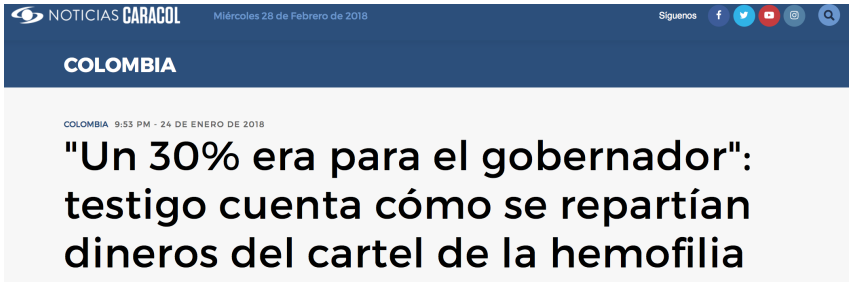
Do ghosts exist?
Corruption in the Colombian education system

Leopoldo Fergusson[†] Arturo Harker[†]
Carlos Molina[‡] Juan Camilo Yamín[†]

[†]Universidad de los Andes [‡]MIT

ASSA 2021 Virtual Annual Meeting
The Political Economy of State Building in Latin America
Jan 3 2020

Ghosts exist

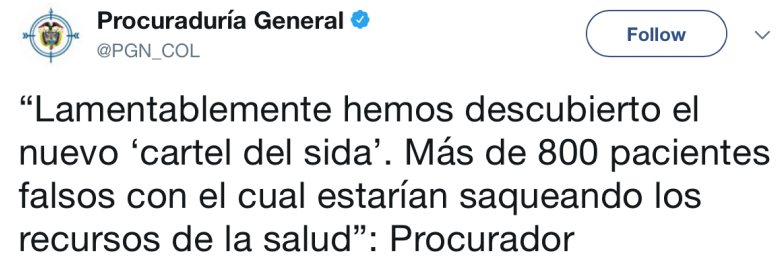

NOTICIAS CARACOL Miércoles 28 de Febrero de 2018 Síguenos

COLOMBIA

COLOMBIA 9:53 PM - 24 DE ENERO DE 2018

"Un 30% era para el gobernador": testigo cuenta cómo se repartían dineros del cartel de la hemofilia

"30% was for the governor," witness describes how the money from the hemophilia cartel was distributed.


Procuraduría General
@PGN_COL
Follow

"Lamentablemente hemos descubierto el nuevo 'cartel del sida'. Más de 800 pacientes falsos con el cual estarían saqueando los recursos de la salud": Procurador

"Unfortunately we have discovered the new 'AIDS cartel'. More than 800 false patients used to steal money from the health system": Attorney General.

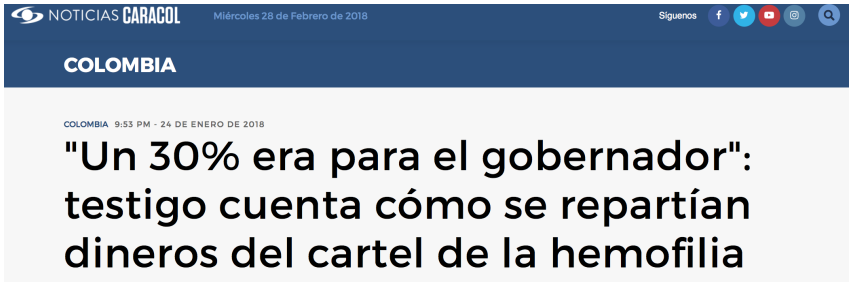

Procuraduría investiga 60 mil casos de "estudiantes fantasma" en todo el país

Judicial 17 Jun 2012 - 5:35 PM

Se habrían desviado 116 mil millones de pesos del sector de la educación.

Attorney General studies 60 thousand cases of "ghost students" in the country

Ghosts exist

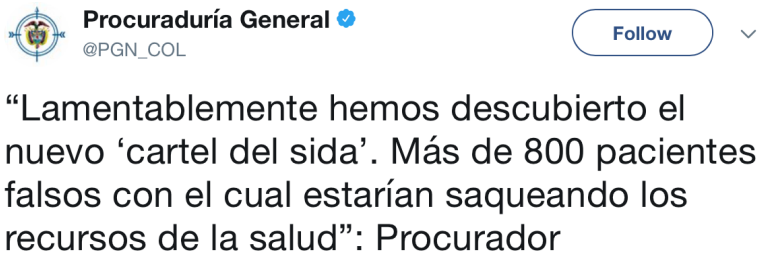

NOTICIAS CARACOL Miércoles 28 de Febrero de 2018 Síguenos

COLOMBIA

COLOMBIA 9:53 PM - 24 DE ENERO DE 2018

"Un 30% era para el gobernador": testigo cuenta cómo se repartían dineros del cartel de la hemofilia

"30% was for the governor," witness describes how the money from the hemophilia cartel was distributed.


Procuraduría General
@PGN_COL
Follow

"Lamentablemente hemos descubierto el nuevo 'cartel del sida'. Más de 800 pacientes falsos con el cual estarían saqueando los recursos de la salud": Procurador

"Unfortunately we have discovered the new 'AIDS cartel'. More than 800 false patients used to steal money from the health system": Attorney General.


Procuraduría investiga 60 mil casos de "estudiantes fantasma" en todo el país

Judicial 17 Jun 2012 - 5:35 PM

Se habrían desviado 116 mil millones de pesos del sector de la educación.

Attorney General studies 60 thousand cases of "ghost students" in the country

$\Uparrow$  $\Rightarrow \Uparrow$ \$ transfers

Corruption for public gain versus for public benefit

- Why $\Uparrow$  ?
  - The most prevalent idea is probably: $\Uparrow$ Capture.
  - But it might also be: $\Uparrow$ Service delivery.
  - Or both: Service delivery $\Leftrightarrow$ $\Uparrow$ Capture.

