ONLINE APPENDIX

Fiscal Rules and the selection of politicians: theory and evidence from Italy

Matteo Gamalerio

Institut d'Economia de Barcelona (IEB), University of Barcelona

&

Federico Trombetta

Department of International Economics, Institutions and Development (DISEIS), Faculty of Political and Social Sciences, Università Cattolica del Sacro Cuore, Milano

A Appendix: Additional tables and figures

Table A1: Descriptive statistics:
Municipalities below 5000 vs. Municipalities above 5000

	(1) Below 5000	(2) obs	(3) Above 5000	(4) obs	(5) p-value
Politicians characteristics					
Female mayors	0.088	4836	0.094	1334	0.230
Age mayors	48.236	4836	47.786	1334	0.023
High skills job mayors	0.228	4836	0.310	1334	0.000
Graduate mayors	0.374	4836	0.516	1334	0.000
Political experience mayors	8.329	4836	8.226	1334	0.381
Female mayoral candidates	0.105	4836	0.110	1334	0.310
Age mayoral candidates	48.110	4836	48.076	1334	0.814
High skills job mayoral candidates	0.213	4836	0.307	1334	0.000
Graduate mayoral candidates	0.356	4836	0.500	1334	0.000
Municipal characteristics					
South	0.253	4836	0.289	1334	0.008
Centre	0.136	4836	0.166	1334	0.006
North-West	0.504	4836	0.307	1334	0.000
North-East	0.107	4836	0.239	1334	0.000
Population density	146.931	4836	496.301	1334	0.000
Area	25.328	4836	43.145	1334	0.000
No profit associations	0.005	4836	0.004	1334	0.000
Firms per capita	0.067	4836	0.076	1334	0.000
Income per capita	8907	4836	9795	1334	0.000
% elderly	0.229	4836	0.177	1334	0.000
% 15-64 years old	0.643	4836	0.677	1334	0.000
% graduate	0.043	4836	0.052	1334	0.000

Notes. Municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. *Below 5000* = 1 for municipalities below 5,000 inhabitants. *Above 5000* = 1 for municipalities above 5,000 inhabitants. Columns (1) and (3) report the mean values for the two samples; *obs* is the number of observations; *p-value* is the p-value of the difference between the means of the two samples.

Table A2: Balance test on municipal covariates

Diff-in-Disc

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Characteristics municipal population							
Dependent variables	% university degree	% 15-64	% 65+	(log) income per capita	# firms	no-profit ass	area	population density
(Post)*(> 5000)	-0.000	-0.007	-0.002	-0.022	0.000	-0.000	0.159	-7.468
	(0.002)	(0.004)	(0.006)	(0.043)	(0.003)	(0.000)	(5.195)	(51.177)
Observations	2,950	2,644	2,509	4,172	2,810	1,716	3,536	3,465
Bandwidth	954.3	859.8	818.2	1346	909.8	568.3	1151	1131
	Panel B: Geographical characteristics municipalities, deficit and re-election/re-run status							
Dependent variables	NE	NW	CEN	SOU	Second term mayor	Re-elected		
(Post)*(> 5000)	-0.065	0.014	0.057	0.032	0.089	-0.031		
	(0.050)	(0.061)	(0.040)	(0.064)	(0.068)	(0.070)		
Observations	3,865	3,772	5,105	3,182	2,706	2,819		
Bandwidth	1246	1220	1644	1034	879.8	1418		

Notes. Diff-in-disc estimates of the impact of fiscal rules on municipal covariates. Municipalities between 0-15,000. Electoral years between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the municipality level are in parentheses.

Table A3: Jump of higher wage indicator at the threshold

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	No	Yes	Yes
Region FE	No	No	Yes
Dependent variable = 1 if mayor paid higher wage			
(> 5000)	0.335	0.329	0.316
	(0.071)	(0.070)	(0.071)
(Post)*(> 5000)	0.041	0.035	0.027
	(0.099)	(0.093)	(0.092)
Observations	1,418	1,418	1,418
Bandwidth	466.3	466.3	466.3

Notes. Diff-in-disc estimates of the impact of fiscal rules on higher wage indicator. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting after 2001. The outcome variable is a dummy variable equal to 1 if mayor receive higher wage, according to the Census population. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Table A4: Cross-sectional RDD coefficients over time

	(1)	(2)	(3)	(4)	(5)
Control Function	Linear	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT	CCT
Election Year FE	No	No	No	No	No
Region FE	No	No	No	No	No
Election	-2	-1	0	1	2
Panel A: Mayoral candidates with university degree					
($>$ 5000)	0.082	0.096	0.011	-0.038	0.032
	(0.046)	(0.044)	(0.048)	(0.049)	(0.080)
Observations	775	802	783	774	267
Bandwidth	1114	1114	1114	1114	1114
Panel B: Mayors with university degree					
($>$ 5000)	0.060	0.069	-0.101	0.004	-0.057
	(0.059)	(0.060)	(0.065)	(0.059)	(0.109)
Observations	984	1,022	1,001	1,005	327
Bandwidth	1425	1425	1425	1425	1425

Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) ($>$ 5000) = 1 for municipalities with more than 5000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in Panel A, and a dummy variable equal to 1 for mayors with a university degree in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Specifically, we run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 4. Robust standard errors clustered at the local labor area level are in parentheses.

Table A5: The effect of fiscal rules on the education of politicians
Difference-in-differences estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Election Year FE	No	No	Yes	No	No	Yes
Municipal FE	No	No	Yes	No	No	Yes
Dependent	Share mayoral candidates with university degree			= 1 for Mayors with university degree		
(>5000)	0.145	0.154	0.028	0.135	0.140	-0.017
	(0.012)	(0.013)	(0.021)	(0.016)	(0.017)	(0.038)
(Post)	0.037	0.035		0.023	0.020	
	(0.005)	(0.006)		(0.007)	(0.008)	
(Post)*(>5000)	-0.010	-0.019	-0.004	0.007	0.001	0.023
	(0.010)	(0.011)	(0.011)	(0.016)	(0.019)	(0.020)
Pre		-0.004			-0.006	
		(0.005)			(0.006)	
(Pre)*(>5000)		-0.018	-0.011		-0.011	0.003
		(0.011)	(0.011)		(0.015)	(0.015)
Observations	26,005	26,005	26,005	26,005	26,005	26,005

Notes. Difference-in-differences estimates of the impact of fiscal rules on the education of politicians. Municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (>5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001; 3) (Pre) = 1 for election immediately before 2001 fiscal rules removal. The outcome variable is the share of mayoral candidates with a university degree in columns 1-3, and a dummy variable equal to 1 for mayors with a university degree in columns 4-6. Robust standard errors clustered at the local labor area level are in parentheses.

Table A6: Introduction of fiscal rules

	(1)	(2)
Dependent Variables	Share mayoral candidates with university degree	= 1 for Mayors with university degree
Control Function	Linear	Linear
Bandwidth	CCT	CCT
Election Year FE	No	No
Region FE	No	No
(>= 1999)*(> 5000)	-0.042	-0.043
	(0.061)	(0.077)
Observations	1,926	2,210
Bandwidth	1335	1534
Mean outcome	0.441	0.464

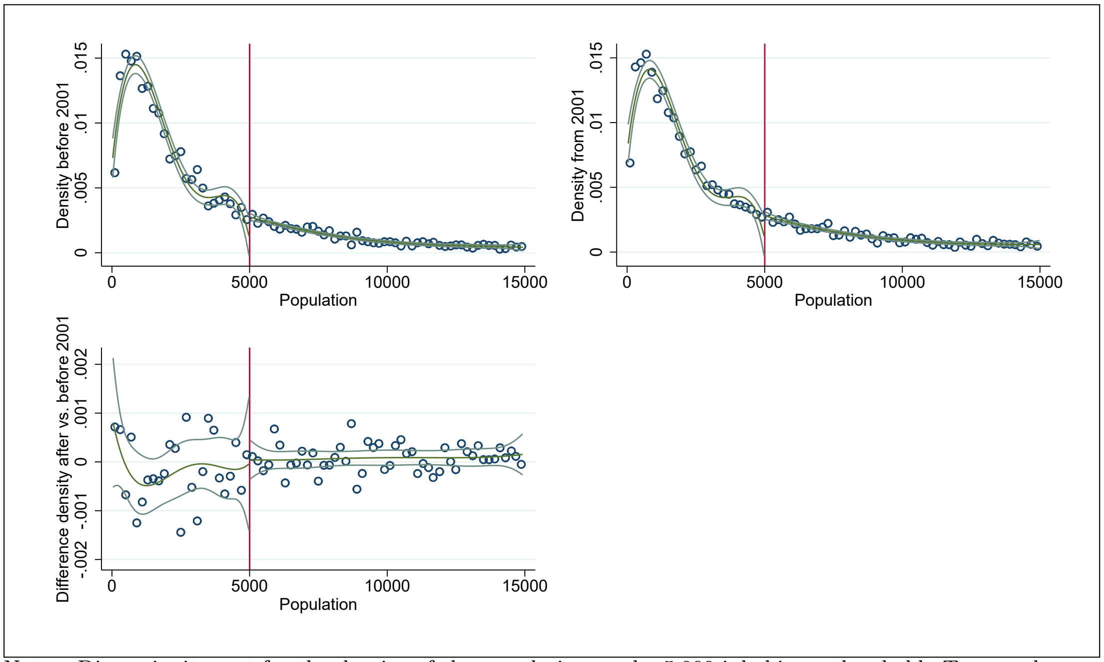
Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2000. Variables in the Table: (>= 1999)*(> 5000)= interaction between dummy = 1 for electoral years 1999-2000 and dummy = 1 for municipalities with more than 5,000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in column 1, while it is equal to 1 for mayors with a university degree in column 2. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Table A7: Heterogeneity analysis based on municipal budget rigidity
Alternative measures

	(1)	(2)	(3)	(4)
Dependent variable	*Mayoral candidates with university degree*		*Mayors with university degree*	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Sample	*Rigidity<median*	*Rigidity>median*	*Rigidity<median*	*Rigidity>median*
Panel A: personnel expenditures as share of current revenues				
(Post)*(> 5000)	-0.221	0.008	-0.224	-0.006
	(0.051)	(0.061)	(0.085)	(0.108)
Observations	1,959	2,031	2,022	1,632
Bandwidth	1127	1500	1160	1217
Mean outcome	0.425	0.516	0.462	0.533
Panel B: debt repayment expenditures as share of current revenues				
(Post)*(> 5000)	-0.125	-0.076	-0.188	-0.066
	(0.052)	(0.091)	(0.089)	(0.107)
Observations	2,214	944	2,110	1,570
Bandwidth	1369	663	1299	1077
Mean outcome	0.474	0.431	0.493	0.486

