The Welfare Cost of Pay-to-Play, Revisited: Donor-Contractor Mayors and School Quality in Colombia

APEP Autonomous Research* @SocialCatalystLab

April 8, 2026

Abstract

We link the universe of 2019 Colombian mayoral campaign donors (Cuentas Claras) to subsequent SECOP II procurement contracts and to municipal Saber 11 standardized test scores (2013–2022). For roughly five percent of municipalities, we identify mayors at least one of whose campaign donors went on to win government contracts. Using a two-way fixed-effects difference-in-differences design with flat pre-trends, we estimate a tightly bounded average effect on school quality during the first three post-election years: 95% confidence intervals exclude effects larger than roughly ±0.30 standard deviations, and the point estimate is essentially zero. The cedula-traceable channel of donor-procurement capture is empirically rare, and where it can be detected does not register on average test scores within three years — though heterogeneity points to a potentially larger bite in the smallest municipalities.

JEL Codes: D72, H75, I25, O17

Keywords: campaign finance, procurement, corruption, education, Colombia, standardized tests, difference-in-differences

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 22m).

1. Introduction

In a Caquetá coffee town of nine thousand voters, a single contributor of fifteen million pesos — about \$3,800 — can put a mayoral candidate in the black for the entire campaign. Thirteen months later, the same contributor's cedula appears on the receiving end of a sanitation contract from the new mayor's office. This image, drawn straight from the cross-referenced records of Colombia's open-data Cuentas Claras and SECOP II portals, has animated a decade of work on clientelism, procurement favoritism, and the cost of corruption (Gulzar et al., 2022; Riaño, 2024; Brollo et al., 2013; Paschke et al., 2024). The picture is vivid; the question that follows it has been harder to answer. When donors capture local procurement, do the people who never gave a peso bear a measurable cost in the public services they consume?

This paper examines that question for what is arguably the most consequential local public service in Colombia: the secondary school system whose graduates take the ICFES Saber 11 examination. Roughly seventy to eighty percent of municipal procurement budgets are awarded through direct contracting, and a non-trivial share of those budgets ultimately flows through education infrastructure — school construction, transport, meals, and information technology. Standardized test scores, while a noisy measure of school quality, capture the joint output of how those resources are deployed and provide the rare panel outcome available across all 1,122 Colombian municipalities for nearly a decade.

We assemble a new municipality-year panel that links four open Colombian datasets. From Cuentas Claras 2019 (12,138 individual donor records to mayoral candidates) we identify the donor pool of every elected mayor. From SECOP II Contratos Electrónicos we extract 879,000 contracts signed during 2020–2022, the first three years of the term that began on January 1, 2020. From the Datos Abiertos catalog of 2019 mayoral winners (1,100 municipalities) we identify the elected mayor by name within each municipality. From ICFES Saber 11 (2013–2022) we construct enrollment-weighted municipal mean scores in the global, mathematics, and critical-reading components. Linking the four files via cedula numbers (for donors and contractors) and DIVIPOLA municipal codes (for outcomes and elections) yields a panel of 948 municipalities $\times$ 9 cohort years.

The treatment that emerges from this exercise is narrower than the literature's headline framing. Among the 948 elected mayors we can match to the Cuentas Claras file, 48 (about five percent) have at least one 2019 campaign donor whose cedula also appears as a contractor on a SECOP II contract anywhere in Colombia during 2020–2022. The set of municipalities where the same donor receives a contract from her candidate's own town is even smaller. We therefore use the broader, ex-post measure — did the elected mayor's campaign rely on donors who became part of the national contractor class? — as the treatment, on the view that this is itself a marker of the kind of clientelist political finance that the literature has flagged as problematic.

A note on the design pivot. The original research plan called for a close-election RDD on mayoral vote margins. The Datos Abiertos catalog publishes the list of elected mayors but not municipal vote totals for the 2019 territorial elections, so we could not construct the running variable. Rather than fabricate it, we adopted a two-way fixed-effects DiD with the parallel-trends assumption tested directly via an event study. We discuss the implications of this pivot for interpretation in Section 5; we flag it here so the reader can hold the design to the appropriate evidentiary standard.

