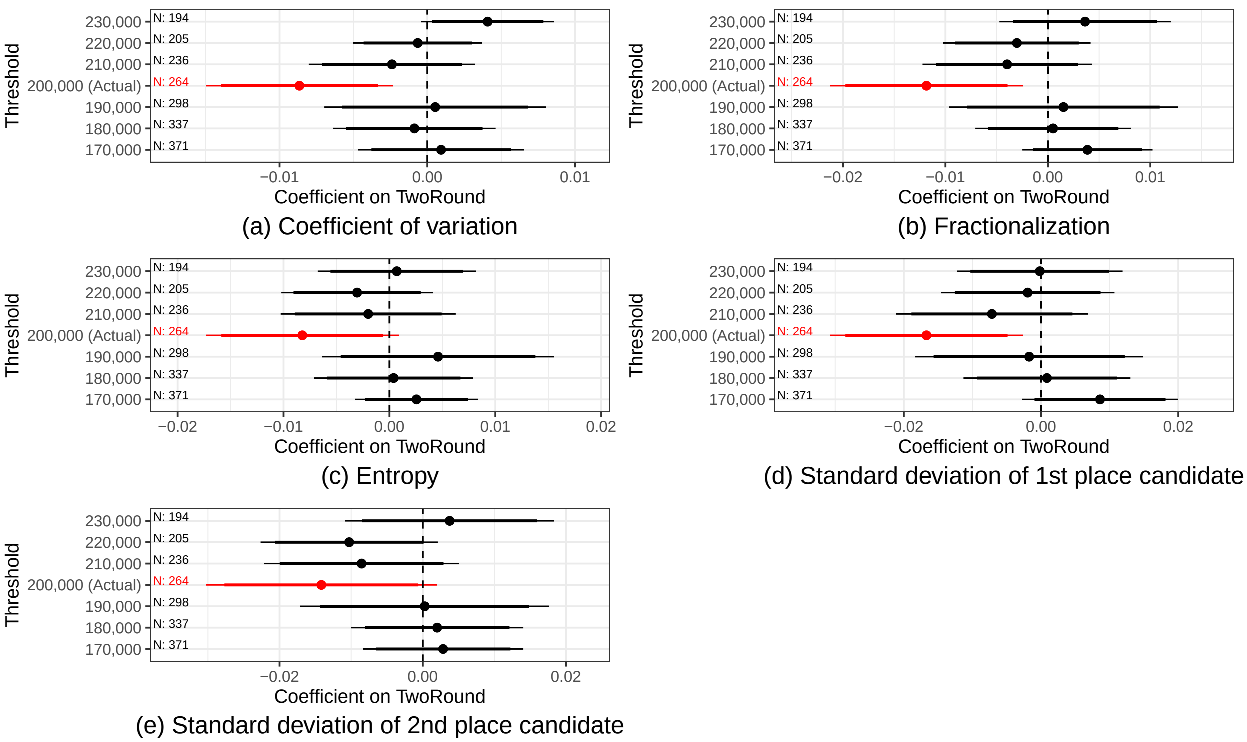
Panel A: Concentration indices of voters for specific candidates				
	Coefficient of variation	Fractionalization	Entropy	
TwoRound	0.004	0.005	0.007	
	(0.003)	(0.004)	(0.004)	
Single-round mean	0.019	0.024	0.027	
Observations	423	423	423	
Municipalities	122	122	122	
Panel B: Standard deviation in vote shares for each candidate				
	1st place candidate	2nd place candidate	3rd place candidate	4th place candidate
TwoRound	0.005	0.003	0.009	0.006
	(0.005)	(0.006)	(0.005)	(0.004)
Single-round mean	0.075	0.071	0.040	0.022
Observations	424	423	400	331
Municipalities	122	122	122	116

Note: The table presents RD estimates at the 285,714 inhabitant threshold on the overall concentration of voters for specific candidates (Panel A) and the standard deviation in a candidate's vote shares for the 1st-4th place candidate across electoral sections (Panel B). At 285,714 inhabitants, a 2004 constitutional amendment changing the size of the local legislature comes into effect. Vote shares are from the first round. Includes only elections after 2004. Observations are at the election level. *Single-round mean* refers to the dependent variable mean for single-round municipalities within the bandwidth. *Estimation method:* Local linear regression with election-year fixed effects and a 125,000 inhabitant bandwidth. To maintain comparability with the baseline estimates, the bandwidth was determined by taking half of the population range of municipalities in the 50,000 voter bandwidth (the smallest municipality is 182,082 inhabitants and the largest 434,474 inhabitants). Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

Table 10 Placebo regression discontinuity estimates on the geographic concentration of voters, at 300,000 inhabitant threshold

Panel A: Concentration indices of voters for specific candidates				
	Coefficient of variation	Fractionalization	Entropy	
TwoRound	0.005	0.005	0.005	
	(0.004)	(0.005)	(0.005)	
Single-round mean	0.019	0.024	0.027	
Observations	471	471	471	
Municipalities	113	113	113	
Panel B: Standard deviation in vote shares for each candidate				
	1st place candidate	2nd place candidate	3rd place candidate	4th place candidate
TwoRound	0.001	0.000	0.001	−0.003
	(0.008)	(0.007)	(0.006)	(0.005)
Single-round mean	0.075	0.072	0.040	0.021
Observations	471	471	444	373
Municipalities	113	113	113	110

Note: The table presents RD estimates at the 300,000 inhabitant threshold on the overall concentration of voters for specific candidates (Panel A) and the standard deviation in a candidate's vote shares for the 1st-4th place candidate across electoral sections (Panel B). At 300,000 inhabitants, a 2000 constitutional amendment placing a cap on local legislator salaries comes into effect. Vote shares are from the first round. Includes only elections after 2000. Observations are at the election level. *Single-round mean* refers to the dependent variable mean for single-round municipalities within the bandwidth. *Estimation method:* Local linear regression with election-year fixed effects and a 125,000 inhabitant bandwidth. To maintain comparability with the baseline estimates, the bandwidth was determined by taking half of the population range of municipalities in the 50,000 voter bandwidth (the smallest municipality is 182,082 inhabitants and the largest 434,474 inhabitants). Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.



(a) Coefficient of variation

(b) Fractionalization

(c) Entropy

(d) Standard deviation of 1st place candidate

(e) Standard deviation of 2nd place candidate

Figure 11 Regression discontinuity coefficients on geographic concentration of voters at different thresholds

Note: The figures plot RD coefficients at different thresholds for the overall concentration of voters for specific candidates (Panels a-c) and the standard deviation in a candidate's vote shares across electoral sections (Panels d-e). Vote shares are from the first round. The thicker horizontal lines represent the 90% confidence interval and the thinner horizontal lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

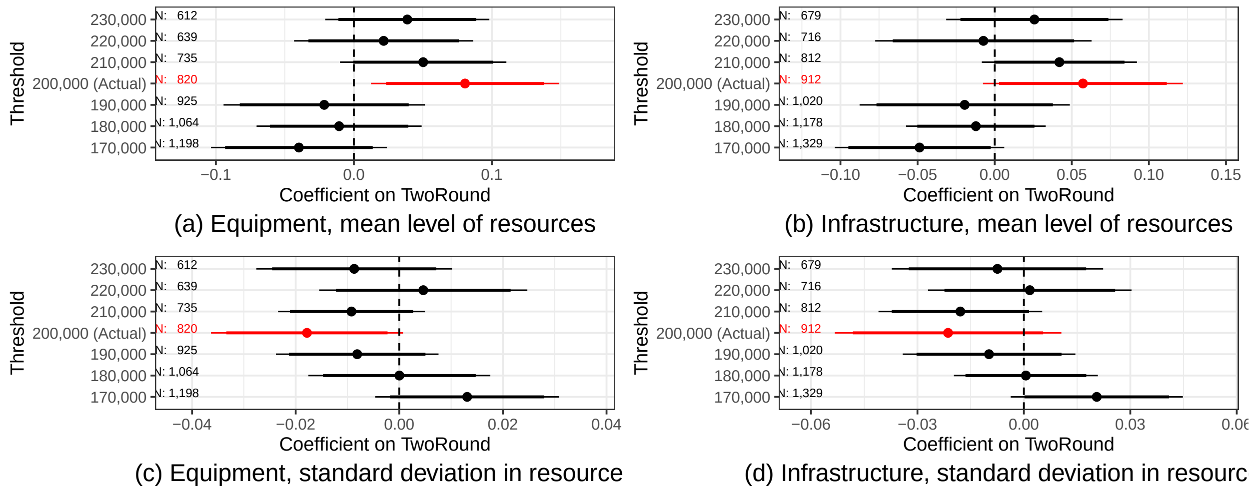


Figure 12 Regression discontinuity coefficients on resources in municipal schools at different thresholds