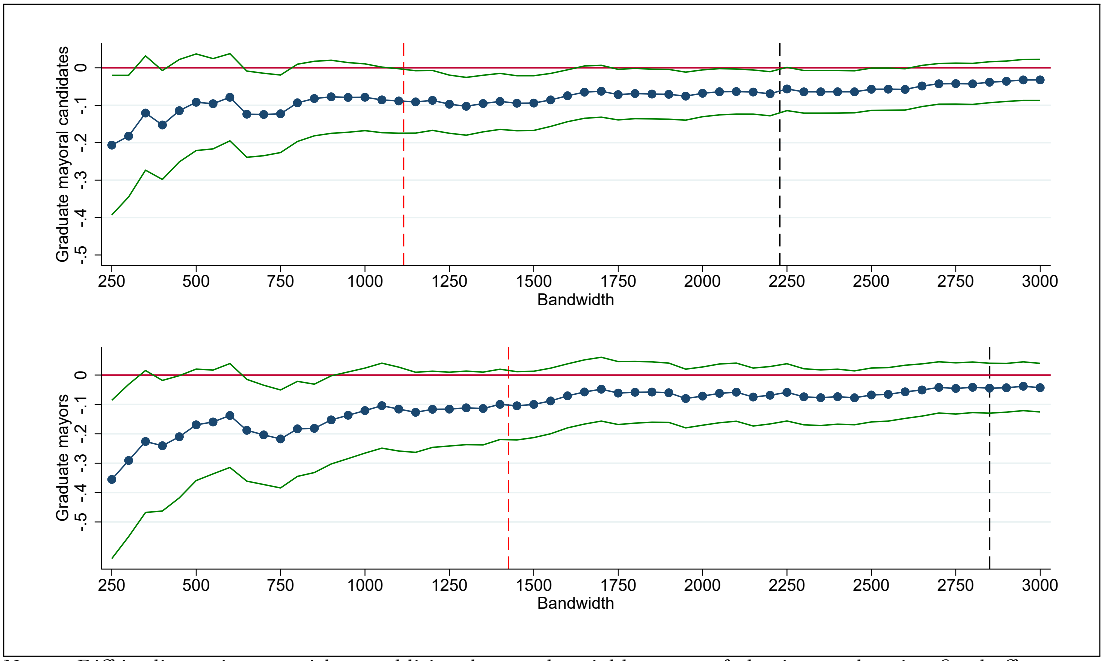
Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Sub-samples: 1) *(Rigidity < median)* = municipalities with a below-median level of personnel (Panel A) or debt (Panel B) expenditures as a fraction of total current revenues; 2) *(Rigidity > median)* = municipalities with an above-median level of personnel (Panel A) and debt (Panel B) expenditures as a fraction of total current revenues. Personnel expenditures as a share of current revenues have an average value of 30.1 percent and debt repayment expenditures as a share of current revenues have an average value equal to 8.2 percent. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in columns 1-2, and a dummy variable equal to 1 for mayors with a university degree in columns 3-4. Year of election and region fixed effects not included. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Figure A1: Density test on the running variable



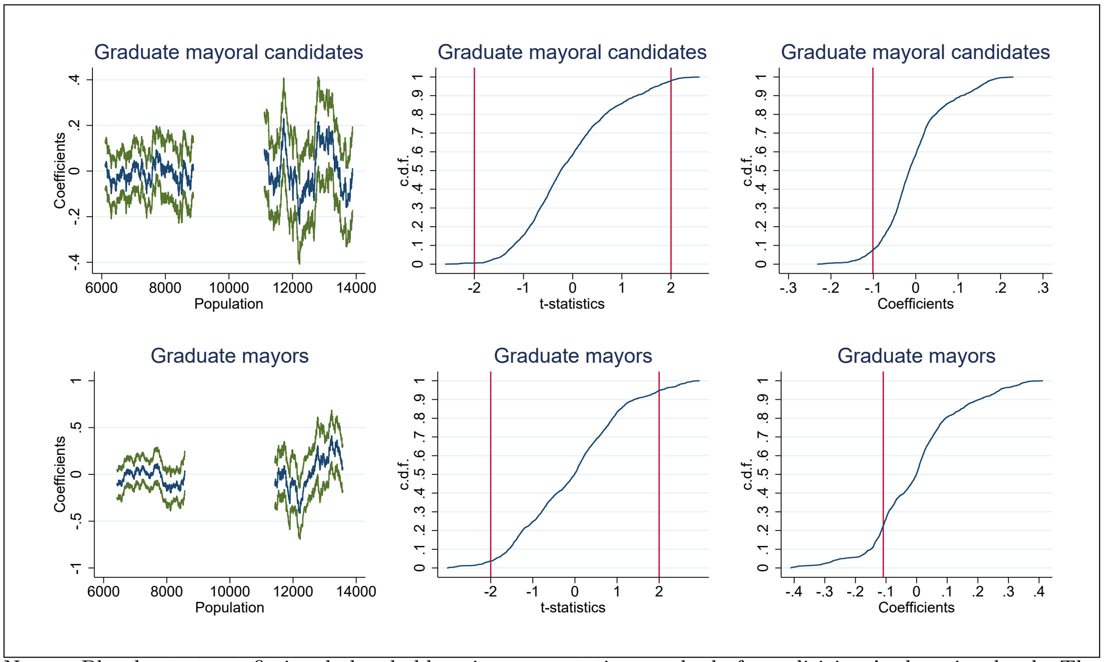
Notes. Discontinuity test for the density of the population at the 5,000-inhabitant threshold. Top graphs: (1) density test for R_{it} before 2001; (2) density test for R_{it} from 2001. Bottom graph: (1) discontinuity test for the difference between the density of average R_{it} from 2001 and the density of average R_{it} before 2001. The central green line represents a split fourth-order polynomial of the outcome variable in the normalized population, fitted separately on each side of the threshold. The grey lines represent the 95 percent confidence interval.

Figure A2: Diff-in-disc estimates: different bandwidths



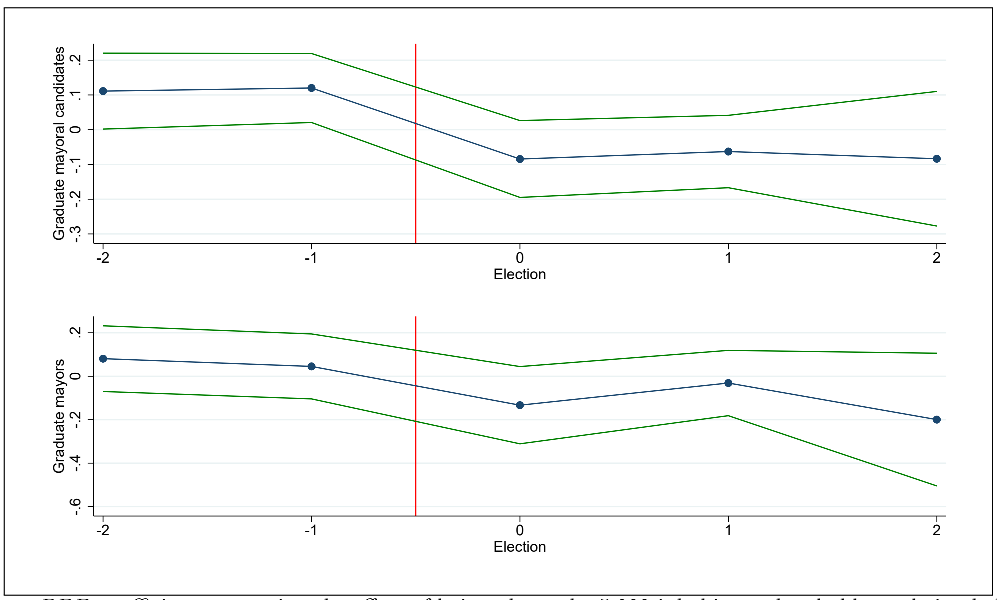
Notes. Diff-in-disc estimates without additional control variables, year of election, and region fixed effects. Horizontal axis: different bandwidths used to estimate the diff-in-disc coefficients. Vertical axis: diff-in-disc coefficients. Dashed red vertical line: optimal bandwidth calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Dashed black vertical line: double the optimal bandwidth. The central blue lines connect the estimated coefficients, while the green lines the 95 percent confidence intervals.

Figure A3: Diff-in-Disc
Placebo thresholds



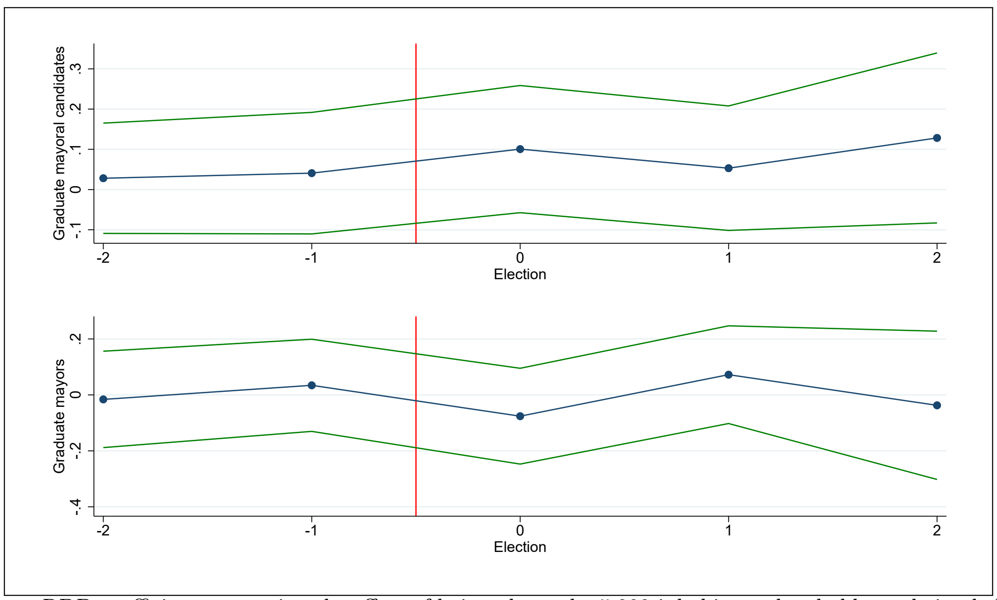
Notes. Placebo tests at fictional thresholds using permutation methods for politicians' education level. The figure reports the estimated coefficients, and c.d.f. of the t-statistics and estimated coefficients of a set of diff-in-disc regressions at 5,542 fictional thresholds for mayoral candidates and 4298 for mayors. The diff-in-disc model is run using a local linear regression with election year and region fixed effects. The graphs on the left report the estimated coefficients at the placebo thresholds with the corresponding population on the x-axis. In these graphs, the central blue lines represent the estimated coefficients, and the green lines the 95 percent confidence intervals. The graphs in the middle report the c.d.f. of the t-statistics associated with these coefficients. The vertical lines in these graphs indicate t-statistics of -2 and 2. The graphs on the right report the c.d.f. of the estimated coefficients. The vertical lines in these graphs indicate the benchmark estimates from Table 4, columns 3.