Corruption for public gain versus for public benefit

- Why $\Uparrow$  ?
  - The most prevalent idea is probably: $\Uparrow$ Capture.
  - But it might also be: $\Uparrow$ Service delivery.
  - Or both: Service delivery $\Leftrightarrow$ $\Uparrow$ Capture.
- This project ask three main questions:
  1. What role does political alignment play in this?
     A: Aligned politicians produce more ghosts.
  2. But then, what do politicians coordinate for?
     A: Mostly capture.
  3. But then, how do they access and hold on to power?
     A: More cheating (electoral fraud).

Related literature and contribution I

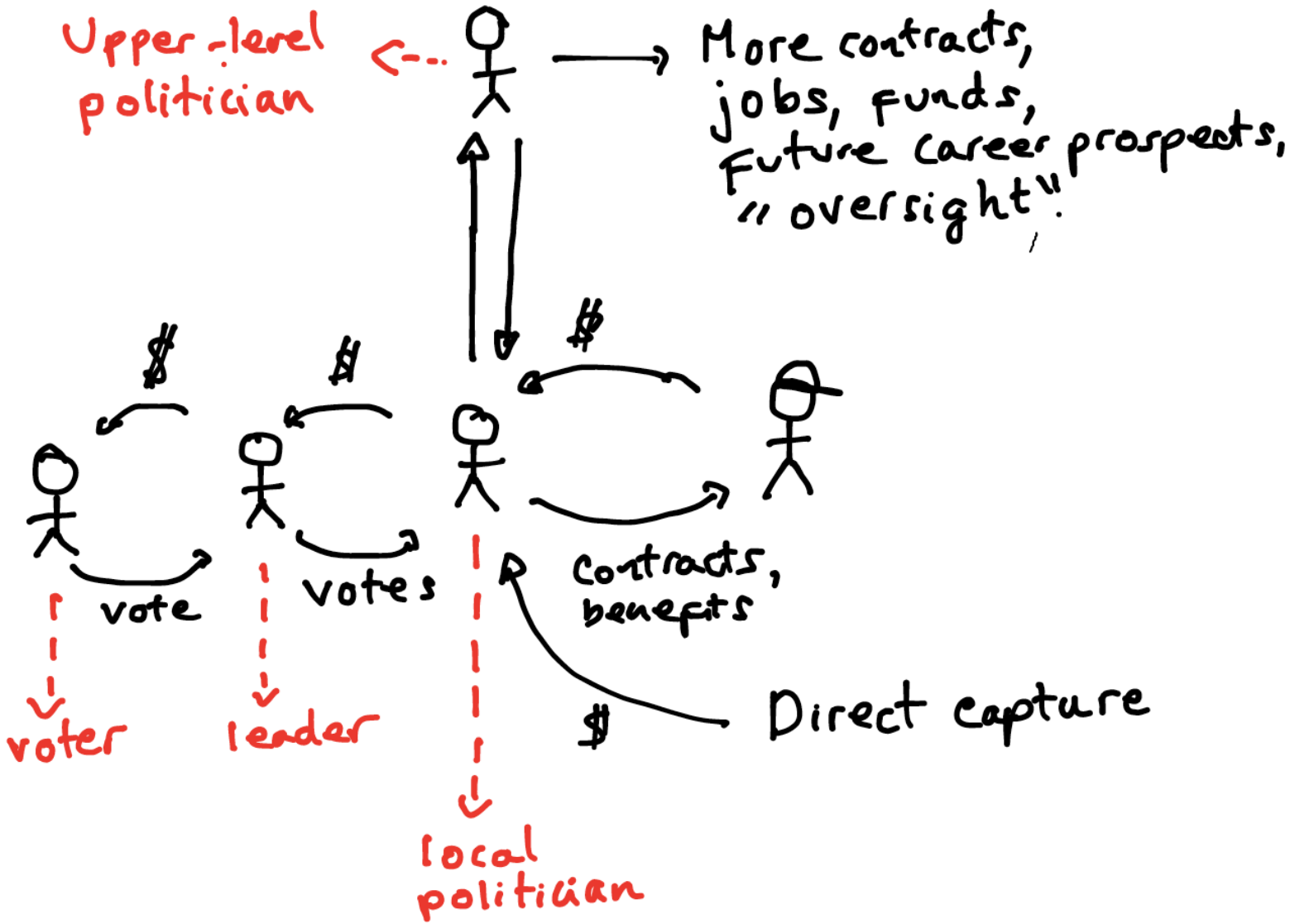
1. Corruption in developing countries and efficiency costs (Olken & Pande, 2012)
   - Local capture by political elites of government transfers (Reinikka & Svensson, 2004).
   - Efficiency costs?
     - Corruption in education can decrease quality
       Ferraz, Finan, and Moreira (2012).
     - Even when trickling down to citizens, ghosts could imply not realizing under-provision
       $\rightarrow$ may lead to inefficient underfunding (Olken, 2007, 2009).
2. Clientelism beyond particularism and implications for corruption (E.g. Maiz & Requejo, 2001; Singer, 2009)
3. Political alignment beyond fiscal transfer consequences (a long list)

Related literature and contribution II

4. Risks of incentive and fixed-rule schemes in financing public goods with weak oversight.
   (Among others, Acemoglu, Fergusson, Robinson, Romero, and Vargas (2016); Bold, Kimenyi, Mwabu, Ng'ang'a, and Sandefur (2013); Behrman, Parker, Todd, and Wolpin (2015))
5. Education: limited effects of more spending in education.