Note: The figures plot RD coefficients at different thresholds for the mean level of resources in schools (Panels a-b) and the standard deviation in resources across schools (Panels c-d). *Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school's equipment and infrastructure elements, calculating the school's percentile in the national distribution, then averaging across schools in the municipality. The thicker horizontal lines represent the 90% confidence interval and the thinner horizontal lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

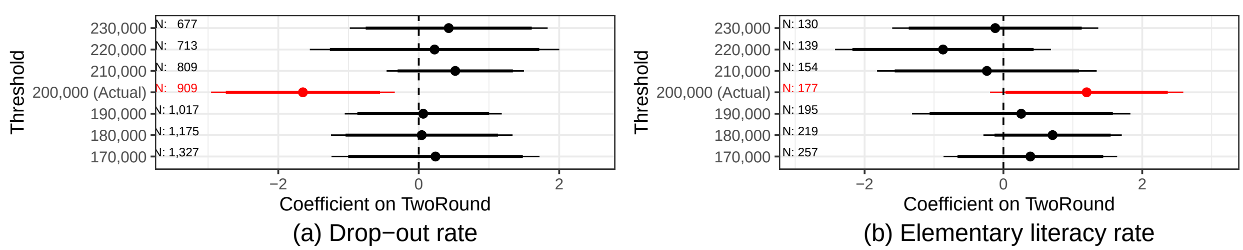
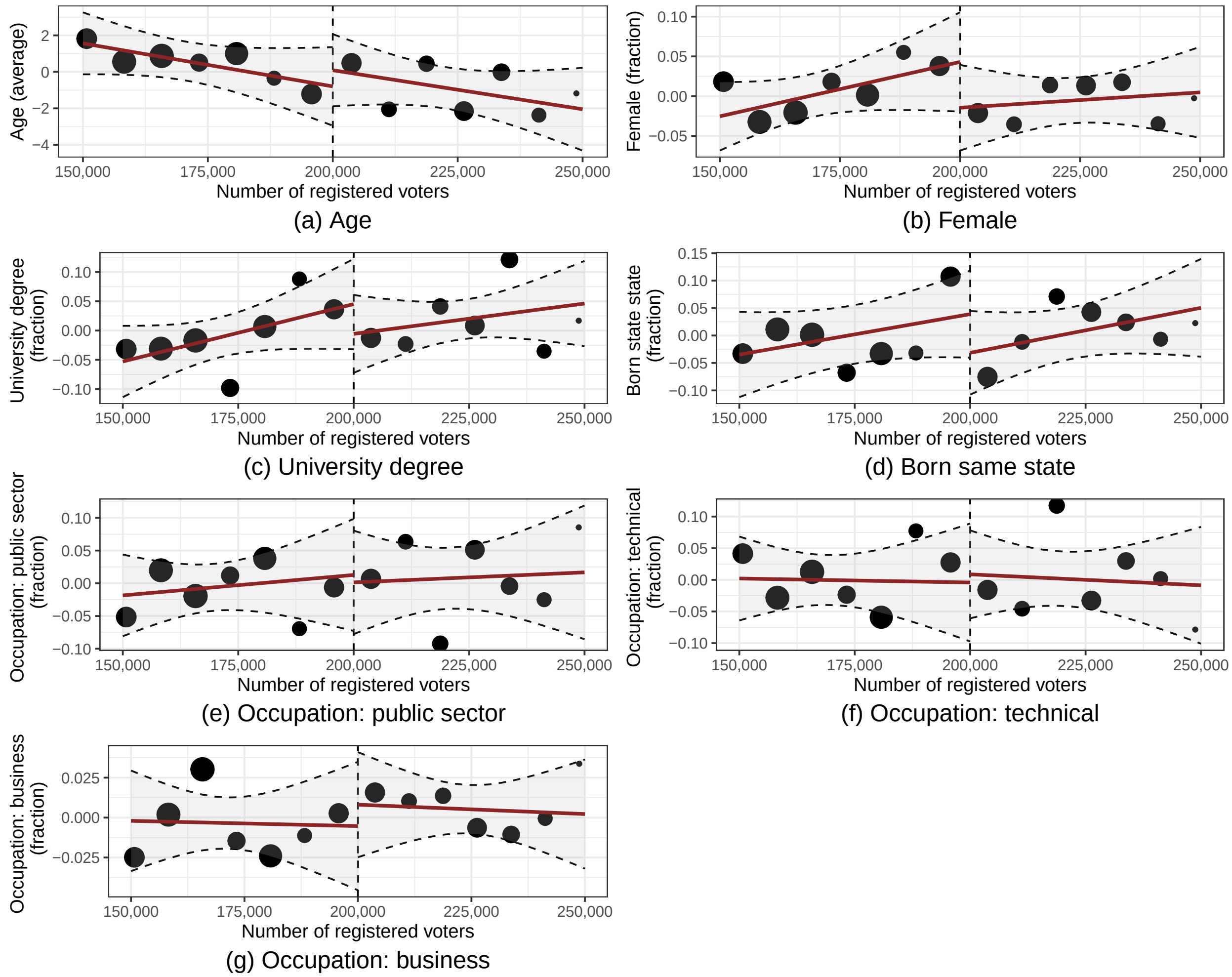


Figure 13 Regression discontinuity coefficients on municipal education outcomes at different thresholds

Note: The figures plot RD coefficients at different thresholds for municipal education outcomes. *Drop-out rate* is the mean rate across schools in the municipality, from the School Census. *Elementary literacy rate* is the literacy rate of cohorts who are of elementary school age during the mayoral term, from the 2000 and 2010 Demographic Census. The thicker horizontal lines represent the 90% confidence interval and the thinner horizontal lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

B.7. Other results on mechanisms



(a) Age

(b) Female

(c) University degree

(d) Born same state

(e) Occupation: public sector

(f) Occupation: technical

(g) Occupation: business

Figure 14 Regression discontinuity plots of characteristics of candidates

Note: The figures plot the average characteristics of candidates. *Born same state* is whether the candidate was born in the same state as the election. *Public sector* includes occupations such as elected positions, judiciary, and workers in public administration. *Technical* includes occupations such as scientists, technicians, and artists. *Business* includes occupations such as administrative positions, workers in commerce and services, and business owners. *Age* is the average age of candidates and *Female*, *University degree*, *Born same state* and previous occupation are fraction of candidates. In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