Figure A4: Cross-sectional RDD coefficients over time (low-rigidity sample)



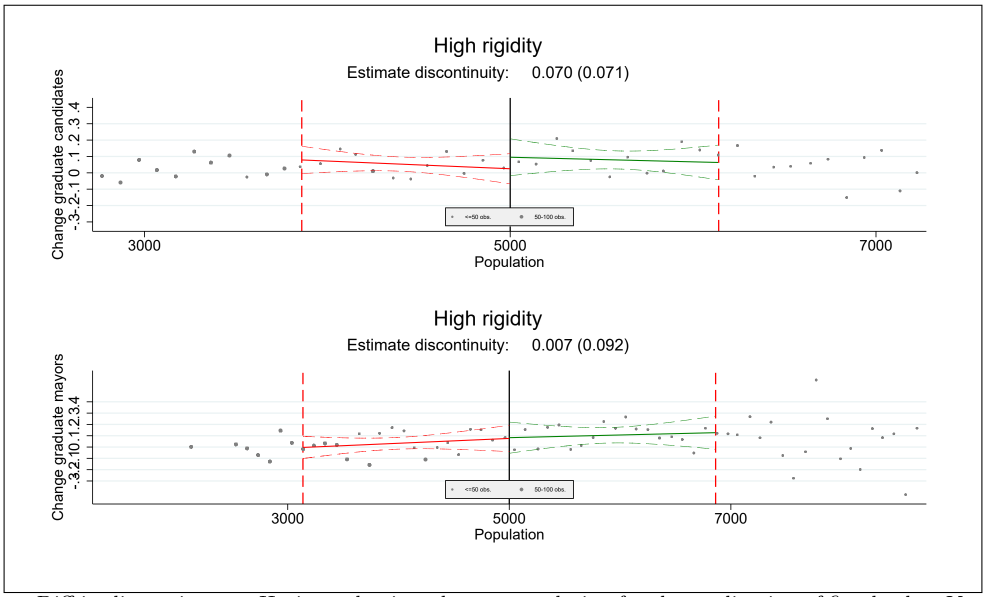
Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. On the x-axis, which goes from -2 to 2, we report the elections before and after the 2001 removal of fiscal rules, where 0 indicates the elections immediately after the relaxation of fiscal rules. We run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 5. The blue lines connect the estimated coefficients, while the green lines represent the 95 percent confidence interval.

Figure A5: Cross-sectional RDD coefficients over time (high-rigidity sample)



Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. On the x-axis, which goes from -2 to 2, we report the elections before and after the 2001 removal of fiscal rules, where 0 indicates the elections immediately after the relaxation of fiscal rules. We run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 5. The blue lines connect the estimated coefficients, while the green lines represent the 95 percent confidence interval.

Figure A6: Diff-in-Disc graphical evidence (high-rigidity sample)



Notes. Diff-in-disc estimates. Horizontal axis: relevant population for the application of fiscal rules. Vertical axis: the change over time in the share of mayoral candidates (top graps) and mayors (bottom graph) with a university degree. Scatter points are averaged over bins of 100 inhabitants. The central line represents a linear regression of the outcome variable in the population, fitted separately on each side of the threshold. The other two dashed lines represent 95 percent confidence intervals. The vertical dashed lines indicates the limit of the optimal bandwidth used in the regressions. Number of observations in each graph: 1) top graph: 3339 observations in total, 1510 within the optimal bandwidth; 2) bottom graph: 7367 observations in total, 2578 within the optimal bandwidth.

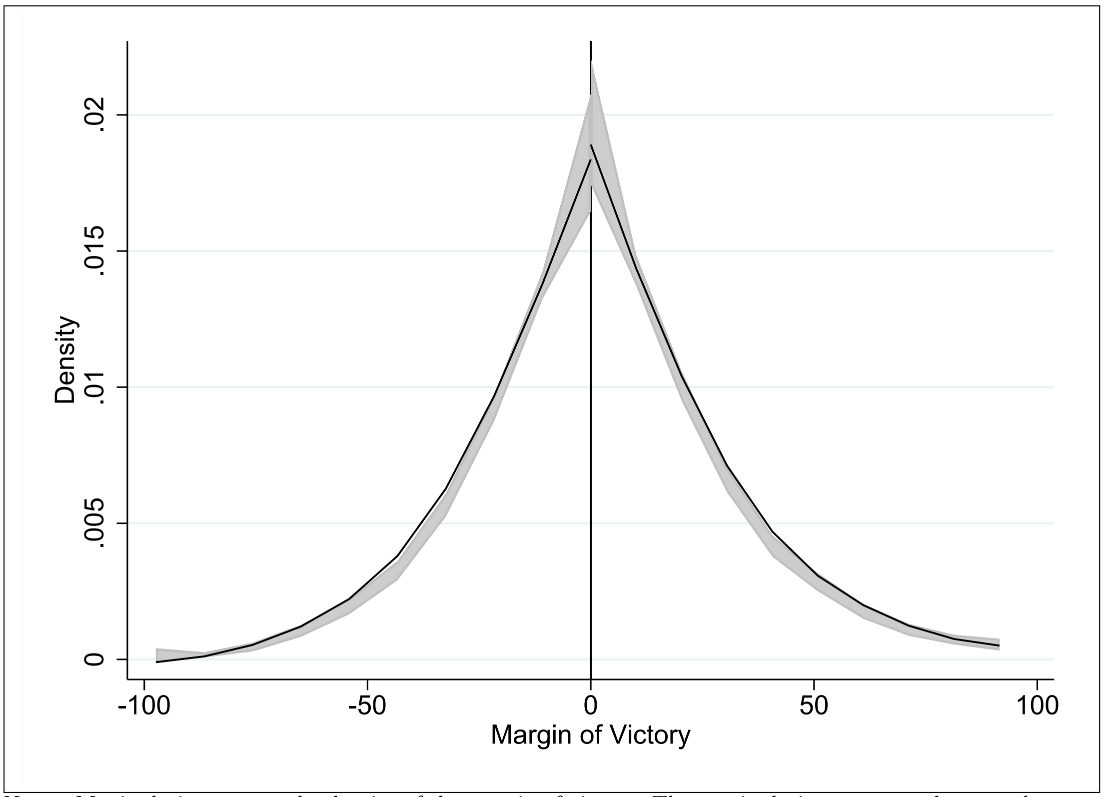
B Appendix: Matching the state of economy

B.1 Regression discontinuity design: main assumptions

There are two main assumptions required for the identification strategy described by model (12) in section IV.C.IV to work correctly. First, there must be no sorting around the threshold $MV_{it} = 0$, such that voters in municipalities with narrow mixed electoral competitions are not able to manipulate the running variable MV_{it}. We test this assumption in Figures B1 and B2, using the test on the continuity of the density of the running variable proposed by Cattaneo, Jansson, and Ma (2018). The evidence in Figures B1-B2 excludes that sorting is happening.

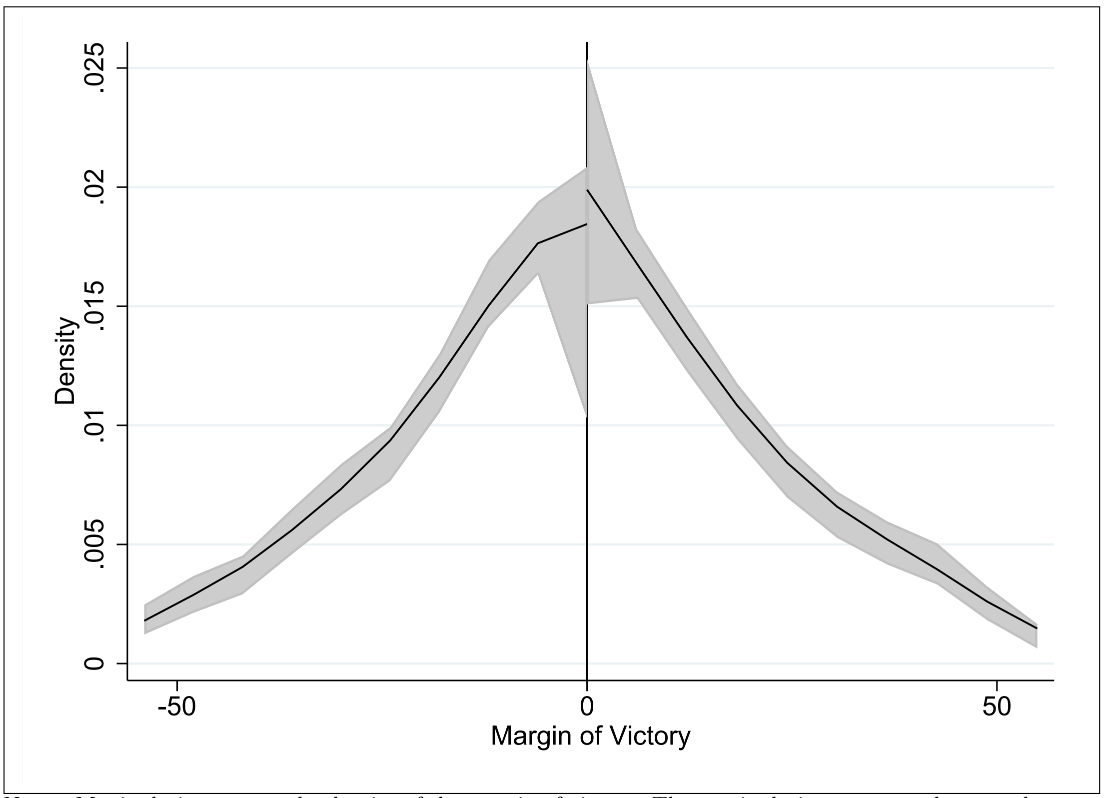
Second, observable municipal characteristics should vary smoothly at the threshold $MV_{it} = 0$. This assumption is required to guarantee that municipalities on one side of the threshold are a proper counterfactual for municipalities on the other side of the cutoff. We test this assumption in Tables B1 and B2, which confirm that municipal covariates are balanced.

Figure B1: Manipulation test on the margin of victory - Municipalities below 5,000



Notes. Manipulation test on the density of the margin of victory. The manipulation test uses the procedure developed by Cattaneo, Jansson, and Ma (2018). T-statistics: the conventional test statistics is 0.501, while the robust one is 0.679.

Figure B2: Manipulation test on the margin of victory - Municipalities above 5,000



Notes. Manipulation test on the density of the margin of victory. The manipulation test uses the procedure developed by Cattaneo, Jansson, and Ma (2018). T-statistics: the conventional test statistics is 0.582, while the robust one is 1.208.

Table B1: Balance test on municipal covariates
RDD, below 5000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Characteristics municipal population				
Dependent variables	% university degree	% 15-64	% 65+	(log) income per capita	# firms	no-profit ass	area	population density
Graduate Mayor	-0.000	0.003	-0.004	-0.015	0.000	0.000	1.547	-6.117
	(0.002)	(0.005)	(0.007)	(0.036)	(0.002)	(0.000)	(2.742)	(22.834)
Observations	2281	2583	2643	2624	2737	3175	2873	3134
Bandwidth	15.24	17.91	18.53	18.32	19.38	23.93	20.79	23.48
				Panel B: Geographical characteristics municipalities				
Dependent variables	NE	NW	CEN	SOU	Past deficit			
Graduate Mayor	0.023	-0.032	-0.048	0.050	0.002			
	(0.036)	(0.061)	(0.043)	(0.060)	(0.005)			
Effective Observations	2897	2766	2336	2525	1850			
Bandwidth	21	19.62	15.63	17.34	17.44			

Notes. RDD estimates of the impact of graduate mayors on municipal covariates. Municipalities below 5,000 inhabitants. Electoral years between 2001 and 2012. Treatment variable: Graduate is a dummy variable equal to 1 when the mayor has a university degree, 0 otherwise. Estimation by RDD using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Municipal dependent variables in Panel A (measured in 2001): 1) share of population with a university degree; 2) share of active population (i.e. population between 15 and 64 years old); 3) share of elderly (i.e. population above 65 years old); 4) log of income per capita; 5) number of firms per capita; 6) number of non-profit associations per capita; 7) area of municipality in square km; 8) population density. In Panel B, the dependent variables are geographical dummy variables for different areas of Italy (i.e. North-West, North-East, Centre, South) and the deficit as a fraction of total revenues from the previous term. Robust standard errors clustered at the local labor area level are in parentheses.