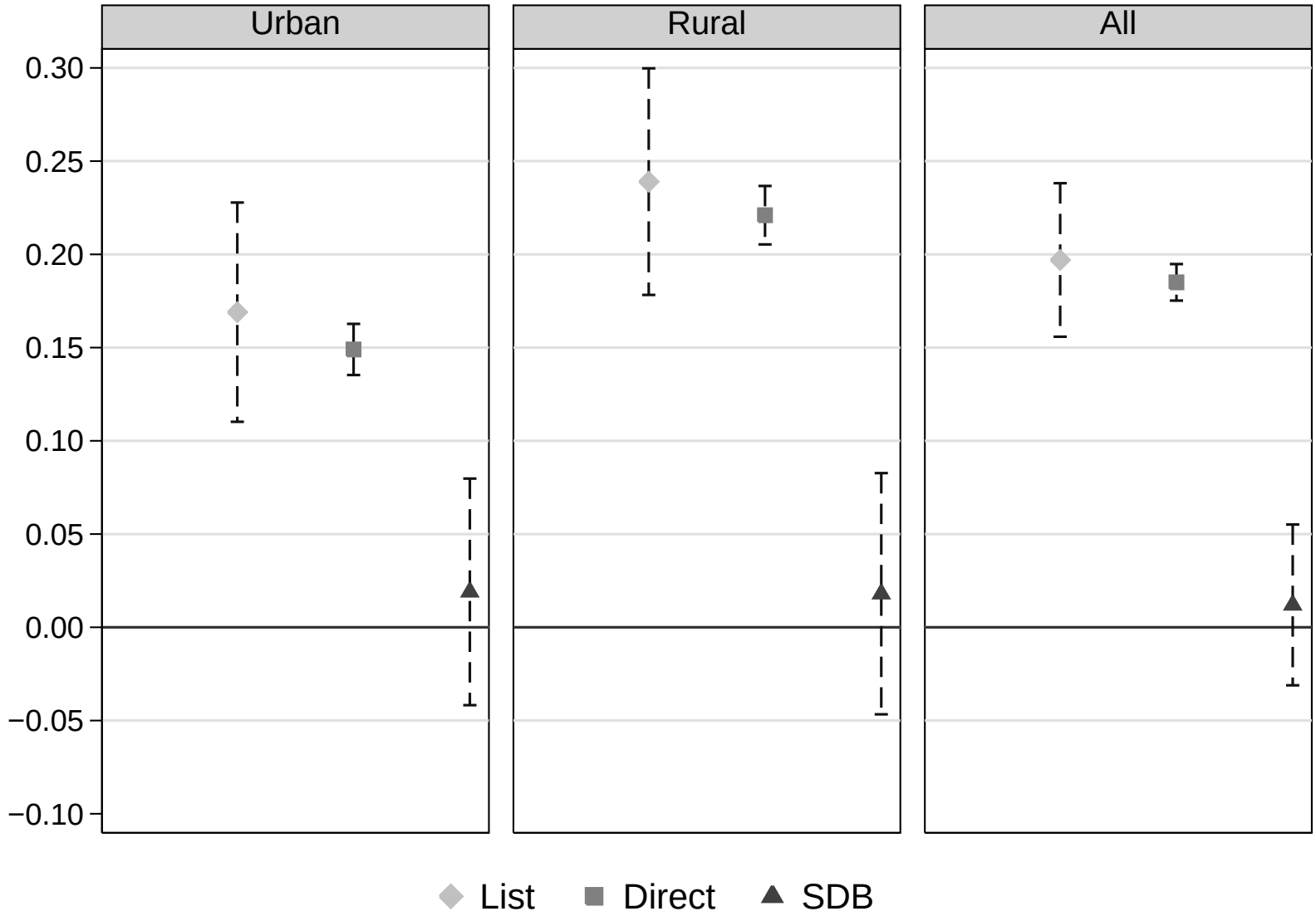
The political landscape

A "model"



The political landscape

Clientelism: vote buying (Fergusson, Molina, & Riaño, 2017)



Contents

Public education: institutional details

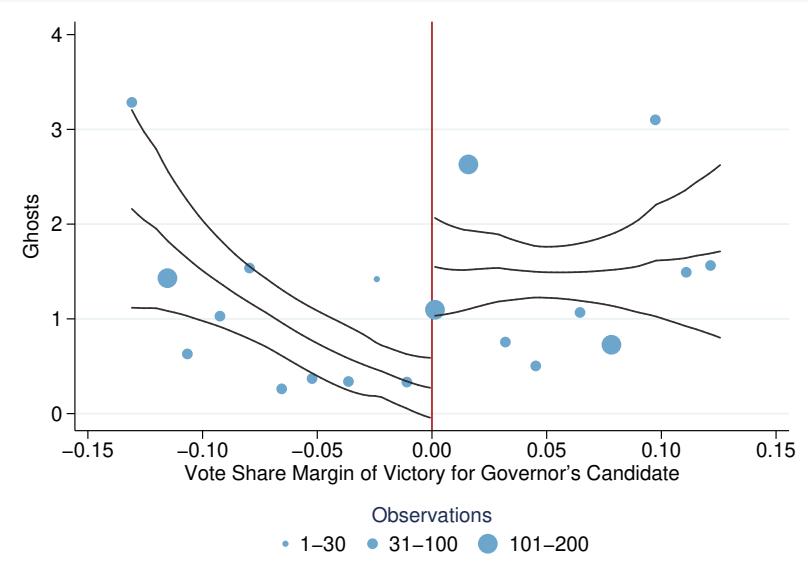
- Most money spent is from central government transfers (88%).
- The rest, from regional government (3%) and local government (9%), is only for investment.
- Enrollment is key in central government $:

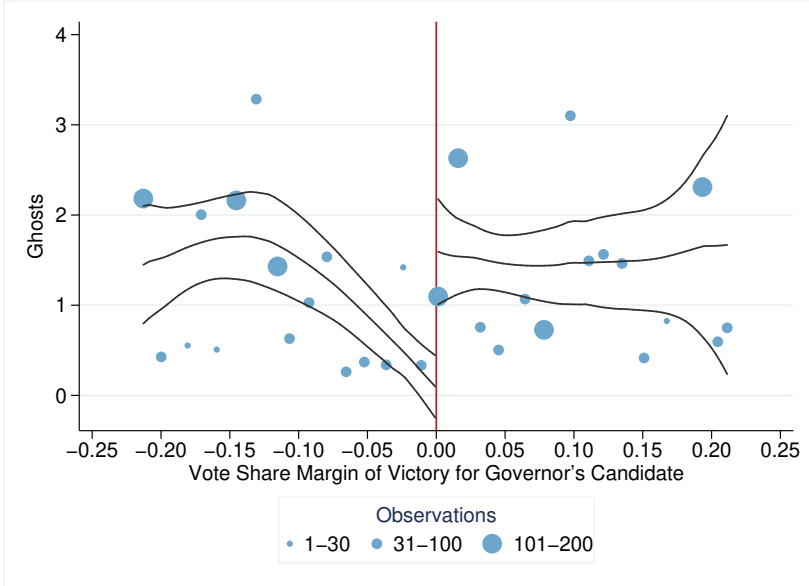
Account	% total resources	Transferred to	Distribution criteria
Payroll	90%	Paid directly by Ministry	# Teachers = f(**enrollment**)
Quality-enrollment (*calidad-matrícula*)	5%	Regional Education Secretaries	performance, poverty, **enrollment**
Quality-access (*calidad-gratuidad*)	5%	Schools	**enrollment**

Audit: census of schools

- Audit study in 2012 financed and managed by the national government (Ministry of Education).
  - Audit firms implementing the study were competitively selected.
  - Ministry functionaries sought to protect from cooptation: no local auditors.
  - Comprehensive audit that nearly reached the goal:
    - 8,167,051 from 8,679,035 students were audited (94.1%).
  - Detailed verification:
    - Face to face verification.
    - Complementary documentary evidence.
- Results
  - 148,410 ghosts.

Main results: ghosts (%): Graphical analysis





Notes: Local polynomial regression. Left: linear fit. Right: quadratic fit. Observations within Calonico et al. (2014) optimal bandwidth displayed. Bin selection method: mimicking variance, evenly spaced using spacings estimators.

Ghosts (%) and alignment

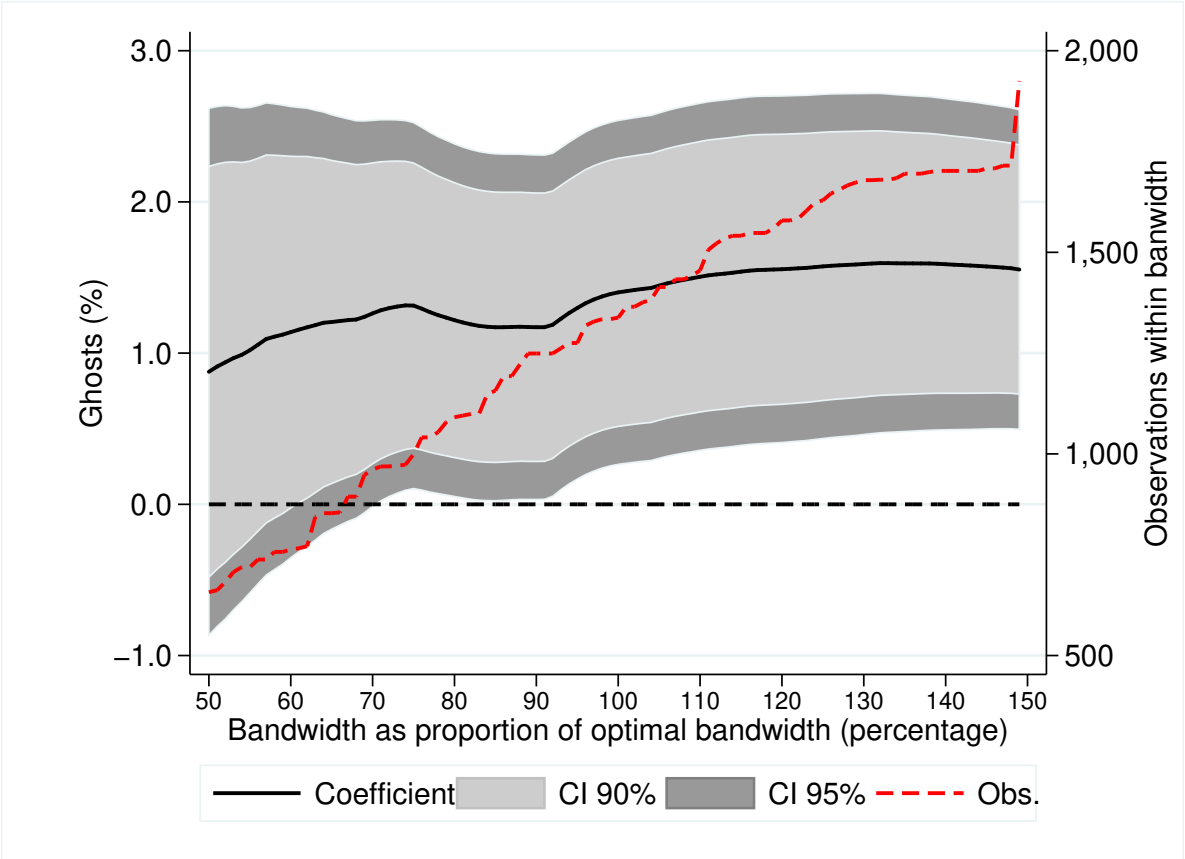
Nonparametric estimators with optimal bandwidth