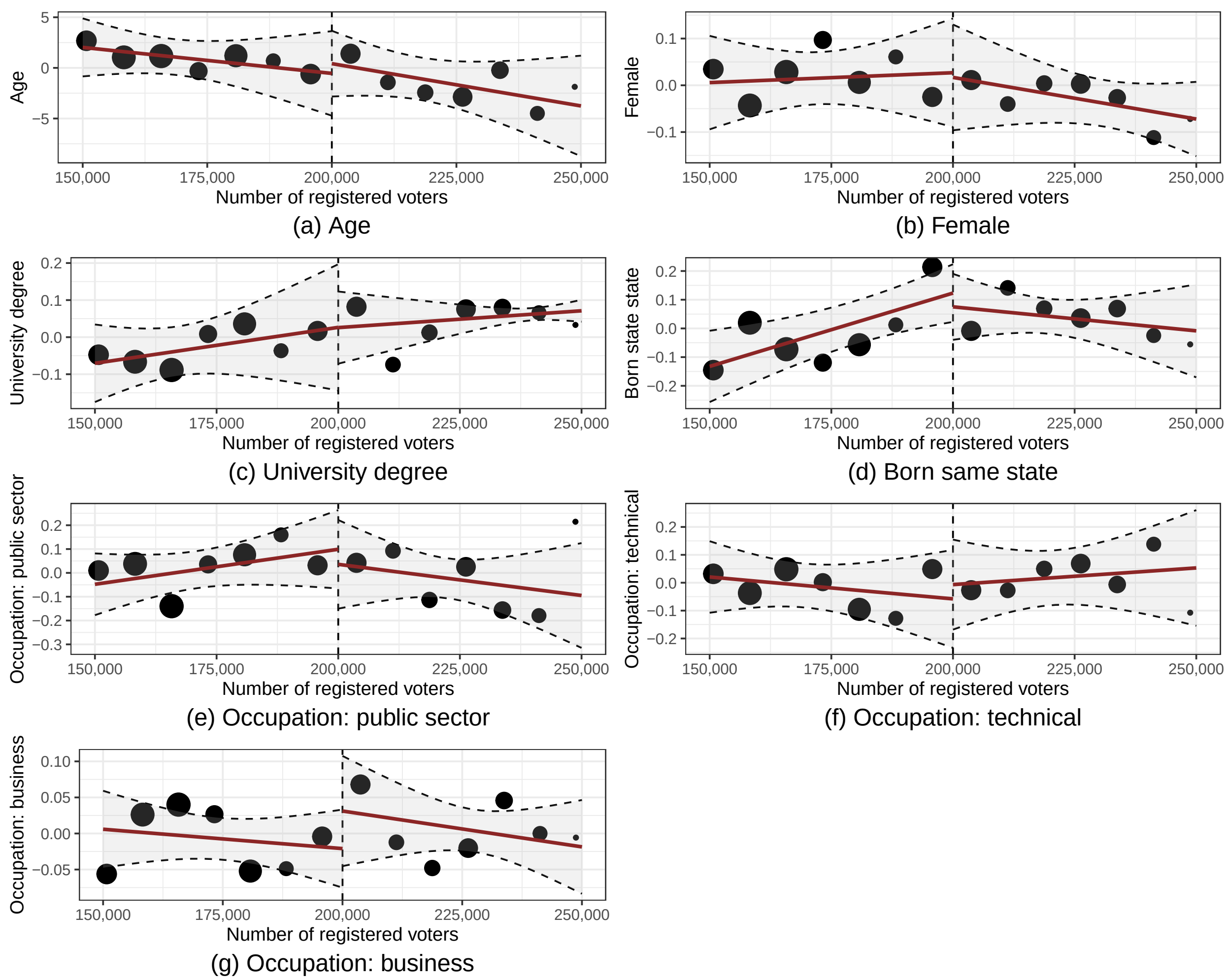


Figure 15 Regression discontinuity plots of characteristics of winners

Note: The figures plot the characteristics of winners. *Born same state* is whether the candidate was born in the same state as the election. *Public sector* includes occupations such as elected positions, judiciary, and workers in public administration. *Technical* includes occupations such as scientists, technicians, and artists. *Business* includes occupations such as administrative positions, workers in commerce and services, and business owners. *Female*, *University degree*, *Born same state* and previous occupation are indicator variables. In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

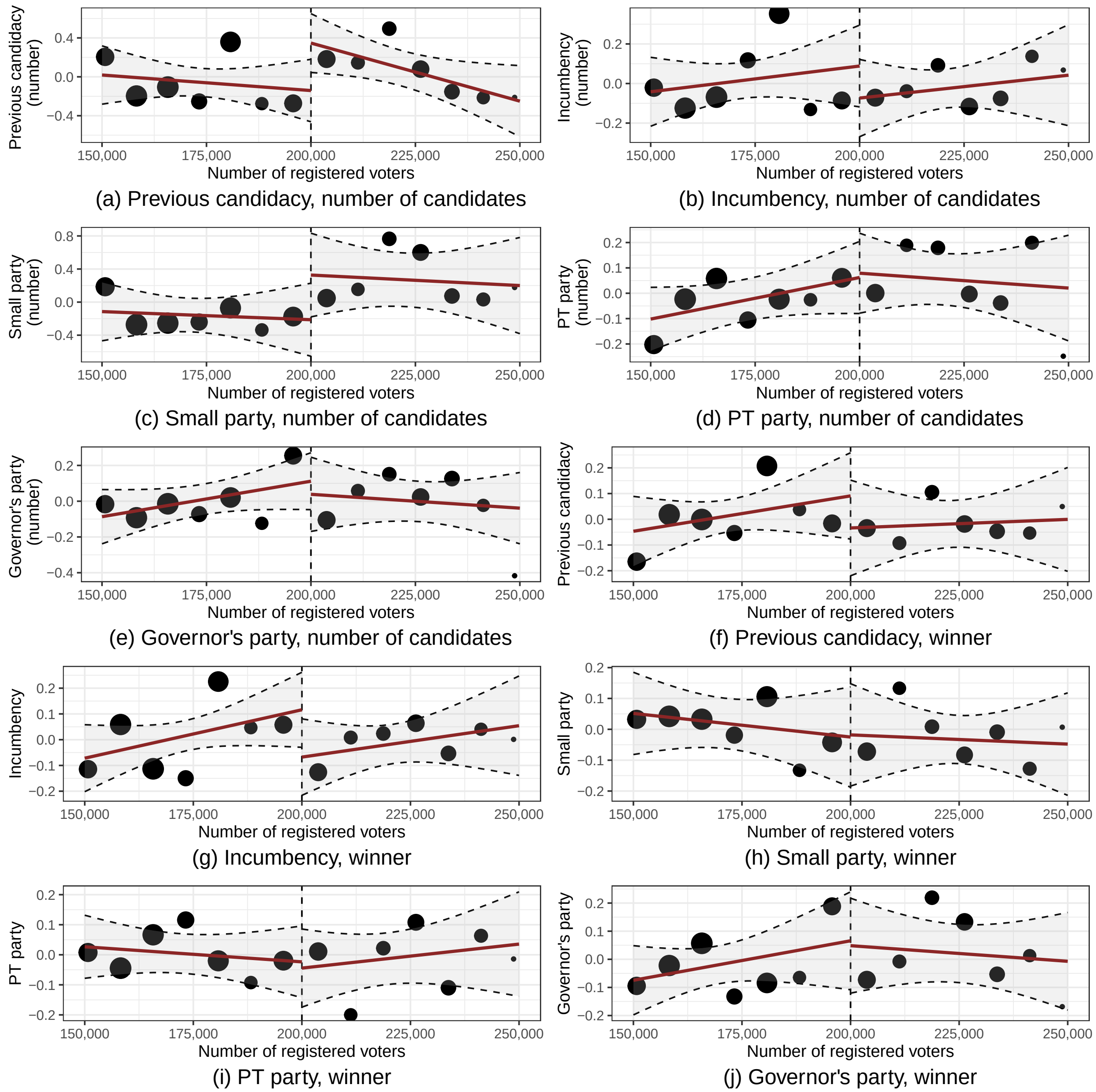


Figure 16 Regression discontinuity plots of political affiliation of candidates

Note: The figures plot the political affiliation of candidates (Panels a-e) and winners (Panels f-j). *Previous candidacy* is whether the candidate ran in a previous mayoral election. *Incumbency* is whether the candidate held the position of mayor in a previous term. *Small party* is any party that is not one of the top 5 parties, by national membership. *PT party* is whether the candidate is from the *Partido dos Trabalhadores*. *Governor's party* is whether the candidate is from the party of the incumbent state governor. *Previous candidacy*, *Incumbency*, and *Governor's party* are unavailable for the 1996 elections. Variables are either the number of candidates with that characteristic (Panels a-e) or an indicator for the winner having that characteristic (Panel f-j). In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