Table B2: Balance test on municipal covariates
RDD, above 5000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Characteristics municipal population							
Dependent variables	% university degree	% 15-64	% 65+	(log) income per capita	# firms	no-profit ass	area	population density
Graduate Mayor	-0.003	-0.001	0.002	-0.009	-0.002	0.000	3.635	-119.001
	(0.003)	(0.005)	(0.006)	(0.058)	(0.004)	(0.000)	(6.463)	(105.907)
Observations	1098	1083	1013	1041	763	1024	1145	1072
Bandwidth	20.95	20.35	18.18	18.87	12.68	18.45	22.24	20.01
	Panel B: Geographical characteristics municipalities							
Dependent variables	NE	NW	CEN	SOU	Past deficit			
Graduate Mayor	0.105	-0.134	0.012	0.001	-0.000			
	(0.077)	(0.098)	(0.059)	(0.074)	(0.004)			
Observations	871	902	1122	1185	683			
Bandwidth	14.95	15.61	21.44	23.60	15.30			

Notes. RDD estimates of the impact of graduate mayors on municipal covariates. Municipalities between 5000 and 15,000 inhabitants. Electoral years between 2001 and 2012. Treatment variable: Graduate is a dummy variable equal to 1 when the mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Municipal dependent variables in Panel A (measured in 2001): 1) share of population with a university degree; 2) share of active population (i.e. population between 15 and 64 years old); 3) share of seniors (i.e. population above 65 years old); 4) log of income per capita; 5) number of firms per capita; 6) number of non-profit associations per capita; 7) area of municipality in square km; 8) population density. In Panel B, the dependent variables are geographical dummy variables for different areas of Italy (i.e. North-West, North-East, Centre, South) and the deficit as a fraction of total revenues from the previous term. Robust standard errors clustered at the local labor area level are in parentheses.

B.2 Additional Tables described in section IV.C.IV (Results on education level and policy choice)

To address potential endogeneity in municipal income growth during a mayor's term, and consequently in the dependent variable shown in Table 6 in section IV.C.IV, we generate an alternative version of this variable. This is done by predicting income growth through regression analysis on pre-determined municipal characteristics, along with regional and year fixed effects. As indicated in Table B3, employing this alternative variable yields similar results.

Table B3: Graduate mayors and matching the state of the economy (with predicted income growth)

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Year of election FE	No	No	Yes
Region FE	No	No	Yes
Mayoral covariates	No	Yes	Yes
Panel A: municipalities below 5000			
Graduate Mayor	0.043	0.050	0.047
	(0.025)	(0.025)	(0.024)
Observations	2790	2756	2642
Bandwidth	19.85	19.50	18.50
Panel B: municipalities above 5000			
Graduate Mayor	-0.041	-0.040	-0.053
	(0.047)	(0.046)	(0.044)
Observations	910	911	785
Bandwidth	15.19	15.20	12.59
P-Value difference Panel A vs. B	0.109	0.086	0.046

Notes. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Dependent variable: probability of matching the state of the economy over the electoral mandate. In this Table, we use a predicted version of income growth. Treatment variable: Graduate is equal to 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Mayoral covariates in columns 2-3: 1) age of the mayor; 2) political experience: years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high skills occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor. Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

The different results in Table 6 for municipalities below and above the

5,000-inhabitant threshold may be due to the different wages paid to the mayors. To rule out this possibility, we expand the initial dataset, including the 2013-2015 period, and repeat the RDD exercise using only those years, during which fiscal rules applied equally across the threshold. Conversely, during these years, the wage increase across the threshold was in place. Table B4 shows that the differences in matching the state of the economy disappear when fiscal rules apply in the same way across the threshold, as none of the estimated coefficients in the Table is statistically different from zero.

Table B4: Graduate mayors and matching the state of the economy
Years 2013-2015

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Year of election FE	No	No	Yes
Region FE	No	No	Yes
Mayoral covariates	No	Yes	Yes
Panel A: municipalities below 5000			
Graduate Mayor	-0.059	-0.053	-0.060
	(0.048)	(0.047)	(0.046)
Observations	1129	1146	1137
Bandwidth	15.52	15.85	15.64
Panel B: municipalities above 5000			
Graduate Mayor	-0.089	-0.061	-0.055
	(0.073)	(0.072)	(0.074)
Observations	476	479	426
Bandwidth	15.21	15.54	13.33
P-Value difference Panel A vs. B	0.736	0.922	0.957

Notes. Municipalities below 15,000 inhabitants. Years 2013-2015. Dependent variable = 1 in the event of above-median deficit coupled with below-median income growth or below-median deficit with above-median income growth. Treatment variable: Graduate is 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Mayoral covariates in columns 2-3: 1) age of the mayor; 2) political experience = years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high-skill occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor.Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

In Table B5, we present additional evidence that more educated mayors are likelier to foster successful municipal administrations. This is based on outcomes that include investment expenditures, measures of fiscal sustainability, and the amount of public services provided. Data on investment expenditures and measures of fiscal sustainability comes from the municipal budget outcomes from the Aida PA database, an online archive managed by the Bureau Van Dijk. The data contains information on the fiscal items of the budgets of all Italian municipalities, covering the years 2000-2012. Data on investment expenditures and measures of fiscal sustainability are derived from municipal budget outcomes, as recorded in the Aida PA database. This online archive, managed by the Bureau Van Dijk, contains information on the fiscal items of budgets for all Italian municipalities, spanning the years 2000-2012.

To measure the amount of public services provided, we use an indicator developed through data collected by the Italian Ministry of Economics and Finance (Opencivitas, 2015). Lockwood et al. (2021) provide an extensive description. The indicator measures the difference between the amount of services provided by one municipality and the standard level of services that should be provided, which, accordingly to the methodology developed by the Italian Ministry of Finance trough the company Sose, corresponds to the average level of services provided by municipalities in the same population bracket. Using this continuous indicator, we build a dummy variable equal to one for municipalities providing a level of public services equal to or greater than the standard level of services.

The results in Panel A of Table B5 show that in municipalities without fiscal rules, graduate mayors are more likely to increase investment expenditures and to provide more public services compared to non-graduate mayors. In addition, we do not find differences in fiscal sustainability measures (i.e., deficit and debt repayment, see Vannutelli, 2022, for more detail) and current expenditures between graduate and non-graduate mayors. This evidence suggests that graduate mayors can produce better outcomes without worsening the sustainability of the municipal administration. Conversely, as shown in Panel B of Table B5, these differences disappear in municipalities with fiscal rules.

Table B5: Performance of graduate mayors

	(1)	(2)	(3)	(4)	(5)
Dep. Variables	Current	Capital	Services	Deficit	Loan
Dep. Variables	Expenditures	Expenditures	Provided		Repayment
Control Function	Linear	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT	CCT
Year of election FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Mayoral covariates	Yes	Yes	Yes	Yes	Yes
Panel A: municipalities below 5000					
Graduate Mayor	8.072	150.462	0.131	-2.241	-4.777
	(36.099)	(79.769)	(0.057)	(5.385)	(11.052)
Observations	2191	2311	1031	1814	3173
Bandwidth	14.47	15.43	24.30	11.50	23.85
Panel B: municipalities above 5000					
Graduate Mayor	-21.218	3.734	0.041	2.689	19.024
	(32.985)	(30.430)	(0.095)	(4.716)	(16.325)
Observations	1079	857	257	949	843
Bandwidth	18.90	13.97	12.43	16.04	13.71
P-Value difference Panel A vs. B	0.549	0.086	0.416	0.491	0.227

Notes. RDD estimates. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Dependent variables: 1) current expenditures = municipal current expenditures; 2) capital expenditures = municipal capital expenditures; 3) services provided = dummy variable equal to 1 if municipality provided in 2010 an average level of public services equal or above the standard level; 4) deficit = total revenues - total expenditures; 5) loan repayment = loan repayment expenditures. Treatment variable: Graduate is equal to 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Mayoral covariates in all columns: 1) age of the mayor; 2) political experience: years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high skills occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor. Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

C Appendix: Alternative stories

We offer more detailed insights into how we address the alternative explanations outlined in section IV.C.V. First, the application of fiscal rules may require the selection of more politically experienced politicians, who may be less educated. To rule out the latter explanation, we run the *Diff-in-Disc* model on other personal characteristics of local politicians, such as past professional background, age, gender, and past political experience. It is important to highlight that, due to data limitations, it was only possible to reconstruct the past political experience for elected mayors, and not for mayoral candidates. We report the results of this exercise in Table C1. For characteristics potentially correlated with education, the estimated coefficient goes in the expected direction (i.e., a decline in the share of politicians from skilled occupations). On the other hand, gender and years of political experience do not seem to be affected by fiscal rules. The lack of an effect for political experience rules out the possibility that the application of fiscal rules may require the selection of more politically experienced politicians.

In addition, in Table C2, we check whether fiscal rules negatively affected municipal councilors' education level. Specifically, as described in section I., our expectation about the effect of fiscal rules was that these should affect politicians in powerful positions, like mayors, rather than politicians in less prominent positions, like municipal councilors. In line with this expectation, Table C2 reports coefficients that, even though negative, are small and not statistically significant.

Third, we show that different non-political outside options for individuals with different education levels are unlikely to explain our results. In principle, fiscal rules may affect the value of public office for individuals with different

Table C1: The effect of fiscal rules on other characteristics

	(1)	(2)	(3)	(4)
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	No	No	No	No
Region FE	No	No	No	No
Dependent Variables	High skill	Age	Female	Pol Experience
Panel A: mayoral candidates				
(Post)*(> 5000)	-0.092	1.090	0.001	
	(0.043)	(0.810)	(0.024)	
Observations	2,944	4,549	3,637	
Bandwidth	952.1	1482	1180	
Mean outcome	0.286	47.95	0.112	
Panel B: mayors				
(Post)*(> 5000)	-0.089	1.277	0.011	-0.592
	(0.062)	(1.445)	(0.034)	(0.750)
Observations	3,510	3,554	3,596	4,156
Bandwidth	1158	1168	1172	1339
Mean outcome	0.309	47.89	0.087	8.182

Notes. Diff-in-disc estimates of the impact of fiscal rules on politicians' characteristics. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variables are 1) high skill: politicians from high-skill occupations; 2) Age: age of the politicians; 3) Female = 1 for female politicians; 4) Pol Experiences = years of political experience at any level of politics (for mayors only). The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

levels of education in the same way. This homogeneous effect could then affect the entry into politics of individuals with different levels of education heterogeneously, given their different outside options. If higher-educated individuals have a better outside option in the labor market compared to less educated ones, the overall effect could be a reduction in the quality of candidates. Table C3 appears to rule out this alternative story. Specifically, in Table C3, we use data on the municipal shares of employed individuals divided by income

Table C2: The effect of fiscal rules on the education of municipal councilors

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	No	Yes	Yes
Region FE	No	No	Yes
(Post)*(> 5000)	-0.026	-0.030	-0.032
	(0.021)	(0.020)	(0.020)
Observations	3,742	3,742	3,742
Bandwidth	1221	1221	1221
Mean outcome	0.263	0.263	0.263

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education level of municipal councilors. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting after 2001. The outcome variable is the share of municipal councilors with a university degree. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

brackets to measure how concentrated opportunities in the labor market are. To do so, we calculate a Herfindahl index of these income brackets' share to measure whether employed individuals are concentrated in one or more specific income brackets. Higher values of this index suggest a greater concentration in one specific bracket and, thus, more homogeneous opportunities for individuals in that municipality, independently of the level of education. The results in Table C3 indicate that the findings are similar across municipalities with low vs. high values of the Herfindahl index, and, if anything, municipalities with a higher Herfindahl index (i.e., where outside options are homogeneous) present stronger results. The fact that the results are stronger in municipalities where outside options are homogeneous suggests that it is unlikely that different options outside of politics across individuals with different levels of education explain our results.