Dependent variable is ghost students per school (in %).						
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. No controls						
Alignment	1.402**	1.272*	1.163	1.670**	1.510*	1.463*
	(0.692)	(0.708)	(0.730)	(0.748)	(0.783)	(0.861)
Observations	4,383	4,383	4,383	4,383	4,383	4,383
Bandwidth	0.130	0.118	0.105	0.242	0.197	0.146
Obs. in bandwidth	1338	1249	1091	2211	1926	1522
Kernel	Triangular	Epanechnikov	Uniform	Triangular	Epanechnikov	Uniform
Local polynomial Order	1	1	1	2	2	2
Panel B. Controls						
Alignment	1.336**	1.514**	1.660**	1.131**	1.206*	1.363**
	(0.561)	(0.708)	(0.659)	(0.514)	(0.696)	(0.682)
Student controls	✓			✓		✓
School controls		✓		✓		✓
Teacher controls			✓	✓		✓
Municipality and party controls					✓	✓
Observations	3,809	3,809	3,809	3,809	3,809	3,809
Bandwidth	0.101	0.140	0.164	0.127	0.0772	0.0781
Obs. in bandwidth	1044	1422	1633	1323	749	755

Notes: Calonico et al. (2014) optimal bandwidth with bias-corrected coefficients and robust standard errors clustered at municipality level. In Panel B, regressions are weighted using a triangular kernel and linear polynomial.

Ghosts (%) and alignment

Robustness to bandwidth choice



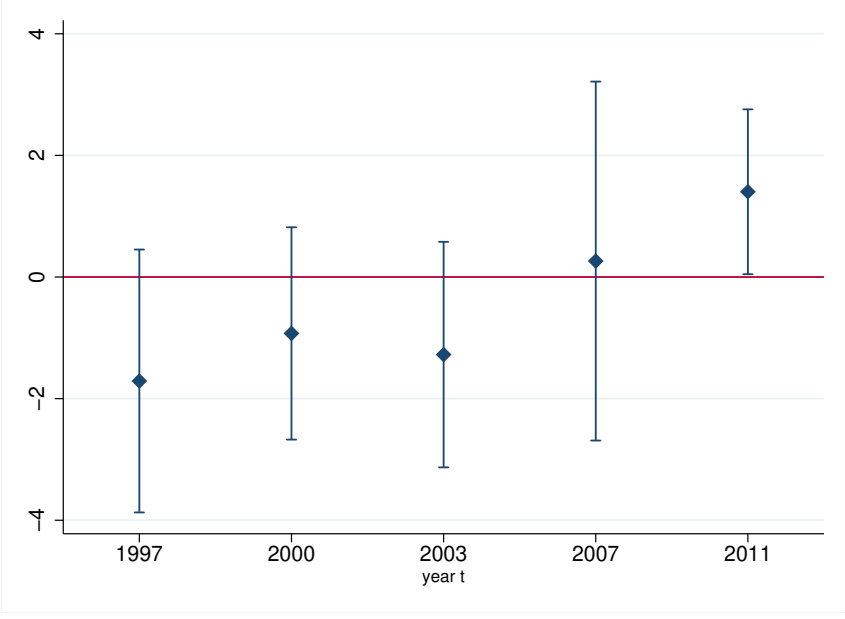
Notes: Nonparametric RDD estimators (triangular kernel and linear polynomial) with Calonico et al. (2014) bias-corrected coefficients (95% and 90% confidence bands) and robust standard errors.