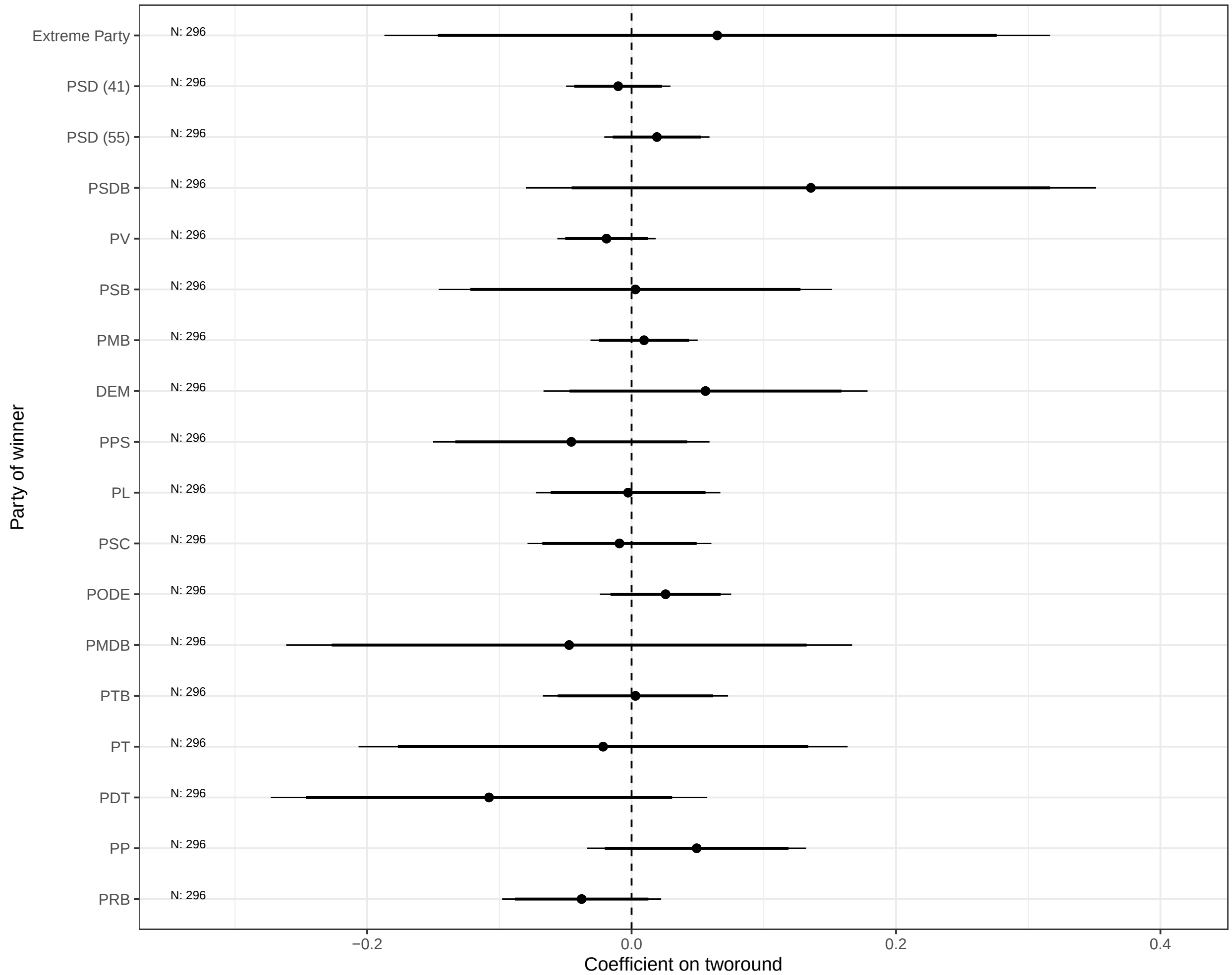


Figure 17 Regression discontinuity coefficients on party identity of winners

Note: The figure plots RD coefficients of the party affiliation of the election winner. *Extreme party* is an indicator for the winner coming from a religious right, secular right, or secular left party, as defined in Codato, Berlatto and Bolognesi (2018) (specifically: PEN, PHS, PL, PRB, PSC, PSDC, PPL, PSOL, DEM, PL, PP, PRTB, PSD, PSL or PTB).[1] The thicker horizontal lines represent the 90% confidence interval and the thinner horizontal lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

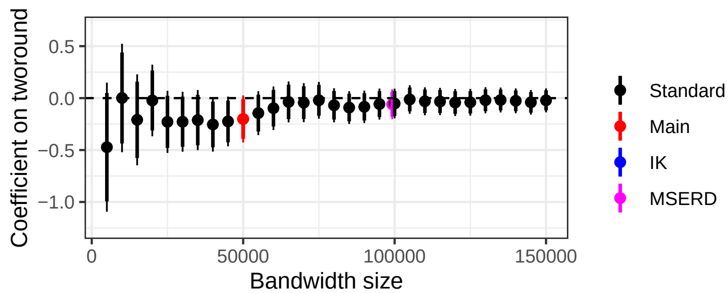


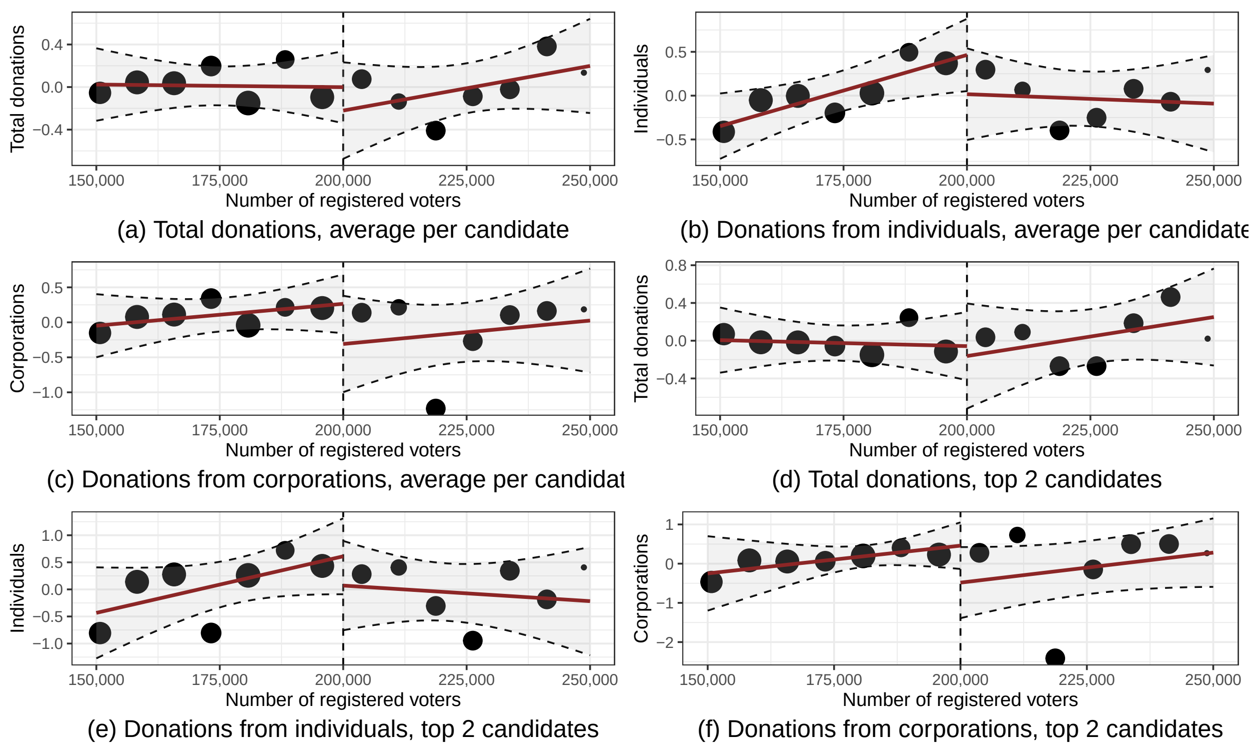
Figure 18 Regression discontinuity coefficients on incumbency status of winner at different bandwidths