Table C3: Effect of fiscal rules and outside option in the private sector

	(1)	(2)	(3)	(4)
Dependent Variables	Share mayoral candidates with university degree		= 1 for Mayors with university degree	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Sample	Herfindal index > median	Herfindal index < median	Herfindal index > median	Herfindal index < median
(Post)*(> 5000)	-0.125	-0.080	-0.131	-0.083
	(0.044)	(0.066)	(0.069)	(0.103)
Observations	2,304	1,384	2,713	1,526
Bandwidth (h)	1245	1162	1510	1263
Mean outcome	0.406	0.571	0.422	0.574

Notes. Diff-in-disc estimates. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in columns 1-2, and a dummy variable equal to 1 for mayors with a university degree in columns 3-4. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Election year and region fixed effects added in all columns. Robust standard errors clustered at the local labor area level are in parentheses.

Fourth, fiscal rules may affect politicians' political orientation, which in turn is correlated with their level of education. As an example, fiscal rules may make political office less attractive for left wing perspective candidates, and this could be positively correlated with income and education (Gethin et al., 2022). Table C4 excludes any effect of fiscal rules on politicians' political orientation.

Fifth, fiscal rules may change the desirability of electing a highly-educated mayor, from the voters' perspective. In particular, they may make competence less important, hence reducing the advantage of highly educated politicians. If true, this effect would also be a potential channel for our result, as highly educated politicians would be discouraged to run, with fiscal rules, because they anticipate the reduction in their electoral advantage. However, it does

Table C4: The effect of fiscal rules on ideology

	(1)	(2)	(3)	(4)
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	No	No	No	No
Region FE	No	No	No	No
Dependent Variables	Left	Right	Center	Civic List
Panel A: mayoral candidates				
(Post)*(> 5000)	0.026	-0.008	0.005	-0.047
	(0.036)	(0.034)	(0.012)	(0.049)
Observations	3,537	4,067	4,778	3,287
Bandwdith	1151	1311	1549	1062
Mean outcome	0.201	0.214	0.021	0.566
Panel B: mayors				
(Post)*(> 5000)	0.033	-0.037	0.010	-0.099
	(0.058)	(0.052)	(0.012)	(0.070)
Observations	4,023	3,841	5,194	3,261
Bandwdith	1305	1245	1680	1060
Mean outcome	0.256	0.175	0.012	0.550

Notes. Diff-in-disc estimates of the impact of fiscal rules on the ideology of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variables are: 1) Left = share of center-left candidates in Panel A, =1 for center-left mayors in Panel B; 2) Right = share of center-right candidates in Panel A, =1 for center-right mayors in Panel B; 3) Center = share of center candidates in Panel A, =1 for center mayors in Panel B; 4) Civic lists = share of independent candidates in Panel A, =1 for independent mayors in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

not seem to be the case in our data. More in detail, in Table C5, we use data at the mayoral candidate level and OLS to show that graduate mayoral candidates have better electoral performances than non-graduate ones independently of whether fiscal rules apply or not, in races where at least one highly educated candidate is running. Specifically, graduate candidates receive more votes, reach a better final ranking position, and are more likely to be elected. The results go in the same direction irrespective of whether we consider municipalities and electoral years with fiscal rules or without them.

In addition, in Table C6 we look at the effect of fiscal rules on the probability of having a mayor with a university degree, splitting the sample between municipalities with a pre-treatment share of high education candidates above and below the median. The effect seems to be stronger in the latter group, although the coefficients are not statistically significant. This suggests that voters may partially "correct" for the reduced number of highly educated candidates, by voting for them when available. However, this compensation appears less feasible when the reduction induced by fiscal rules implies that no highly educated candidates are running. Those pieces of evidence should be seen as suggestive, rather than causal, as fiscal rules may change not only the number of high-education candidates, but also their type (in dimensions other than education), and this may be endogenous as well. However, the fact that voters do not change their behaviour seems to suggest that any endogenous selection process on characteristics different from education is not too relevant. One possible explanation for the fact that voters do not seem to change their behaviour with fiscal rules is that fiscal policies are just one of the several tasks a mayor is supposed to do. Hence, voters may think that human capital has a positive impact on other tasks as well, hence keeping (roughly) the same preferences even when fiscal rules constrain fiscal policies.

Sixth, educated mayors may be more corrupt than non-graduate ones. Daniele and Giommoni (2020) show that the introduction of fiscal rules should make it more challenging to extract rents, reducing the office value for individuals attracted by them. If graduates are more corrupt than non-graduates, the introduction of fiscal rules may make them less interested in entering politics. However, this does not seem to be the case. Using the Mafia index built by Calderoni (2011), which quantifies the presence of Mafia-style criminal organizations in Italian provinces, we run model (11) splitting the sample between municipalities in provinces below vs. above the median of mafia presence. As shown in table C7, the negative effect of fiscal rules on the education of mayoral candidates is driven by municipalities in provinces with low mafia presence. These are the municipalities where corruption is less of an issue. Furthermore, as we can see from Table C8, graduate mayors do not appear to be more corrupt than non-graduate ones. More in detail, to measure corruption, we use the web archive of one of the leading Italian newspapers (La Repubblica) to find episodes of corruption linked to the mayors in the analysis. Using an algorithm based on the mayor's first and last names, the name of the city, the years of the legislature, and a series of keywords related to episodes of corruption, we create a database of newspaper articles reporting episodes of corruption linked to the mayors in the dataset. We use this database to create a dummy variable equal to 1 for mayors found to be corrupt, and 0 otherwise. The coefficients reported in Table C8 are estimated using this dummy variable as the dependent variable.

Finally, as described and tested in section IV.C.I and Table A6, we do not find interactive effects between the 1999 introduction of fiscal rules and the differential wage paid across the 5000 inhabitants threshold. To further check that this is the case, in Table C9, we replicate the main analysis of Tables 4 and 5 by keeping only the electoral years from 1999 (i.e., excluding prior elections in which fiscal rules were not implemented in any municipality) and the sub-sample of municipalities that effectively held an election in either the electoral years 1999 or 2000 (i.e., election years in which fiscal rules applied uniformly across the 5000 inhabitants threshold). The idea of this exercise is to repeat the analysis by keeping a pre-treatment period in which the application of fiscal rules is constant over time and across the threshold. As we can see from Panel A of Table C9, the results for the mayoral candidates (i.e., the main focus of our theoretical and empirical analysis) are essentially unchanged, even though less precisely estimated due to the lower number of observations. The results for mayors in Panel B are somehow weaker (i.e., smaller coefficients and not statistically significant), but they are qualitatively similar (i.e., negative and economically significant coefficients in the entire and low rigidity samples, small and positive coefficients in the high rigidity sample).

Table C5: Candidate level regressions: graduate vs. non-graduate candidates

	(1)	(2)	(3)
Dependent Variables	Vote Shares	Ranking Position	=1 if elected Mayor
Panel A: all elections			
Graduate	5.951	-0.222	0.086
	(0.239)	(0.011)	(0.007)
Observations	41,086	41,185	41,185
Panel B: fiscal rules applied			
Graduate	6.016	-0.242	0.088
	(0.384)	(0.020)	(0.011)
Observations	14,080	14,092	14,092
Panel C: fiscal rules did not applied			
Graduate	5.890	-0.210	0.085
	(0.280)	(0.013)	(0.008)
Observations	27,006	27,093	27,093

Notes. OLS estimates. Municipalities below 15,000 inhabitants. Electoral terms between 1993 and 2012. Only electoral races with at least one graduate candidate. Dependent variables: 1) vote shares = vote share taken by mayoral candidate; 2) ranking position = position of the candidate in the final ranking of mayoral candidates; 3) =1 if elected mayor = 1 if candidate elected mayor. Independent variable reported in the Table is = 1 for mayoral candidates with a university degree, 0 otherwise. Election year and region fixed effects included in all columns. Mayoral candidate covariates included in all columns: 1) high skills job = 1 if candidate worked in a high-skill occupation in the past; 2) female = 1 if candidate is a woman; 3) age = age of the mayoral candidate; 4) independent = 1 if candidate is not affiliated to a national political party; 5) unemployed = 1 if candidate is unemployed. Municipal covariates in all columns (measured in 2001, except for numbers 5 and 6, which are measured in 2005): 1) share of population with a university degree; 2) share of active population (i.e. population between 15 and 64 years old); 3) share of seniors (i.e. population above 65 years old); 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of non-profit associations per capita; 7) area of municipality in square km; 8) population density. Robust standard errors clustered at the municipality level are in parentheses.

Table C6: Effect of fiscal rules and pre-treatment share of graduate candidates

	(1)	(2)	(3)	(4)
Dependent Variable:	= 1 mayors with university degree			
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	No	No	Yes	Yes
Region FE	No	No	Yes	Yes
Sample	Low share graduate	High share graduate	Low share graduate	High share graduate
(Post)*(> 5000)	-0.126	-0.057	-0.115	-0.077
	(0.101)	(0.085)	(0.102)	(0.085)
Observations	1,119	2,323	1,119	2,323
Bandwidth	1075	1143	1075	1143
Mean outcome	0.224	0.634	0.224	0.634

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of mayoral candidates. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. High and low share of graduate candidates before treatment is measured using elections between 1993 and 2000. The sample split is at the median. Variables: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in all columns. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Election year and region fixed effects added in columns 3 and 4. Robust standard errors clustered at the local labor area level are in parentheses.

Table C7: The role of criminal organizations

	(1)	(2)	(3)	(4)
Dependent variable	*Mayoral candidates with university degree*		*Mayors with university degree*	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Sample	*Mafia index<median*	*Mafia index>median*	*Mafia index<median*	*Mafia index>median*
(Post)*(> 5000)	-0.141	0.001	-0.274	0.051
	(0.062)	(0.048)	(0.096)	(0.086)
Observations	1,476	2,468	1,722	1,994
Bandwidth	927.8	1613	1078	1304
Mean outcome	0.409	0.513	0.432	0.545

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Sub-samples: Mafia index<median if municipality located in a province with a low presence of Mafia-style criminal organizations; Mafia index>median if municipality located in a province with a high presence of Mafia-style criminal organizations. The mafia index comes from Calderoni (2011). Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in column 1-2 and is equal to 1 for mayors with a university degree in column 3-4. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Table C8: The effect of graduate mayors on corruption

		(1)	(2)	(3)	(4)
Control Function		Linear	Linear	Linear	Linear
Bandwidth		CCT	CCT	CCT	CCT
Year of election FE		No	Yes	No	Yes
Region FE		No	Yes	No	Yes
Covariates		No	Yes	No	Yes
Municipalities		Below 5000		Above 5000	
Dependent variable = 1 if mayor corrupt					
Graduate Mayor		-0.008	0.006	-0.008	-0.027
		(0.015)	(0.014)	(0.041)	(0.038)
Effective Observations		2654	2319	1015	907
Bandwidth		18.60	15.49	17.52	15.12
	Descriptive statistics dummy variable for corruption				
	Mean	St. deviation	Min	Max	Observations
	0.098	0.206	0	1	6694

Notes. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Treatment variable: Graduate is a dummy variable =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Mayoral covariates included in columns 2 and 4: 1) female = 1 if mayor is a woman; 2) age = age of the mayor at the beginning of the term; 3) political experience = years of past political experience of the mayor at any level of politics; 4) left = 1 for center-left mayor; 5) high skills job = 1 if mayor worked in a high-skill occupation in the past. Robust standard errors clustered at the local labor area level are in parentheses.