More robustness

Predicting ghosts and previous alignment

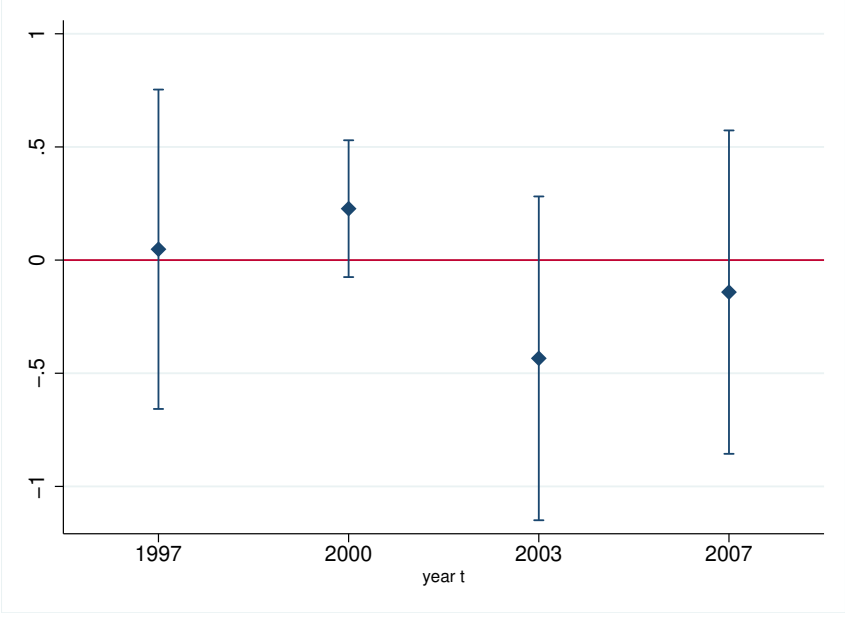
Panel A

$\text{Ghosts}_{2012} = \text{f}(\text{align}_t)$



Panel B

$\text{Align}_t = \text{f}(\text{align}_{2011})$

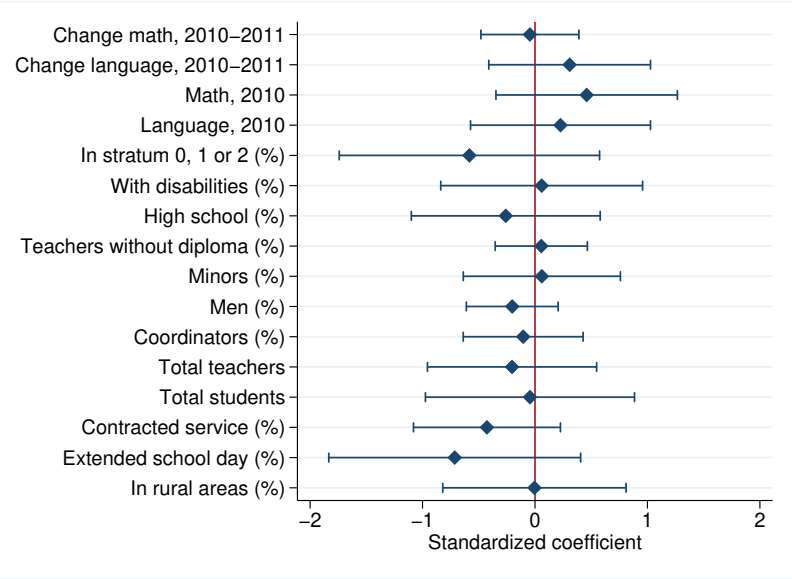


Notes: Nonparametric RDD estimators with Calonico et al. (2014) optimal bandwidth, bias-corrected coefficients (95% confidence bands) and robust standard errors (clustered at the municipality level in Panel A). In Panel A the dependent variable is ghost students from the 2012 census (in percent) and the treatment variable is mayor and governor party alignment in each election year marked in the x-axis. In Panel B the dependent variable is alignment between mayor and governor in each election marked in the x-axis and the treatment variable is mayor and governor party alignment in 2011.