Note: The figure plots RD coefficients at different bandwidths for the incumbency status of the winner. *Incumbency* is whether the winner held the position of mayor in a previous term. The thicker vertical lines represent the 90% confidence interval and the thinner vertical lines represent the 95% confidence interval. *Estimation method:* Local linear regression with election-year fixed effects and with the specified voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.

Table 11 Regression discontinuity estimates on resources in municipal schools, by incumbency status of mayor

	Mean level of resources		*Standard deviation in resources*	
	Equipment	Infrastructure	Equipment	Infrastructure
TwoRound	0.077	0.072	−0.019	−0.036
	(0.037)	(0.036)	(0.010)	(0.016)
TwoRound * Incumbent	0.025	0.007	−0.002	−0.008
	(0.024)	(0.021)	(0.008)	(0.012)
Single-round mean	0.738	0.731	0.121	0.157
Observations	789	789	789	789
Municipalities	76	76	76	76

Note: The table presents RD estimates on the mean level (first two columns) and standard deviation (last two columns) in resources in municipal schools. *Equipment* and *Infrastructure* are indices constructed by taking the first principal component of a school's equipment and infrastructure elements, calculating the school's percentile in the national distribution, then averaging across schools in the municipality. *Incumbent* is an indicator for whether the mayor is an incumbent and thus term-limited. Observations are at the year level. *Single-round mean* refers to the dependent variable mean for single-round municipalities within the bandwidth. *Estimation method:* Local linear regression with election-year fixed effects and a 50,000 voter bandwidth. Population density included as a control separately across the cutoff. Standard errors clustered at the municipality level.



(a) Total donations, average per candidate

(b) Donations from individuals, average per candidate

(c) Donations from corporations, average per candidat

(d) Total donations, top 2 candidates

(e) Donations from individuals, top 2 candidates

(f) Donations from corporations, top 2 candidates

Figure 19 Regression discontinuity plots of campaign donations received by candidates

Note: The figures plot log donation levels, in *reais*, received by candidates. Donors identified as *Individual* and *Corporation* depending on whether the donor filed an individual or corporate identification number. In each panel, each point plots an average value within a 7,500 voter bin. Variables on the vertical axis are residualized by population density and election-year fixed effects. Diameter of the points is proportional to the number of observations. Confidence intervals (dashed lines) represent the 95% confidence intervals of a local linear regression (solid red line) with standard errors clustered at the municipality level.

C. Theory Appendix

This section contains a stylized model where two-round elections create incentives for politicians to appeal to a broader group of voters and provide public goods differently. The model adapts a standard probabilistic voting model and follows the setup in Genicot, Bouton and Castanheira (2021) by allowing for targeting of government interventions to specific localities within a municipality. The model is extended by (i) introducing a third non-strategic candidate who appeals to a single locality, (ii) allowing candidates to exert effort to increase the municipal budget, and (iii) adapting it to the context of single- and two-round elections.

C.1. The environment

Consider an election with three politicians and $J \geq 3$ localities within a municipality. Politicians are indexed by $c \in \{A, B, C\}$, and localities by $j \in \{1, 2, \cdots, J\}$. Each locality has a continuum of voters of mass $1/J$.

Prior to election day, each politician simultaneously announces a platform that describes (i) the total government budget, G^c, and (ii) the allocation of the government budget to each locality, $\mathbf{q}^c = (q_1^c, q_2^c, \ldots, q_J^c)$, where $q_j^c \geq 0$. The politician's budget constraint is:

$$\sum_{j=1}^{J} q_j^c \leq G^c$$

Since each locality has the same number of voters, each voter receives the same fraction of the government budget allocated to their locality. I assume without loss of generality that voters care about the total amount allocated to their locality, q_j^c. In promising a certain budget, politicians face a cost that is quadratic in the size of the budget:

$$\frac{1}{2}\kappa \left(G^c\right)^2$$

for a constant κ. Platforms are binding for politicians between rounds and after the election.[1]

The third candidate C is a non-strategic candidate with the following platform:

$$\mathbf{q}^C = \left(0, 0, \ldots, 0, G^C\right)$$

[1] This is not unrealistic, as the time between rounds is short compared to the length of the campaign (three weeks in Brazil). This assumption can be relaxed as long as there is some continuity between the two rounds – for example, if voters' second round vote depends on a candidate's policy proposal in both rounds, or if candidates are constrained in the extent to which their proposals can change between rounds.

I assume that G^C is the highest offer in locality J, ie. that candidate A and B's equilibrium allocations to J are smaller than G^C (see Appendix C.5). The strategy of candidate C aligns with an interpretation where C is a small candidate and A and B are front–runners in relation to C.

Voters obtain utility $u(q_j)$ from government spending q_j, where $u(q_j)$ is strictly increasing and concave in q_j. In addition to the policy component of voters' preferences, there is an individual v_i and municipality shock δ toward candidate A, which are independently and uniformly distributed across voters and rounds:

$$v_i \sim U\left[-\frac{1}{2\psi}, \frac{1}{2\psi}\right] \qquad\qquad \delta \sim U\left[-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right]$$

Voters cast a ballot for the politician who offers them the highest payoff. In localities $j \in \{1, \ldots, J-1\}$, ie. where candidate C has not allocated resources, this amounts to voting for either A or B.[2] In locality J, ie. where candidate C is dominant, voters randomize between voting for C with probability $1-\alpha$ and for either A or B with probability α, depending on whether A or B offers the higher payoff, where $0 < \alpha < 1$.[3]

Thus, voters will vote for A if and only if:

$$u(q_j^B) \leq u(q_j^A) + v_i + \delta \tag{C.1}$$

In localities $j \in \{1, \ldots, J-1\}$, all voters for whom this is true vote for A. In locality J, a fraction α of voters for whom this is true vote for A.

C.2. Preliminaries

Define $\Delta u_j^{cd} \equiv u(q_j^c) - u(q_j^d)$. Let π_{jt}^c be candidate c's vote share in locality j in round $t \in \{1, 2\}$ and π_t^c be the total vote share in the municipality for candidate c in round t.

C.2.1. *Assumption: Swing-able voters.–* To derive candidates' vote shares in different election rounds, condition C.1 corresponds to voters for whom $v_i \geq u(q_j^B) - u(q_j^A) - \delta$. To ensure $0 < \pi_{j1}^A < 1$, or that there are voters to be swung in every locality, we need that:

$$u(q_j^B) - u(q_j^A) - \delta \in \left(-\frac{1}{2\psi}, \frac{1}{2\psi}\right)$$

[2] Since $u(\cdot)$ is strictly increasing, candidates A and B will invest a non-zero amount in these localities, and voters will always vote for either A or B.