Table C9: The effect of fiscal rules on the education of politicians
Election years from 1999

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Sample	*All sample*	*Rigidity<median*	*Rigidity>median*
Panel A: mayoral candidates with university degree			
(Post)*(> 5000)	-0.084	-0.182	0.032
	(0.067)	(0.072)	(0.111)
Observations	1,675	1,167	686
Bandwidth	1114	1279	1140
Mean outcome	0.445	0.411	0.494
Panel B: mayors with university degree			
(Post)*(> 5000)	-0.057	-0.084	0.025
	(0.083)	(0.093)	(0.122)
Observations	2,171	1,267	1,196
Bandwidth	1425	1386	1862
Mean outcome	0.449	0.434	0.468

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education level of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1999 and 2012 and only municipalities that voted in election year 1999 or election year 2000. Variables in the table: 1) $(> 5000) = 1$ for municipalities with more than 5000 inhabitants; 2) $(\text{Post}) = 1$ for electoral terms starting after 2001. The outcome variable is the share of mayoral candidates with a university degree in Panel A, and a dummy variable equal to 1 for mayors with a university degree in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

D Additional details on empirical models

D.1 Difference-in-discontinuity model

We estimate the Difference-in-Discontinuity (*Diff-in-Disc*) model described in equation (11) with a local linear regression (Gelman and Imbens, 2018), using the subsample of observations that lie within the interval $R_{it} \in [-h, +h]$ around the threshold. The optimal bandwidth h is calculated using the MSE-optimal bandwidth following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). More in detail, to leverage the panel structure of our dataset, which includes multiple electoral years and time observations for each municipality, we follow the approach of Grembi et al. (2016).

Specifically, we estimate model (11) using the statistical software Stata and the command "regress", assigning equal weight to all observations within the optimal bandwidth h. Accordingly, the MSE-optimal bandwidth h is calculated in Stata using the "rdrobust" command (Calonico et al., 2017), with the option set for a rectangular kernel.

D.2 Regression discontinuity design model

We estimate model (12) using local linear regression (Gelman and Imbens, 2018) on a subsample of municipalities within the interval $MV_{it} \in [-h, +h]$. The optimal bandwidth h is determined based on the MSE-optimal bandwidth criteria from Calonico, Cattaneo, and Titiunik (2014), and Calonico, Cattaneo, and Farrell (2018). This estimation is conducted in Stata with the "rdrobust" command (Calonico et al., 2017). In line with the guidance of Calonico, Cattaneo, and Titiunik (2014), and Cattaneo, Idrobo, and Titiunik (2020), we employ robust inference methods and we weight observations by their proximity to the cutoff using a triangular kernel. The "rdrobust" command provides RDD estimates with a conventional variance estimator (Conventional), bias-corrected RDD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RDD estimates with a robust variance estimator (Robust). For simplicity, in all tables that present estimates from model (12), we report RDD estimates using the conventional variance estimator (Conventional). We have confirmed that the results and evidence from both bias-corrected RDD estimates with a conventional variance estimator (Bias-corrected) and those with a robust variance estimator (Robust) are essentially identical.

E Appendix to the theoretical framework

E.1 Formal analysis of the model

To ease the notation, we define the expected payoffs from being in office as follows:

$$h := E + k((1-\tau)\phi^H + \tau) \tag{E.1}$$

$$l := E + k((1-\tau)\phi^L + \tau) \tag{E.2}$$

$$f := E + k(1-\tau)(1-p) \tag{E.3}$$

E.1.1 No fiscal rules

When there are no fiscal rules, any elected politician is free to choose the policy once in office. As a consequence, at the policy stage biased politicians choose $x = 1$, unbiased politicians choose $x = s$ and hence they pick the correct policy with probability ϕ^Γ.

Lemma E1. *Without fiscal rules, there is a PBNE whose policy choice is as follows*

- *Biased politicians always choose* $x = 1$*;*
- *Unbiased politicians choose* $x = s$*.*

Proof of Lemma E1. Once in office, politicians learn their bias and there is no trade off with respect to their favourite policy. Hence, biased politicians choose $x = 1$ irrespective of the state. Unbiased politicians always choose $x = s$, because $\phi^L > max[p, 1-p]$, hence the signal realization always indicates the most likely state of the world.

To see this, note that if $\phi > p$

$$Pr(\theta = 0|s = 0) = \frac{\phi(1-p)}{\phi(1-p) + (1-\phi)p} > \frac{(1-\phi)p}{(1-\phi)p + \phi(1-p)} = Pr(\theta = 1|s = 0)$$

by Bayes' rule, and if $\phi > 1 - p$

$$Pr(\theta = 1|s = 1) = \frac{\phi p}{\phi p + (1-\phi)(1-p)} > \frac{(1-\phi)(1-p)}{\phi p + (1-\phi)(1-p)} = Pr(\theta = 0|s = 1)$$

Ex ante (i.e. before observing the signal realization), a politician with signal precision ϕ expects to choose the policy that matches the state, if she follows the signal realization (i.e. if she chooses $x = s$), with probability ϕ. To see this, note that, from an ex ante perspective,

$$Pr(s = \theta) = (1-p)Pr(s = 0|\theta = 0) + pPr(s = 1|\theta = 1) = \phi$$

■

The voter anticipates the equilibrium choices described above. Since higher educated unbiased politicians behave in a better way, in expectation, V prefers to elect the candidate with $\Gamma = H$ when the election is contested.

Lemma E2. *If there are two candidates of different education level, $\gamma_H^i = 1$.*

Proof of Lemma E2. At the voting stage, V anticipates the policy choices outlined in Lemma E1. Suppose two candidates of different education level run: from V's point of view, the expected utility of choosing the H candidate is $\mathbb{E}u^V(\Gamma = H) = \tau p + (1-\tau)\phi^H$, because the biased politician matches the state with probability p and the unbiased one with probability ϕ^Γ. It is easy to see that $\mathbb{E}u^V(\Gamma = H) > \mathbb{E}u^V(\Gamma = L) = \tau p + (1-\tau)\phi^L$ because $\phi^H > \phi^L$. ■

Combining these results, we can derive the relevant thresholds in w.

Lemma E3. *Without fiscal rules, there exists a symmetric PBNE where the entry threshold of politicians is defined by*

$$\bar{w}_H = \frac{4W^H \left(E + ((1-\tau)\phi^H + \tau)k\right)}{4W^H + (E + ((1-\tau)\phi^H + \tau)k)}$$

$$\bar{w}_L = \frac{4W^H - \left(E + ((1-\tau)\phi^H + \tau)k\right)}{4W^H + (E + ((1-\tau)\phi^H + \tau)k)} \frac{4W^L \left(E + ((1-\tau)\phi^L + \tau)k\right)}{4W^L + (E + ((1-\tau)\phi^L + \tau)k)}$$

Proof of Lemma E3. Start from an H politician. She compares (1) and (2), choosing to enter when $w^i \leq \gamma_H^i(E + k\mathbb{E}_{b,\theta,s}u_H^P)$. Given Lemma E1, it is clear that $\mathbb{E}_{b,\theta,s}u_H^P = ((1-\tau)\phi^H + \tau)$. Given Lemma E2, it is clear that $\gamma_H^i = 1 - \frac{1}{2}\frac{1}{2}p_H^c$, where p_H^c is the conjectured probability that an opponent perspective candidate of H type chooses to run. Moving to an L politician, the logic on $\mathbb{E}_{b,\theta,s}u_L^P$ is the same. However, she knows she can win office only if H does not run, hence with probability $\left(1 - \frac{1}{2}\frac{1}{2}p_L^c - \frac{1}{2}p_H^c\right)$, because the L type loses for sure against an H opponent and with probability 0.5 against a L opponent. In a symmetric equilibrium, strategies must be the same for players of the same type and conjectured probabilities of running must be correct, thus $p_L^c = \frac{\bar{w}_L}{W^L}$ and $p_H^c = \frac{\bar{w}_H}{W^H}$. As a consequence, the symmetric equilibrium thresholds are the solution of the following system of equations:

$$\bar{w}_H = \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^H + \tau)k\right)$$
$$\bar{w}_L = \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L} - \frac{1}{2}\frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^L + \tau)k\right)$$

We solve the system starting from $\bar{w}_H$ and using (E.1) and (E.2) to ease the notation.

$$\begin{aligned}
\bar{w}_H &= \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}\right) h \\
\bar{w}_H &= \left(\frac{4W^H - \bar{w}_H}{4W^H}\right) h \\
\bar{w}_H(4W^H + h) &= 4W^H h \\
\bar{w}_H &= \frac{4W^H h}{4W^H + h}
\end{aligned}$$

Substituting in the second equation, we solve for $\bar{w}_L$:

$$\begin{aligned}
\bar{w}_L &= \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L} - \frac{1}{2}\frac{\bar{w}_H}{W^H}\right) l \\
\bar{w}_L &= \left(1 - \frac{\bar{w}_L}{4W^L} - \frac{1}{2W^H}\frac{4W^H h}{4W^H + h}\right) l \\
\bar{w}_L \left(1 + \frac{l}{4W^L}\right) &= \left(1 - \frac{2h}{4W^H + h}\right) l \\
\bar{w}_L \left(\frac{4W^L + l}{4W^L}\right) &= \left(\frac{4W^H - h}{4W^H + h}\right) l \\
\bar{w}_L &= \frac{4W^H - h}{4W^H + h}\frac{4W^L l}{4W^L + l}
\end{aligned}$$

Hence, we find that in our symmetric equilibrium (which is unique conditional on our indifference breaking assumptions)

$$\begin{aligned}
\bar{w}_H &= \frac{4W^H h}{4W^H + h} \\
\bar{w}_L &= \frac{4W^H - h}{4W^H + h}\frac{4W^L l}{4W^L + l}
\end{aligned}$$

■

Lemma E4. *Conditional on the assumptions on the tie-breaking rules, there are no symmetric PBNE leading to strategies different than those described in*

Lemma E3.

Proof of Lemma E4. To prove the statement, note the following:

- Policy choices are strictly dominant strategies for the different types of politicians, once the type is realized, hence they are the sole sequentially rational strategy and they must be the same in every equilibrium;

- Given the anticipated and uniquely defined policy choices, the voting choice of the voter is uniquely defined, meaning that in every equilibrium the voter would have a unilateral profitable deviation with any alternative choice than opting for the H candidate whenever available. When there is only one candidate the voter does not play any role.