Balance on observable variables

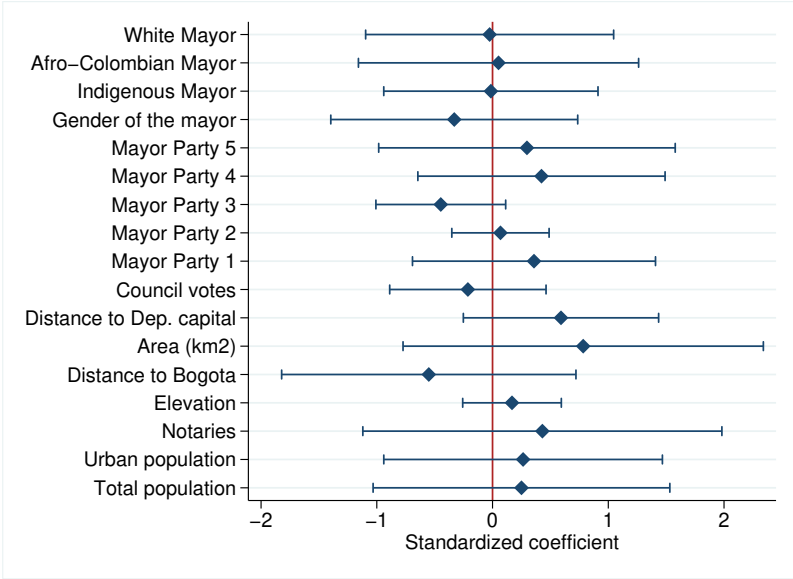
Panel A

School-level covariates



Panel B

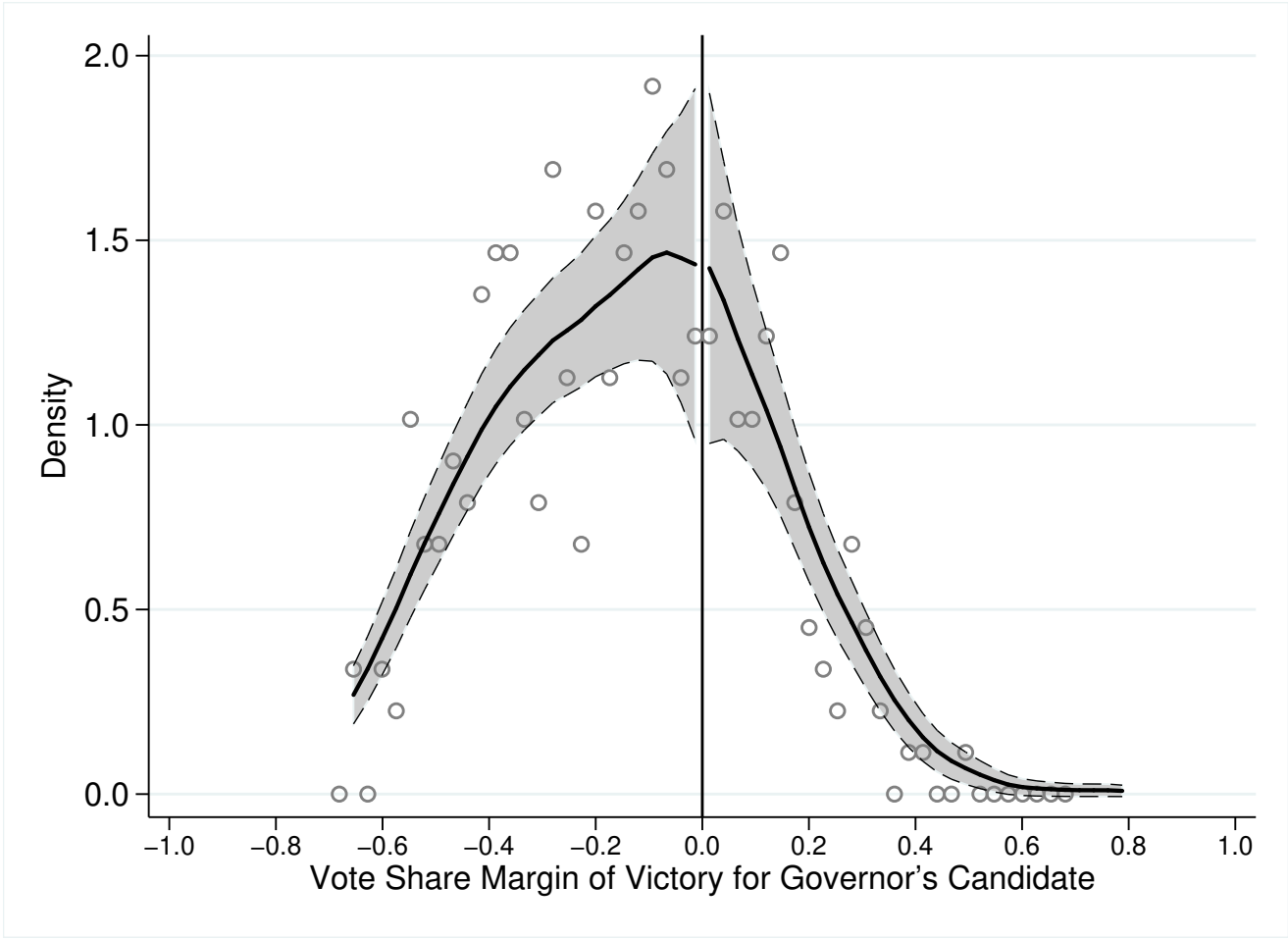
Municipality-level covariates



Notes: Nonparametric RDD estimators with Calonico et al. (2014) optimal bandwidth, bias-corrected coefficients (95% confidence bands) and robust standard errors (clustered at the municipality level in Panel A). All effects are standardized.

Manipulation?

McCrary (2008) density test



Note: Discontinuity estimate: 0.0268 (s.e. 0.2647)

What makes it worse?

- Autonomy:
  "You can take advantage of the certified municipality as a channel for good and bad things, but you can certainly take a lot of advantage."

  (Former Education Secretary)
- Municipal institutional weakness.
- School quality (staff qualification).

Corruption for personal gain or for public benefit?

- Thus far this could be
  - Aligned mayors increase ghosts because connections to higher echelons of power provide more opportunities for clientelism.
  - Diversion is stronger with more autonomy because this facilitates maximizing private rents.
  - And where institutions are weak because this reduces chances of getting caught.

Corruption for personal gain or for public benefit?

- Thus far this could be
  - Aligned mayors increase ghosts because connections to higher echelons of power provide more opportunities for clientelism.
  - Diversion is stronger with more autonomy because this facilitates maximizing private rents.
  - And where institutions are weak because this reduces chances of getting caught.
- Or
  - Officials fabricate ghosts to obtain resources and productively invest them.
  - Aligned mayors could have more interest to do this (improved coordination, credit claiming).
  - Stronger effects in autonomous municipalities reflects more incentives where there is more flexibility to productively use resources.
  - Larger impact in weakly institutionalized places reflects more urgent needs for extra funds.

Examining other implications

- Do these places receive better/more education?
  - Not better (test scores), nor more (ghost-adjusted enrollment rates).
- Do these politicians engage more in fraud, are officials more often accused of abuses?
  - Higher risk of fraud, more complaints.
- How do they do later on?
  - We could expect either, depending on the strength of democratic control versus entrenchment of clientelistic machine and favors returned.
  - Some, weaker evidence, that they do better electorally and in jobs.

To sum up

- Diversion of government funds may be used for private economic and political gain and/or to pass some of the benefits on to citizens.
- Our findings are consistent with the following key elements, which may arise where politics is highly clientelistic:
  1. Resource extraction is particularly valuable to politicians whose political network is connected to higher echelons of power.
  2. Areas with weaker accountability and more discretion for diversion of resources for the reproduction of the clientelistic network engage more in this form of corruption.
  3. A substantial part of the money is being diverted for political and economic gain rather than to improve the quantity or quality of the service.
- 1. is concerning because precisely that type of politician may persist in power.
- 2 and 3 are a problem because this exacerbate inequalities.

References I

Acemoglu, D., Fergusson, L., Robinson, J. A., Romero, D., & Vargas, J. F. (2016). *The perils of high-powered incentives: Evidence from colombia's false positives* (Tech. Rep.). NBER Working Paper # 22617.

Behrman, J. R., Parker, S. W., Todd, P. E., & Wolpin, K. I. (2015). Aligning learning incentives of students and teachers: results from a social experiment in mexican high schools. *Journal of Political Economy*, *123*(2), 325–364.

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2013). Scaling up what works: Experimental evidence on external validity in kenyan education.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326.

Fergusson, L., Molina, C., & Riaño, J. F. (2017). *I sell my vote, and so what? A new database and evidence from Colombia* (Documentos Cede No. 2017-20). Universidad de los Andes-Cede.

Ferraz, C., Finan, F., & Moreira, D. B. (2012). Corrupting learning: Evidence from missing federal education funds in brazil. *Journal of Public Economics*, *96*(9-10), 712–726.

Maiz, R., & Requejo, R. (2001). Clientelism as a political incentive structure for corruption. In *Trabajo presentado en la conferencia de primavera del european consortium for political research, grenoble.*

McCrary, J. (2008, February). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2). Retrieved from `http://eml.berkeley.edu/~jmccrary/mccrary2006_DCdensity.pdf`

Olken, B. A. (2007). Monitoring corruption: evidence from a field experiment in Indonesia. *Journal of political Economy*, *115*(2), 200–249.

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public economics*, *93*(7-8), 950–964.

Olken, B. A., & Pande, R. (2012). Corruption in developing countries. *Annual Review of Economics*, *4*(1), 479-509. Retrieved from `http://www.annualreviews.org/doi/abs/10.1146/annurev-economics-080511-110917` doi: 10.1146/annurev-economics-080511-110917

Reinikka, R., & Svensson, J. (2004, May). Local capture: Evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics*, *119*(2), 678-704. Retrieved from `http://ideas.repec.org/a/tpr/qjecon/v119y2004i2p678-704.html`

Singer, M. (2009). Buying voters with dirty money: The relationship between clientelism and corruption.