[3] Assuming $\alpha > 0$ allows a non-zero first order condition for locality J in the single-round election, which enables a direct comparison between single- and two-round elections. This assumption can be relaxed and will yield the same predictions.

Let $u(y)$ be the largest possible utility coming from the allocation of government resources. This assumption is satisfied if:

$$\delta \in \left(u(y) - u(0) - \frac{1}{2\psi}, u(y) - u(0) + \frac{1}{2\psi} \right)$$
$$\iff \frac{1}{2\gamma} + u(y) - u(0) < \frac{1}{2\psi}$$

In other words, that swings in municipality vote shares are smaller than the variation in individual preferences. Note that this implies that $\gamma > \psi$ since $u(y) - u(0) > 0$.

C.2.2. *Vote shares with three candidates.–* In single-round elections and the first round of two-round elections, π_{j1}^{c} is given by:

$$\pi_{j1}^{A} = \begin{cases} \frac{1}{2} + \psi\left(\Delta u_j^{AB} + \delta\right) & \text{if } j \in \{1, \dots, J-1\} \\ \alpha\left(\frac{1}{2} + \psi\left(\Delta u_j^{AB} + \delta\right)\right) & \text{if } j = J \end{cases}$$
$$\pi_{j1}^{B} = \begin{cases} \frac{1}{2} + \psi\left(\Delta u_j^{BA} - \delta\right) & \text{if } j \in \{1, \dots, J-1\} \\ \alpha\left(\frac{1}{2} + \psi\left(\Delta u_j^{BA} - \delta\right)\right) & \text{if } j = J \end{cases}$$
$$\pi_{j1}^{C} = \begin{cases} 0 & \text{if } j \in \{1, \dots, J-1\} \\ 1-\alpha & \text{if } j = J \end{cases}$$

Candidates' total vote share in the municipality π_1^c is given by:

$$\pi_1^A = \left(\frac{J-1+\alpha}{J}\right)\left(\frac{1}{2} + \psi\delta\right) + \frac{\psi}{J}\left(\sum_{j=1}^{J-1} \Delta u_j^{AB} + \alpha \Delta u_J^{AB}\right)$$
$$\pi_1^B = \left(\frac{J-1+\alpha}{J}\right)\left(\frac{1}{2} - \psi\delta\right) + \frac{\psi}{J}\left(\sum_{j=1}^{J-1} \Delta u_j^{BA} + \alpha \Delta u_J^{BA}\right)$$
$$\pi_1^C = \frac{1-\alpha}{J}$$

The probability that candidates A and B attain a vote share above θ is given by:

$$Pr\left(\pi_1^A \geq \theta\right) \equiv Pr\left[\delta \geq \frac{1}{\psi}\left(\frac{J}{J-1+\alpha}\theta - \frac{\psi}{J-1+\alpha}\left(\sum_{j=1}^{J-1} \Delta u_j^{AB} + \alpha \Delta u_J^{AB}\right) - \frac{1}{2}\right)\right]$$
$$= \frac{1}{2} + \frac{\gamma}{\psi}\left[\frac{1}{2} - \frac{J}{J-1+\alpha}\theta + \frac{\psi}{J-1+\alpha}\left(\sum_{j=1}^{J-1} \Delta u_j^{AB} + \alpha \Delta u_J^{AB}\right)\right]$$

$$Pr\left(\pi_1^B \geq \theta\right) \equiv Pr\left[\delta \leq \frac{1}{\psi}\left(\frac{1}{2} - \frac{J}{J-1+\alpha}\theta + \frac{\psi}{J-1+\alpha}\left(\sum_{j=1}^{J-1} \Delta u_j^{BA} + \alpha \Delta u_j^{BA}\right)\right)\right]$$
$$= \frac{1}{2} + \frac{\gamma}{\psi}\left[\frac{1}{2} - \frac{J}{J-1+\alpha}\theta + \frac{\psi}{J-1+\alpha}\left(\sum_{j=1}^{J-1} \Delta u_j^{BA} + \alpha \Delta u_J^{BA}\right)\right]$$

C.2.3. *Vote shares with two candidates.–* In the second round of two-round elections, π_{j2}^c and π_2^c are given by:

$$\pi_{j2}^A = \frac{1}{2} + \psi\left(\Delta u_j^{AB} + \delta\right) \qquad \pi_{j2}^B = \frac{1}{2} + \psi\left(\Delta u_j^{BA} - \delta\right)$$
$$\pi_2^A = \frac{1}{2} + \psi\delta + \frac{\psi}{J}\sum_{j=1}^{J} \Delta u_j^{AB} \qquad \pi_2^B = \frac{1}{2} - \psi\delta + \frac{\psi}{J}\sum_{j=1}^{J} \Delta u_j^{BA}$$

The probability of attaining a vote share above θ is given by:

$$Pr\left(\pi_2^A \geq \theta\right) \equiv Pr\left[\delta \geq \frac{1}{\psi}\left(\theta - \frac{1}{2} - \frac{\psi}{J}\sum_{j=1}^{J} \Delta u_j^{AB}\right)\right]$$
$$= \frac{1}{2} + \frac{\gamma}{\psi}\left(\frac{1}{2} - \theta + \frac{\psi}{J}\sum_{j=1}^{J} \Delta u_j^{AB}\right)$$
$$Pr\left(\pi_2^B \geq \theta\right) \equiv Pr\left[\delta \leq \frac{1}{\psi}\left(\frac{1}{2} - \theta + \frac{\psi}{J}\sum_{j=1}^{J} \Delta u_j^{BA}\right)\right]$$
$$= \frac{1}{2} + \frac{\gamma}{\psi}\left(\frac{1}{2} - \theta + \frac{\psi}{J}\sum_{j=1}^{J} \Delta u_j^{BA}\right)$$

C.2.4. *Assumption: Contestable localities.–* For $0 < Pr(\pi_1^A \geq \theta) < 1$ and $0 < Pr(\pi_1^B \geq \theta) < 1$, or that all localities are contestable in the first round, we need that:

$$\frac{1}{\psi}\left[\frac{J}{J-1+\alpha}\theta - \frac{1}{2} - \frac{\psi}{J-1+\alpha}\left(\sum_{j=1}^{J-1} \Delta u_j^{AB} + \alpha \Delta u_J^{AB}\right)\right] \in \left(-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right)$$
$$\frac{1}{\psi}\left[\frac{1}{2} - \frac{J}{J-1+\alpha}\theta + \frac{\psi}{J-1+\alpha}\left(\sum_{j=1}^{J-1} \Delta u_j^{BA} + \alpha \Delta u_J^{BA}\right)\right] \in \left(-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right)$$

which corresponds to the following condition for the first round:

$$\theta \in \left(\left(\frac{J-1+\alpha}{J}\right)\left(-\frac{\psi}{2\gamma} + \frac{1}{2}\right), \left(\frac{J-1+\alpha}{J}\right)\left(\frac{\psi}{2\gamma} + \frac{1}{2}\right)\right) \tag{C.2}$$

For $0 < Pr\left(\pi_2^A \geq \theta\right) < 1$ and $0 < Pr\left(\pi_2^B \geq \theta\right) < 1$, or that all localities are contestable in the second round, we need that:

$$\frac{1}{\psi}\left[\theta - \frac{1}{2} - \frac{\psi}{J}\sum_{j=1}^{J}\Delta u_j^{AB}\right] \in \left(-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right)$$
$$\frac{1}{\psi}\left[\frac{1}{2} - \theta + \frac{\psi}{J}\sum_{j=1}^{J}\Delta u_j^{BA}\right] \in \left(-\frac{1}{2\gamma}, \frac{1}{2\gamma}\right)$$

which corresponds to the following condition for the second round:

$$\theta \in \left(-\frac{\psi}{2\gamma} + \frac{1}{2}, \frac{\psi}{2\gamma} + \frac{1}{2}\right) \tag{C.3}$$

Claim C.1. $\theta = \frac{1}{2}\left(1 - \frac{1-\alpha}{J}\right)$, *the vote share required to attain the most votes with three candidates, satisfies condition (C.2).*

Both the upper and lower inequalities are satisfied because:

$$-\frac{\psi}{2\gamma} + \frac{1}{2} < \frac{1}{2} < \frac{\psi}{2\gamma} + \frac{1}{2}$$

Claim C.2. $\theta = \frac{1}{2}$, *the vote share required to win in the first round of a two-round election, satisfies condition (C.2).*

The lower inequality is satisfied:

$$\left(\frac{J-1+\alpha}{J}\right)\left(-\frac{\psi}{2\gamma} + \frac{1}{2}\right) < \frac{1}{2}$$

because $\frac{J-1+\alpha}{J} < 1$ and $-\frac{\psi}{2\gamma} + \frac{1}{2} < \frac{1}{2}$.