- Move now to perspective candidates' entry decision. For every conjectured strategy, every candidate's strategy is a threshold strategy. First, consider the H candidate. For any conjectured strategy of the opponents, given the way we assume the voter breaks indifferences, her expected payoffs from running are uniquely defined as $(1-\frac{1}{2}\frac{1}{2}p^c_H)(E+((1-\tau)\phi^H+\tau)k)$. Furthermore, $(1-\frac{1}{2}\frac{1}{2}p^c_H)(E+((1-\tau)\phi^H+\tau)k)>0$ and $(1-\frac{1}{2}\frac{1}{2}p^c_H)(E+((1-\tau)\phi^H+\tau)k) < W^H$, hence in every equilibrium there must exist a unique type of w^i, strictly between 0 and W^H, such that $w^i = (1-\frac{1}{2}\frac{1}{2}p^c_H)(E+((1-\tau)\phi^H+\tau)k)$. For every type above it, the unique best response is not to run. For every type below it, the unique best response is to run. By assumption, type $w^i = (1-\frac{1}{2}\frac{1}{2}p^c_H)(E+((1-\tau)\phi^H+\tau)k) := \bar{w}_H$ chooses to run.

- Consider now the L perspective candidate. She knows she will win only if H does not run. Furthermore, she wins with probability $\frac{1}{2}$ against a low education opponent. In every equilibrium, given the way we assume the voter breaks indifferences, this happens with probability $1-\frac{1}{2}p^c_H-\frac{1}{2}\frac{1}{2}p^c_L \in (0,1)$, and this conjecture must be correct. Furthermore, in every equilibrium her expected payoff from being in office is uniquely defined as $(E+((1-\tau)\phi^L+\tau)k)$. Furthermore, $\left(1-\frac{1}{2}p^c_H-\frac{1}{2}\frac{1}{2}p^c_L\right)(E+((1-\tau)\phi^L+\tau)k) > 0$ and $\left(1-\frac{1}{2}p^c_H-\frac{1}{2}\frac{1}{2}p^c_L\right)(E+((1-\tau)\phi^L+\tau)k) < W^L$, hence in every equilibrium there must exists a unique type of w^i, strictly between 0 and W^L, such that $w^i = \left(1-\frac{1}{2}p^c_H-\frac{1}{2}\frac{1}{2}p^c_L\right)(E+((1-\tau)\phi^L+\tau)k)$. For every type above it, the unique best response is not to run. For every type below it, the unique best response is to run. By assumption on the tie breaking rule, type $w^i = \left(1-\frac{1}{2}p^c_H-\frac{1}{2}\frac{1}{2}p^c_L\right)(E+((1-\tau)\phi^L+\tau)k) := \bar{w}_L$ chooses to run.

- In every symmetric equilibrium, it must be that conjectures are correct and candidates with the same education level choose the same strategy. Hence, it must be that $p^c_H = \frac{\bar{w}_H}{W^H}$ and $p^c_L = \frac{\bar{w}_L}{W^L}$. As a consequence, $\gamma^i_H = 1-\frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}$ and $\gamma^i_L = 1-\frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L}-\frac{1}{2}\frac{\bar{w}_H}{W^H}$. This leads to the system of equations described in lemma E3, whose solution is unique.

The same logic applies to the equilibrium in case of fiscal rules. ■

E.1.2 Fiscal rules

If fiscal rules are present, all politicians in office are constrained to choose $x=0$. As a consequence,

Lemma E5. *When fiscal rules are in place, equilibrium entry thresholds are as follows:*

$$\bar{w}^{FR}_H = \frac{4W^H\left(E+(1-\tau)(1-p)k\right)}{4W^H+\left(E+(1-\tau)(1-p)k\right)}$$

$$\bar{w}^{FR}_L = \frac{4W^H-\left(E+(1-\tau)(1-p)k\right)}{4W^H+\left(E+(1-\tau)(1-p)k\right)}\frac{4W^L\left(E+(1-\tau)(1-p)k\right)}{4W^L+\left(E+(1-\tau)(1-p)k\right)}$$

Proof of Lemma E5. Given our assumption that, even in case of fiscal rules, the H candidate is chosen over an L candidate, the proof for this Lemma follows the same logic as the proof of Lemma E3. The sole difference is that now $\mathbb{E}_{b,\theta,s}u_H^P = \mathbb{E}_{b,\theta,s}u_L^P = (1-\tau)(1-p)$. The reason is that now both types of politicians, being constrained to play $x = 0$, derive utility only if $\theta = 0$ and they are unbiased. ■

E.2 Proof of the main proposition

Proof of Proposition 1. The proposition implies a comparison between $\hat{\lambda}$ and $\hat{\lambda}^{FR}$, defined using equation (5) and replacing the relevant p_H and p_L. We have:

$$\begin{aligned}
&\hat{\lambda} > \hat{\lambda}^{FR} \\
&\frac{p_L^{FR}}{p_H^{FR}} > \frac{p_L}{p_H} \\
&\frac{\bar{w}_H}{\bar{w}_L} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}}
\end{aligned} \tag{E.4}$$

Substituting the relevant thresholds (7), (8), (9) and (10), and using (E.1), (E.2) and (E.3) to ease the notation, we have that:

$$\frac{\bar{w}_H}{\bar{w}_L} = \frac{4W^H h}{4W^H + h}\frac{4W^H + h}{4W^H - h}\frac{4W^L + l}{4W^L l} > \frac{4W^H f}{4W^H + f}\frac{4W^H + f}{4W^H - f}\frac{4W^L + f}{4W^L f} = \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}} \tag{E.5}$$

$$\frac{(4W^L + l)h}{(4W^H - h)l} > \frac{(4W^L + f)}{(4W^H - f)}$$

Furthermore, the RHS of (E.5) is increasing in f and $f \leq l$ because $\phi^L >$

$max[p, 1-p]$. Hence, the RHS is below $\frac{(4W^L+l)}{(4W^H-l)}$. Note, however, that:

$$\frac{(4W^L+l)h}{(4W^H-h)l} > \frac{(4W^L+l)}{(4W^H-l)}$$
$$\frac{h}{(4W^H-h)l} > \frac{1}{(4W^H-l)}$$
$$(4W^H-l)h > (4W^H-h)l$$
$$h > l$$

that always holds because $\phi^H > \phi^L$. ■

Equation (E.5) is useful to capture the two channels through which fiscal rules act. Each side is composed by two elements whose comparison, individually taken, points toward $\hat{\lambda} > \hat{\lambda}^{FR}$. First, we have that $\frac{h}{l} > \frac{f}{f} = 1$, because $\phi^H > \phi^L$ and fiscal rules shut down the difference in expected payoffs from office between the two types of politicians. In words, the ratio between expected payoffs from being in office for H over L types is higher without fiscal rules, implying that their presence should discourage H types relatively more (note that fiscal rules reduce both h and l). Furthermore, the condition $\frac{h}{l} > \frac{f}{f}$ can be re-written as $\frac{h-f}{h} > \frac{l-f}{l}$. In other words, fiscal rules reduce $\hat{\lambda}$ through the first channel as long as the expected cost they imply for politicians in office, relative to their payoff from office without fiscal rules, is higher for H than for L types. In our model, this is always the case.

Second, we have that $\frac{(4W^L+l)}{(4W^H-h)} > \frac{(4W^L+f)}{(4W^H-f)}$, because $h > f$ and $l > f$. This part is a consequence of the strategic considerations of different types related with the running probability of the opponent. More in detail, fiscal rules unambiguously decrease the equilibrium p_H. However, a reduction in p_H is good news for L types, because they may win with higher chances.

Finally, we show that $\hat{\lambda}$ corresponds to the expected share of H candidates in any given municipality where at least one candidate runs.

Lemma E6. *The municipality-level expected share of H candidates conditional, on having at least one candidate, is* $\hat{\lambda} = \frac{p_H}{p_H + p_L}$.

Proof of Lemma E6. Define the expected share of H candidates conditional on having at least one candidate running as:

$$\hat{S} := \frac{[0.25 * 1 * (1 - (1 - p_H)^2) + 0.5 * 1 * p_H(1 - p_L) + 0.5 * 0.5 * p_H p_L]}{0.25(1 - (1 - p_H)^2) + 0.25(1 - (1 - p_L)^2) + 0.5(1 - (1 - p_H)(1 - p_L))} \tag{E.6}$$

To see that this is the expected share conditional on at least one perspective candidate running, note that at municipality level the share can be 1 with probability $\frac{1}{4}(1-(1-p_H)^2)+\frac{1}{2}p_H(1-p_L)$, i.e. when there are two H perspective candidates and at least one of them run or when there are one H and L candidate and only the H candidate runs. The share is 0.5 with probability $\frac{1}{2}p_H p_L$ (i.e. there are one H and L perspective candidate and both of them run), zero with probability $\frac{1}{4}(1 - (1 - p_L)^2) + \frac{1}{2}p_L(1 - p_H)$ and undefined (define it as $S = \emptyset$) when no perspective candidate runs, i.e. with probability $\frac{1}{4}(1-p_H)^2+\frac{1}{4}(1-p_L)^2+\frac{1}{2}(1-p_H)(1-p_L)$. Then, the expected share conditional on $S \neq \emptyset$ is

$$\begin{aligned}\hat{S} = E(S|S \neq \emptyset) &= 1 * Pr(S = 1|S \neq \emptyset) + 0.5 * Pr(S = 0.5|S \neq \emptyset) \\ &= 1 * \frac{Pr(S = 1 \cap S \neq \emptyset)}{Pr(S \neq \emptyset)} + 0.5 * \frac{Pr(S = 0.5 \cap S \neq \emptyset)}{Pr(S \neq \emptyset)}\end{aligned}$$

Substituting the relevant probabilities, we obtain (E.6), where the denominator is the total probability of having at least one candidate running.

To complete the proof, note that:

$$
\begin{aligned}
\hat{S} &= \frac{0.25(1-(1-p_H)^2)+0.5p_H(1-p_L)+0.5*0.5p_Hp_L}{0.25(1-(1-p_H)^2)+0.25(1-(1-p_L)^2)+0.5(1-(1-p_H)(1-p_L))} \\
&= \frac{1-(1-p_H)^2+2p_H(1-0.5p_L)}{1-(1-p_H)^2+1-(1-p_L)^2+2-2(1-p_H)(1-p_L)} \\
&= \frac{1-(1-2p_H+p_H^2)+2p_H-p_Hp_L}{4-((1-p_H)+(1-p_L))^2} \\
&= \frac{4p_H-p_H^2-p_Hp_L}{4-(2-(p_H+p_L))^2} \\
&= \frac{p_H(4-p_H-p_L)}{4-4-(p_H+p_L)^2+4(p_H+p_L)} \\
&= \frac{p_H(4-p_H-p_L)}{(p_H+p_L)(4-(p_H+p_L))} \\
&= \frac{p_H}{p_H+p_L} := \hat{\lambda}
\end{aligned}
$$

■

E.3 Rigid municipalities

With respect to the baseline model, we add a second group of municipalities, those that are characterized by a high share of rigid expenditures, such as personnel and debt repayment expenditures, which cannot be adjusted in the short run. This implies that they cannot adjust their policy choice quickly. For simplicity, we model this as a constraint to keep the policy constant irrespective of the state of the world. We show that the introduction of fiscal rules is always expected to have a bigger effect on the probability that a candidate is an H type in non-rigid municipalities. Intuitively, the ability to get a better signal about the state of the world does not matter in case of rigidity and in case of fiscal rules. As long as choosing the right policy is valuable for motivating perspective candidates, the constrain imposed by fiscal rules has a stronger discouraging effect on highly educated perspective candidates in previously unconstrained municipalities.