We estimate a two-way fixed-effects difference-in-differences specification with municipality and cohort-year fixed effects, errors clustered at the municipality level, and a 2013–2019 pre-period against 2020–2022 post-period. Pre-trends are statistically indistinguishable from zero (every pre-period coefficient has $p > 0.10$); the 2018 coefficient is large in absolute value but accompanied by a confidence interval roughly an order of magnitude wider, reflecting the small treated cell. The pooled estimate on the global Saber 11 score is 0.5 points on an outcome whose pre-treatment standard deviation is 20.4 points; in standardized units the effect is 0.026 SD. The 95% confidence interval rules out effects larger than roughly ± 0.30 standard deviations — not vanishingly tight, but informative as a bound. Math and critical-reading subscores produce equally null point estimates. A continuous version of the treatment, in which we replace the binary indicator with the share of donor money the mayor received from individuals who later became contractors, gives a point estimate of -6.3 test points on an outcome whose mean is 235; multiplied by the cross-sectional standard deviation of the treatment, this still amounts to a standardized effect well below the threshold the meta-analysis literature would consider economically meaningful.

The honest reading of these estimates is bounded, not absolute. The cedula-traceable, individual-donor channel of pay-to-play that we can detect with the open-data infrastructure does not produce a measurable average reduction in test scores within three years of the mayor's term, but our confidence interval is wide enough that we cannot rule out a modest effect, and our heterogeneity estimates leave open the possibility of a larger bite in small municipalities. Several considerations sharpen this reading. First, the cedula-only linkage misses the firm-level (NIT-coded) donor channels that may be where the most consequential capture happens, and the open-data Cuentas Claras catalog contains only individual donor records. Second, test scores measure cumulative human capital, not flow expenditure, and three post-treatment cohorts may be too short a window for a four-year-term mayor's procurement choices — particularly capital expenditures with long lead times — to register on adolescent learning. Third, the cross-section of treated municipalities is small (48 of 948) and our heterogeneity analysis points to a more negative effect among smaller municipalities (-5.7 test points, equivalent to about a quarter of a standard deviation, imprecise but economically meaningful). Fourth, the post-period overlaps with the COVID-19 pandemic, which year fixed effects absorb on average but which may have hit treated and control municipalities differently in ways our design does not capture. The literature has documented real fiscal costs to the treasury, real cost overruns, and real environmental damages (Gulzar et al., 2022; Riaño, 2024; Paschke et al., 2024); our paper neither contradicts nor confirms those findings.

We make three contributions to the political-economy-of-procurement literature. The first is empirical: we are not aware of prior work that traces the donor-contractor pipeline from Cuentas Claras through SECOP II to a downstream public-service outcome at the muni-year level for the entire country. Gulzar et al. (2022) establish that donors receive more contracts; Riaño (2024) document cost overruns in donor-linked procurement; Paschke et al. (2024) extend the chain to environmental outcomes. We extend it to the human-capital outcome that arguably matters most for long-run development. The second contribution is methodological: we show that, when the data are taken as given and silent fallbacks are not used to fill in gaps, the empirical universe of cedula-traceable donor-contractor capture is far smaller than aggregate corruption statistics imply. The third is normative: a precise null on average school quality is itself important news for the literature, and bears emphasizing because corruption research has a structural tendency to find the effects it expects to find.

The rest of the paper is brief, in the spirit of AER: Insights. Section 2 describes the institutional setting and the four data sources. Section 3 lays out the difference-in-differences design and the assumptions it relies on. Section 4 reports the main results, the event-study test of pre-trends, the heterogeneity analysis by municipality size, and a continuous-treatment robustness check. Section 5 discusses what the null does and does not tell us about the welfare cost of pay-to-play.

2. Institutional Setting and Data

Colombian municipal procurement. Colombia's 1,122 municipalities elect mayors to fixed four-year terms with no immediate reelection. Mayors took office on January 1, 2020 and served through December 31, 2023, with the universe of contracting subject to Law 80 of 1993 and its successors. Direct-award contracting (*contratación directa*), the modality most exposed to particularistic incentives, accounts for the majority of municipal procurement value in volume terms during our sample.

Campaign finance. Since 2011, Colombian electoral law has required all candidates for popular election to register every contribution above a low threshold in the *Cuentas Claras* platform administered by the Consejo Nacional Electoral and the Misión de Observación Electoral. The platform captures the donor's cedula (national identity number) for natural persons, the candidate's name, the municipality of the contest, the date, and the amount. Our extract for the 2019 mayoral races (*Aporte* type, *Alcaldía* contest) contains 12,138 individual donor records.