The upper inequality is equivalent to:

$$\frac{1}{2} < \left(\frac{J-1+\alpha}{J}\right)\left(\frac{\psi}{2\gamma} + \frac{1}{2}\right)$$
$$\iff J > (1-\alpha)\left(\frac{\gamma+\psi}{\psi}\right)$$

which is true so long as J is large enough and γ/ψ is not too large.

Claim C.3. $\theta = \frac{1}{2}$, *the vote share required to win in the second round of a two-round election,* *satisfies condition (C.3).*

Both the upper and lower inequalities are satisfied because:

$$-\frac{\psi}{2\gamma}+\frac{1}{2}<\frac{1}{2}<\frac{\psi}{2\gamma}+\frac{1}{2}$$

C.2.5. *Assumption: C never makes it to the second round.*– With three candidates, I assume that candidate C always receives the lowest vote share: C never wins a single-round election nor makes it to the second round in a two-round election. To ensure this, the probability that candidates A and B attain vote shares above candidate C's must be 1, or π_1^C does not satisfy condition (C.2):

$$\frac{1-\alpha}{J}\leq\left(\frac{J-1+\alpha}{J}\right)\left(-\frac{\psi}{2\gamma}+\frac{1}{2}\right)\qquad\text{or}\qquad\frac{1-\alpha}{J}\geq\left(\frac{J-1+\alpha}{J}\right)\left(\frac{\psi}{2\gamma}+\frac{1}{2}\right)$$

The first inequality (left) and second inequality (right) are equivalent to:

$$J\geq(1-\alpha)\left(\frac{2\gamma}{\gamma-\psi}+1\right)\qquad\qquad J\leq(1-\alpha)\left(\frac{2\gamma}{\gamma+\psi}+1\right)$$

The first inequality is much more likely to be satisfied, which is true so long as J is large enough and $2\gamma/(\gamma-\psi)$ is not too large.

C.3. Equilibrium strategies

Candidates' payoff is 1 if they win the election and 0 otherwise, minus the effort cost incurred during the campaign. Candidates maximize their expected payoff, so this amounts to maximizing the probability of winning minus the effort cost.

C.3.1. *In a single-round election.*– Candidate C attains a vote share of $\frac{1-\alpha}{J}$, so the probability of winning is the probability of attaining over half of the remaining votes:

$$Pr\left(\pi_1^c\geq\frac{1}{2}\left(1-\frac{1-\alpha}{J}\right)\right)\qquad\text{for } c\in\{A,B\}$$

For $c\in\{A,B\}$ and $d\in\{B,A\}$, the maximization problem is:

$$\max_{G^c,\mathbf{q}^c=\left(q_1^c,\ldots,q_J^c\right)}\ \frac{1}{2}+\left(\frac{\gamma}{J-1+\alpha}\right)\left(\sum_{j=1}^{J-1}\Delta u_j^{cd}+\alpha\Delta u_J^{cd}\right)-\frac{1}{2}\kappa\left(G^c\right)^2\qquad\text{s.t.}\sum_j q_j^c\leq G^c$$

which corresponds to the following first-order conditions:

$$
\begin{aligned}
&\left(\frac{\gamma}{J-1+\alpha}\right) u'(q_j^c) = \lambda_{1R} && \text{for } j \in \{1,\ldots,J-1\} \\
&\left(\frac{\gamma}{J-1+\alpha}\right) \alpha u'(q_j^c) = \lambda_{1R} && \text{for } j = J \\
&\kappa G^c = \lambda_{1R} && \text{(C.4)}
\end{aligned}
$$

where λ_{1R} is the Lagrange multiplier of the budget constraint in a single-round system.

The ratio in marginal utilities between localities is:

$$
\begin{aligned}
&\text{between } j \text{ and } j'\text{:} \quad \frac{u'(q_j^c)}{u'(q_{j'}^c)} = 1 && \forall \quad j, j' \in \{1,\ldots,J-1\} \\
&\text{between } j \text{ and } J\text{:} \quad \frac{u'(q_j^c)}{u'(q_J^c)} = \alpha && \forall \quad j \in \{1,\ldots,J-1\} \qquad \text{(C.5)}
\end{aligned}
$$

Since $u(\cdot)$ is strictly increasing and strictly concave, equation C.5 yields the following prediction:

Prediction C.4. *In a single-round election, for all $j \in \{1,\ldots,J-1\}$, we have that $q_j^c > q_J^c$ for $c \in \{A, B\}$.*

C.3.2. *In a two-round election.–* The probability of winning is the probability of attaining a vote share above $\frac{1}{2}$ in the first round or second round, should one occur:

$$
\begin{aligned}
&Pr(c \text{ wins in 1st round}) + Pr(\text{second round occurs}) \cdot Pr(c \text{ wins 2nd round}) \\
&= Pr\left(\pi_1^c \geq \frac{1}{2}\right) + \left(1 - Pr\left(\pi_1^c \geq \frac{1}{2}\right) - Pr\left(\pi_1^d \geq \frac{1}{2}\right)\right) Pr\left(\pi_2^c \geq \frac{1}{2}\right) \\
&\quad \text{for } c \in \{A, B\} \text{ and } d \in \{B, A\}
\end{aligned}
$$

The maximization problem is:

$$
\begin{aligned}
\max_{G^c, \mathbf{q}^c = \left(q_1^c,\ldots,q_J^c\right)} \quad & \left(\frac{1}{2} + \frac{\gamma}{\psi}\left[\frac{1}{2}\left(\frac{\alpha-1}{J-1+\alpha}\right) + \left(\frac{\psi}{J-1+\alpha}\right)\left(\sum_{j=1}^{J-1} \Delta u_j^{cd} + \alpha \Delta u_J^{cd}\right)\right]\right) \\
& + \frac{\gamma}{\psi}\left(\frac{1-\alpha}{J-1+\alpha}\right)\left[\frac{1}{2} + \frac{\gamma}{J}\sum_{j=1}^{J} \Delta u_j^{cd}\right] - \frac{1}{2}\kappa \left(G^c\right)^2 \\
\text{s.t.} \quad & \sum_j q_j^c \leq G^c
\end{aligned}
$$

which corresponds to the following first-order conditions:

$$\begin{aligned}
&\left(\frac{\gamma}{J-1+\alpha}\right)\left(1+\frac{(1-\alpha)\gamma}{\psi J}\right)u'(q_j^c)=\lambda_{2R} && \text{for } j\in\{1,\dots,J-1\}\\
&\left(\frac{\gamma}{J-1+\alpha}\right)\left(\alpha+\frac{(1-\alpha)\gamma}{\psi J}\right)u'(q_j^c)=\lambda_{2R} && \text{for } j=J\\
&\kappa G^c=\lambda_{2R} && \text{(C.6)}
\end{aligned}$$

where λ_{2R} is the Lagrange multiplier of the budget constraint in a two-round system.