Rigidity as $x = 0$. Assume that rigid municipalities are constrained to the policy $x = 0$ irrespective of the true state of the world, even in the absence of fiscal rules, as they cannot adapt their expenditures quickly when they should respond to negative shocks. Given the above, it is easy to see that, in rigid municipalities, the expected payoff conditional on being in office is the same for every education level, and it is $(E + (1 - \tau)(1 - p)k) = f$, irrespective of whether fiscal rule are in place or not.

Proposition E1. *When rigidity implies $x = 0$, the probability that a candidate is highly educated in rigid municipalities is the same with or without fiscal rules.*

Proof of Proposition E1. Define $\hat{\lambda}^{R0}$ the probability that a candidate is highly educated in those municipalities. Given the exogenous constraint to $x = 0$ irrespective of fiscal rules, we have $\bar{w}_H^{R0} = \bar{w}_H^{FR}$ and $\bar{w}_L^{R0} = \bar{w}_L^{FR}$, hence if we substitute in equation (5) we obtain $\hat{\lambda}^{R0} = \hat{\lambda}^{FR}$. ■

Rigidity as $x = 1$. Assume that those rigid municipalities are constrained to the policy $x = 1$ irrespective of the true state of the world. It is easy to see that, in those rigid municipalities, the expected payoff conditional on being in office is $E + k(\tau + (1 - \tau)p) := r1$. Equilibrium thresholds are the solution of the same system of equations as above, where h and l are both replaced by $r1$. We first show that the comparison between the probability that a candidate is an H type in those municipalities and in municipalities with fiscal rules is in general ambiguous. Second, we show that the probability that a candidate is an H type in non-rigid municipalities is always higher than in rigid municipalities, implying that any negative effect of fiscal rules is stronger in non-rigid municipalities. Define $\hat{\lambda}^{R1}$ the probability that a candidate is an H type in rigid municipalities.

Proposition E2. *When rigidity implies $x = 1$, the probability that a candidate is highly educated in rigid municipalities is higher than the probability that a candidate is highly educated with fiscal rules if $\tau > (1-\tau)(1-2p)$.*

Proof of Proposition E2. The proposition implies a comparison between $\hat{\lambda}^{R1}$ and $\hat{\lambda}^{FR}$. We have:

$$\hat{\lambda}^{R1} > \hat{\lambda}^{FR}$$
$$\frac{\bar{w}_H^{R1}}{\bar{w}_L^{R1}} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}}$$
$$\frac{4W^L + r1}{4W^{H-r1}} > \frac{4W^L + f}{4W^{H-f}}$$
$$4W^H(r1 - f) > 4W^L(f - r1)$$

The inequality is true iff $r1 > f$, i.e. iff $E+k(\tau+(1-\tau)p) > E+k(1-\tau)(1-p)$. This simplifies to $\tau > (1-\tau)(1-2p)$. ■

Proposition E3. *When rigidity implies $x = 1$, the probability that a candidate is highly educated in non-rigid municipalities is always higher than the probability that a candidate is highly educated in rigid municipalities.*

Proof of Proposition E3. The proposition implies a comparison between $\hat{\lambda}$ and $\hat{\lambda}^{R1}$. We have:

$$\hat{\lambda} > \hat{\lambda}^{R1}$$
$$\frac{\bar{w}_H}{\bar{w}_L} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}}$$
$$\frac{h(4W^L + l)}{l(4W^H - h)} > \frac{4W^L + r1}{4W^H - r1}$$

Note that the RHS is increasing in $r1$, and that $r1 < l$. Hence, the LHS is higher than the upper bound of the LHS. To see this, note that

$$\frac{h(4W^L + l)}{l(4W^H - h)} > \frac{4W^L + l}{4W^H - l}$$
$$\frac{h}{l(4W^H - h)} > \frac{1}{4W^H - l}$$
$$h(4W^H - l) > l(4W^H - h)$$
$$h > l$$

■

E.4 Discussion on the theoretical framework

In this appendix we further discuss some of the assumptions and implications of the model.

E.4.1 The necessity of policy-motivated politicians

This section shows that some degree of policy motivation (irrespective of its direction) is necessary for our result.

Corollary E1. *If $k = 0$, the probability that a candidate is highly educated is the same with and without fiscal rules.*

Proof of Corollary E1. Substituting $k = 0$ in the LHS of equation (E.4) and on the relevant equations of Lemma E3 and E5, we obtain

$$h := E$$

$$l := E$$

$$f := E$$

Hence, $\hat{\lambda} = \hat{\lambda}^{FR}$. ■

Intuitively, when $k = 0$, fiscal rules have no effect on the incentives of H politicians: they get E for being in office irrespective of the policy they choose. Hence, their probability of running is the same, and nothing changes for L politicians as well.

On the other hand, the observed effect of fiscal rules holds if politicians are purely policy motivated and if the winning probability does not enter in their decision. In particular:

Corollary E2. *Assume that $E = 0$ and $k > 0$. In this case, $\hat{\lambda} > \hat{\lambda}^{FR}$.*

Proof of Corollary E2. Substituting $E = 0$ in (E.4) and on the relevant equations of Lemma E3 and E5, it is still true that $h := k(\tau + (1-\tau)\phi^H) > l := k(\tau + (1-\tau)\phi^L)$. Therefore, we can follow the same steps as in the proof of proposition 1 and conclude that $\hat{\lambda} > \hat{\lambda}^{FR}$. ■

We also show that the effect of fiscal rules, driven by $\phi^H > \phi^L$, survives even if we assume that candidates keep their salary if they lose, so effectively they do not take into account γ_Γ^i in their decision.

Corollary E3. *Assume that politicians receive w^i, instead of 0, when they run and lose. In this case, $\hat{\lambda} > \hat{\lambda}^{FR}$.*

Proof of Corollary E3. With this assumption, equations (1) and (3) become $\gamma^i\left(E + k\mathbb{E}_{b,\theta,s}u_H^P\right) + (1-\gamma^i)w^i$ and $\gamma^i\left(E + k\mathbb{E}_{b,\theta,s}u_L^P\right) + (1-\gamma^i)w^i$ respectively. This means that $\bar{w}_\Gamma$ does not depend on γ anymore. Therefore, it is straightforward to see that in this case $\bar{w}_H = E + ((1-\tau)\phi^H + \tau)k$, $\bar{w}_L = E + ((1-\tau)\phi^L + \tau)k$ and $\bar{w}_H^{FR} = \bar{w}_L^{FR} = E + (1-\tau)(1-p)$. Replacing in equation (E.4), we obtain

$$\frac{\left(E + ((1-\tau)\phi^H + \tau)k\right)}{(E + ((1-\tau)\phi^L + \tau)k)} > 1$$

Hence the result holds. ∎

E.4.2 Education and bias

Suppose bias is correlated with education, i.e. we have τ_H and τ_L. We show that it is always possible to find a range of values in τ_H, τ_L where the main result of the paper holds. We keep assuming that H politicians are preferred by V.[1]

Proposition E4. *Assume $\tau_H \neq \tau_L$. For every τ_H, it is always possible to find a range of values of τ_L such that $\hat{\lambda} > \hat{\lambda}^{FR}$.*

[1] This translates into the assumption that $(1-\tau_H)\phi^H + \tau_H p > (1-\tau_L)\phi^L + \tau_L p$, i.e. $\tau_L > \tau_H \frac{\phi^H - p}{\phi^L - p} - \frac{\phi^H - \phi^L}{\phi^L - p}$.

Proof of Proposition E4. To ease the notation, we define the following:

$$\begin{aligned}
h_\tau &:= E + k((1-\tau_H)\phi^H + \tau_H) \\
l_\tau &:= E + k((1-\tau_L)\phi^L + \tau_L) \\
f_H &:= E + k(1-\tau_H)(1-p) \\
f_L &:= E + k(1-\tau_L)(1-p)
\end{aligned}$$

Using (E.4), but noticing that we cannot simplify the RHS as before, we have that $\hat{\lambda} > \hat{\lambda}^{FR}$ iff

$$\frac{h_\tau(4W^L + l_\tau)}{l_\tau(4W^H - h_\tau)} > \frac{f_H(4W^L + f_L)}{f_L(4W^H - f_H)}$$
$$h_\tau(4W^L + l_\tau) f_L(4W^H - f_H) > f_H(4W^L + f_L) l_\tau(4W^H - h_\tau)$$

Note that $h_\tau > f_H$ and $l_\tau > f_L$, therefore a sufficient condition for the inequality to hold is

$$h_\tau f_L > f_H l_\tau \tag{E.7}$$

We now show that, for every τ_H and every combination of parameters, there exists a set of values of τ_L where (E.7) holds. First, note that if $\tau_H > \tau_L$ then $h_\tau > l_\tau$ and $f_L > f_H$, therefore the inequality is always satisfied. Consider now the case of $\tau_H < \tau_L$. Noticing that the LHS of (E.7) is increasing in τ_H and the RHS is decreasing in τ_H, we set τ_H to zero and look for a condition on τ_L such that the inequality holds. Higher τ_H are only going to relax this condition. Substituting $\tau_H = 0$ in (E.7) and using the definitions outlined above, we have that the inequality holds iff

$$
\begin{aligned}
&(E + k\phi^H)(E + k(1 - \tau_L)(1 - p)) > (E + k(1 - p))(E + k(\phi^L + (1 - \phi^L)\tau_L))\\
&(E + k(1 - p))k(\phi^H - \phi^L) > \tau_L \left[(E + k(1 - p))k(1 - \phi^L) + (E + k\phi^H)(1 - p)k\right]\\
&\tau_L < \frac{(E + k(1 - p))(\phi^H - \phi^L)}{E(2 - \phi^L - p) + k(1 - p)(1 + \phi^H - \phi^L)} := \bar{\tau}_L
\end{aligned}
$$

Note that $\bar{\tau}_L$ is strictly positive for every combination of parameters, therefore we have a non-empty set of values of τ_L such that the main result of the paper holds. ∎

References cited only in the Online Appendix

Calderoni F. (2011). "Where is the mafia in Italy? Measuring the presence of the mafia across Italian provinces." *Global Crime*.

Calonico, S., Cattaneo, M.D., Farrell, M. H. and Titiunik R. (2017). "rdrobust: Software for Regression Discontinuity Designs." *The Stata Journal* 17(2): 372–404.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). "Manipulation Testing based on Density Discontinuity." *Stata Journal*, 18(1): 234-261.

Cattaneo, M. D., Idrobo, N. and Titiunik R. (2020). "A Practical Introduction to Regression Discontinuity Designs: Foundations." Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge: Cambridge University Press.

Lockwood, B., Porcelli, F., Redoano, M., Bracco, E., Liberini, F., and Sgroi, D. (2021). "The Effects of Social Capital on Government Performance and Turnover: Theory and Evidence from Italian Municipalities." Working paper.

OpenCivitas, (2015). "2010 Comuni – Fabbisogni, caratteristiche e prestazioni generali." Distributed by OpenCivitas, Roma. Retrieved from https://www.opencivitas.it/it/open-data, last accessed: 2022-12-11.