Descriptive Statistics I

	Full sample	Close races $< 13\%$	
		Parties aligned	Parties not aligned
Variable	(1)	(2)	(3)
A. School level			
Outcomes			
Ghosts_{2012} (%)	1.348	1.525	1.117
	(4.774)	(4.532)	(4.741)
Test score_{2012}	0.000	0.098	-0.102
	(1.000)	(1.092)	(0.896)
Test score_{2013}	-0.000	0.157	-0.153
	(1.000)	(1.054)	(0.880)
Test score_{2014}	0.000	0.156	-0.200
	(1.000)	(1.079)	(0.921)
Δ $\text{Test score}_{2012\text{-}2010}$	0.068	0.011	0.151
	(0.834)	(0.783)	(0.837)
Δ $\text{Test score}_{2013\text{-}2010}$	0.060	0.061	0.058
	(0.819)	(0.809)	(0.753)
Δ $\text{Test score}_{2014\text{-}2010}$	0.090	0.087	0.059
	(0.802)	(0.807)	(0.758)
Mayor-governor party alignment			
Aligned (dummy)	0.313	1.000	0.000
	(0.464)	(0.000)	(0.000)

Descriptive Statistics II

	Full sample	Close races < 13%	
		Parties aligned	Parties not aligned
Variable	(4)	(5)	(6)
B. Municipal level			
Outcomes			
Ghosts_{2012} (%)	1.266	1.712	0.892
	(2.346)	(4.147)	(1.253)
$\text{Total coverage rate}_{2012}$	0.879	0.870	0.900
	(0.188)	(0.203)	(0.210)
$\text{Total coverage rate}_{2013}$	0.883	0.882	0.902
	(0.195)	(0.204)	(0.234)
$\text{Risk of electoral fraud}_{2018}$	0.265	0.396	0.242
	(0.442)	(0.494)	(0.432)
$\text{All complaints}_{2012\text{-}2014}$	30.111	46.057	16.773
	(66.223)	(116.996)	(15.236)
$\text{Educ complaints}_{2012\text{-}2014}$	0.045	0.113	0.000
	(0.260)	(0.467)	(0.000)
$\text{Educ complaints}_{2012\text{-}2014}$ (%)	0.002	0.005	0.000
	(0.012)	(0.023)	(0.000)
Mayor-governor party alignment			
Aligned (dummy)	0.298	1.000	0.000
	(0.458)	(0.000)	(0.000)

Descriptive Statistics

Municipality level

	Mean	*Median*	*S.D.*	*Min*	*Max*
Ghosts (%)	1.266	1.266	2.346	0	23.00
Fiscal dependency	0.719	0.719	0.182	0.0948	0.981
Per capita property tax	25.25	25.25	37.04	0	344.3
Coca presence	57.58	57.58	346.6	0	4,846
Notaries	0.748	0.748	0.765	0	7
Displaced rate	336.8	336.8	524.1	0	3,649
Rural population	13,096	13,096	13,766	247	134,910
Urban population	22,776	22,776	56,267	159	534,956
Total population	35,872	35,872	61,843	1,015	535,642
Births	519.4	519.4	947.7	11	7,765
Mortality rate	24.71	24.71	9.964	7.770	63.16
Area	6.229	6.229	1.134	2.833	11.09
Elevation	754.2	754.2	1,565	2	25,221
Distance to Bogota	390.4	390.4	190.8	49.91	1,271
Car theft	2.280	2.280	9.497	0	103
Number of schools	2.042	2.042	6.387	0	61
Investment on education	6.380	6.380	19.96	0.0496	162.2
SGP on education	0.155	0.155	0.154	0	0.910

Back

Extensive margin, outliers, donut hole, large cities

Dependent variable is presence of ghost dummy (col. 1) and share of ghosts students (in %) per school (cols. 2 to 10).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Extensive margin	Dropping outliers *Extreme values*			Donut hole *% of the optimal bandwith*			Dropping large cities *Municipalities with a pop. of over...*		
		1 %	3 %	5 %	1 %	3 %	5 %	1 M	500 K	100 K
Alignment	0.0653	1.003**	0.756*	0.778**	1.411**	1.304*	1.331*	1.402**	1.444**	2.149**
	(0.186)	(0.464)	(0.403)	(0.305)	(0.698)	(0.685)	(0.723)	(0.692)	(0.673)	(1.049)
Observations	4,383	1,182	1,161	1,135	4,380	4,348	4,334	4,383	4,193	3,531
Bandwidth	0.112	0.112	0.112	0.112	0.130	0.125	0.124	0.130	0.151	0.120
Obs. in bandwidth	1193	1182	1161	1135	1338	1278	1278	1338	1479	987

Notes: Regressions are weighted using a triangular kernel and assuming a linear polynomial. Robust standard errors clustered at municipality level. Optimal bandwidth in all columns.

Municipal level regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Alignment	3.872*	3.931*	2.232	5.270*	5.413*	3.786
	(2.382)	(2.380)	(1.919)	(2.851)	(2.867)	(2.566)
Observations	332	332	332	332	332	332
Bandwidth	0.119	0.107	0.126	0.159	0.145	0.165
Obs. in bandwidth	109	96	117	143	133	153
Kernel	Triangular	Epanechnikov	Uniform	Triangular	Epanechnikov	Uniform
Local polynomial Order	1	1	1	2	2	2

Notes: Mean of dependent variable is 1.266 and standard deviation is 2.346. Bias corrected coefficients and robust standard errors clustered at municipality level (Calonico, Cattaneo, & Titiunik, 2014). Optimal bandwidth in all columns. * 10%, ** 5%, ***1%.

Back