The ratio in marginal utilities between localities is:

$$\begin{aligned}
&\text{between } j \text{ and } j': && \frac{u'(q_j^c)}{u'(q_{j'}^c)}=1 && \forall \quad j,j'\in\{1,\dots,J-1\}\\
&\text{between } j \text{ and } J: && \frac{u'(q_j^c)}{u'(q_J^c)}=\frac{\alpha+\frac{(1-\alpha)\gamma}{\psi J}}{1+\frac{(1-\alpha)\gamma}{\psi J}} && \forall \quad j\in\{1,\dots,J-1\} && \text{(C.7)}
\end{aligned}$$

Since $u(\cdot)$ is strictly increasing and strictly concave, equation C.7 yields prediction C.5:

PREDICTION C.5. *In a two-round election, for all $j\in\{1,\dots,J-1\}$, we have that $q_j^c>q_J^c$ for $c\in\{A,B\}$.*

C.4. Comparing single- to two-round elections

I compare two outcomes under a single- and two-round election: (i) politician's allocations to localities and (ii) politician's choice of the overall budget. To simplify notation, denote the equilibrium allocations and overall budget as q_j^{1R} and G^{1R} (single-round elections) and q_j^{2R} and G^{2R} (two-round elections).

C.4.1. *Preliminaries.–* I first establish three lemmas.

LEMMA C.6. $\dfrac{u'(q_j^{1R})}{u'(q_j^{2R})}\dfrac{G^{2R}}{G^{1R}}>1$ *for all $j\in\{1,\dots,J\}$.*

Proof. For $j\in\{1,\dots,J-1\}$, combining the first-order conditions in equations C.4 and C.6:

$$\frac{u'(q_j^{1R})}{u'(q_j^{2R})}\frac{G^{2R}}{G^{1R}}=1+\frac{(1-\alpha)\gamma}{\psi J}>1$$

For $j = J$, combining the first-order conditions in equations C.4 and C.6:

$$\frac{u'(q_J^{1R})}{u'(q_J^{2R})}\frac{G^{2R}}{G^{1R}} = 1 + \frac{(1-\alpha)\gamma}{\alpha\psi J} > 1$$

□

Lemma C.7. $\frac{u'(q_j^{1R})}{u'(q_j^{2R})} < \frac{u'(q_J^{1R})}{u'(q_J^{2R})}$.

Proof. Comparing the ratio of marginal utilities in equations (C.5) and (C.7), the ratio is smaller in the single-round system compared to the two-round system:

$$\frac{u'(q_j^{1R})}{u'(q_J^{1R})} < \frac{u'(q_j^{2R})}{u'(q_J^{2R})} \iff \alpha < \frac{\alpha + \frac{(1-\alpha)\gamma}{\psi J}}{1 + \frac{(1-\alpha)\gamma}{\psi J}}$$

which is true because $\alpha < 1$. □

Lemma C.8. *If $q_j^{1R} > q_j^{2R}$ for one $j \neq J$ then $q_{j'}^{1R} > q_{j'}^{2R}$ for all other $j' \in \{1, \ldots, J-1\}$.*

Proof. If $q_j^{1R} > q_j^{2R}$, then $u'(q_j^{1R}) < u'(q_j^{2R})$ because $u(\cdot)$ is strictly concave. From the first-order conditions in equations C.4 and C.6, the marginal utilities between all $j, j' \in \{1, \ldots, J-1\}$ are equal. Then we must have that $u(q_{j'}^{1R}) < u(q_{j'}^{2R})$ and that $q_{j'}^{1R} > q_{j'}^{2R}$. □

C.4.2. *Allocations to localities.–* Prediction C.9 states that candidates promise more to locality J in a two-round election than in a single-round election.

Prediction C.9. $q_J^{1R} < q_J^{2R}$ *for all $j \in \{1, \ldots, J-1\}$.*

Proof. I prove by contradiction. Assume that $q_J^{1R} \geq q_J^{2R}$. Then $u'(q_J^{1R}) \leq u'(q_J^{2R})$. By lemma C.6, we must have that $G^{2R} > G^{1R}$. To satisfy the budget constraint, we must have that $q_j^{1R} < q_j^{2R}$ for some $j \neq J$ and, by lemma C.8, for all $j \neq J$. Then $u'(q_j^{1R}) > u'(q_j^{2R})$. However, this violates lemma C.7, and so we must have $q_J^{1R} < q_J^{2R}$. □

In a single-round election, because voters in locality J strongly favor candidate C, the marginal return to allocating resources there is low and candidates A and B ignore these voters. In the two-round election, there is the possibility of a second round where C is not present which results in a higher marginal return to allocating resources to that locality. As a result, while not offering a completely equitable distribution, candidates A and B solicit more votes from locality J in a two-round election, even in the first round. This increased allocation in two-round elections to locality J is a force to reduce inequality in allocations across localities.

C.4.3. *Overall government budget.–* Prediction C.10 states that the overall government budget promised is higher in two-round elections than in single-round elections.

PREDICTION C.10. $G^{1R} < G^{2R}$.

Proof. I prove by contradiction. Assume that $G^{1R} \geq G^{2R}$. Since $q_J^{1R} < q_J^{2R}$, to satisfy the budget constraint, we must have that $q_j^{1R} > q_j^{2R}$ for some $j \neq J$ and, by lemma C.8, for all $j \neq J$. Then $u'(q_j^{1R}) < u'(q_j^{2R})$. However, this violates lemma C.6, and so we must have $G^{1R} < G^{2R}$. □

In a two-round election, candidates face higher incentives to invest in *all* localities. These incentives result from the fact that, in two-round elections, there is a conditionality to winning: to win in the first round, candidates must not only attain the most votes, but must attain a majority of votes, otherwise candidates must compete again in a second round. As a result, candidates in two-round elections exert more effort to increase the government budget.

C.5. Assumption: Candidate C's budget

In general, for every utility function $u(\cdot)$, there exists a G^C such that G^C is the highest offer in locality J. I show this for the case where $u(\cdot) = \beta \ln(\cdot)$ and for the two-round election (since q_J is higher in a two-round election).

The first order conditions with respect to q_J and G are:

$$\begin{aligned}
&\left(\frac{\gamma}{J-1+\alpha}\right)\left(\alpha + \frac{(1-\alpha)\gamma}{\psi J}\right)\frac{\beta}{q_J} = \lambda_{2R} \\
&\kappa G = \lambda_{2R}
\end{aligned}$$

We can write G as a function of q_J:

$$G = \frac{1}{\kappa}\left(\frac{\gamma}{J-1+\alpha}\right)\left(\alpha + \frac{(1-\alpha)\gamma}{\psi J}\right)\frac{\beta}{q_J}$$

Since $q_J < q_j$ for all $j \neq J$ (prediction C.5), then $q_J < G/J$, implying:

$$q_J < \left(\frac{1}{\kappa J}\left(\frac{\gamma}{J-1+\alpha}\right)\left(\alpha + \frac{(1-\alpha)\gamma}{\psi J}\right)\beta\right)^{1/2} \equiv \Gamma$$

So long as $G^C > \Gamma$, then candidate C's allocation to locality J will be the highest offer there. This will be true so long as κ or J is large enough and γ/ψ is not too large.

References

Codato, Adriano, Fábia Berlatto, and Bruno Bolognesi. 2018. "Tipologia dos políticos de direita no Brasil: uma classificação empírica." *Análise Social*, 53: 870–897.

Genicot, Garance, Laurent Bouton, and Micael Castanheira. 2021. "Electoral Systems and Inequalities in Government Interventions." *Journal of the European Economic Association*, 19: 3154–3206.