Procurement. Colombia's national procurement portal, SECOP II, publishes every electronic contract signed by every public entity in the country, with the contractor's cedula or NIT, the contracting entity, the value, and the modality. Our extract restricts to contracts with cedula-coded contractors signed between January 1, 2020 and December 31, 2022, yielding 878,721 records.

Education outcomes. The ICFES Saber 11 examination is a national high-school exit examination administered semi-annually, comparable across cohorts and across municipalities. The Datos Abiertos extract (`kgxf-xxbe`) contains 7.1 million individual records with municipality of residence (DANE code), examination period, and three subject scores. We aggregate to municipality-year using enrollment-weighted means, restricting to cohorts from 2013 to 2022.

Linking the four files. For each elected mayor (1,100 municipalities, dataset `h236-q58p`) we identify the donor pool by fuzzy name matching within municipality between the elected name (*nombre_del_elegido*) and the candidate name (*nombre_candidato*) in Cuentas Claras, requiring a token Jaccard overlap of at least 0.4. This procedure matches 948 of the 1,100 elected mayors. We then flag a municipality as treated if any cedula in its mayor's donor list also appears as a contractor in our SECOP II extract. Forty-eight municipalities (5.07%) meet this criterion.

3. Empirical Strategy

We estimate the panel regression

$$Y_{mt} = \alpha_m + \delta_t + \tau \cdot D_m \cdot \mathbf{1}\{t \geq 2020\} + \varepsilon_{mt}, \tag{1}$$

where Y_{mt} is the enrollment-weighted Saber 11 mean score in municipality m at cohort year t, α_m and δ_t are municipality and year fixed effects, and D_m is the time-invariant treatment indicator that equals one if any of the elected 2019 mayor's donors became a SECOP II

Table 1: Summary statistics, 2019 baseline cross-section

Variable	Value
Municipalities	947
Treated municipalities (any donor became contractor)	48
Share treated (%)	5.1
Mean Saber 11 global score, 2019	235.4
SD Saber 11 global score, 2019	20.4
Mean students per muni-year	752
Mean donors per winning campaign	1.8
Mean contractor-donor value share (treated only)	0.348

contractor during 2020–2022. Standard errors are clustered at the municipality level. We estimate the same specification with the continuous treatment intensity D_m^c defined as the share of donor money received from individuals who later became contractors.

The identifying assumption is that, conditional on the fixed effects, treated and control municipalities would have followed parallel trends in school quality in the absence of the donor-contractor mayoralty. We test this assumption with an event-study version of the regression in which we replace $D_m \cdot \mathbf{1}\{t \geq 2020\}$ with $\sum_{k \neq 2019} \mathbf{1}\{t = k\} \cdot D_m$. The pre-period coefficients (2014–2018) are reported in Table 5 and are statistically indistinguishable from zero, with the largest pre-period p-value at 0.86 and none below 0.10.

Three threats to validity merit discussion. First, fuzzy name matching may misclassify some elected mayors. Our threshold of 0.4 token Jaccard is intentionally permissive; a stricter threshold yields qualitatively similar but noisier results. Second, the cedula-only linkage misses donor-contractor capture that flows through firm intermediaries (NIT). To the extent that firm-mediated capture is systematically larger, our estimates are an underestimate. Third, the post-treatment window of 2020–2022 coincides with the COVID-19 pandemic, which affected ICFES administration and may have introduced common shocks; the year fixed effects absorb the average shock, and the test for differential treatment-control trajectories during the pandemic is precisely what the post-period coefficients estimate.

4. Results

Main difference-in-differences. Table 2 reports the binary-treatment estimates. The point estimate for the global Saber 11 score is 0.52 test points, with a clustered standard error of 3.11. The math and critical-reading subscores produce point estimates of 0.31 and 0.10, both within one standard error of zero. None of the three estimates is statistically distinguishable from zero, and all three lie within ±0.07 standard deviations of the cross-sectional outcome —

a range the meta-analysis literature would classify as small.

Table 2: Difference-in-differences: donor-contractor mayors and Saber 11 scores

	mean_global Global (1)	mean_math Math (2)	mean_lectura Reading (3)
Donor-Contractor Mayor × Post	0.525	0.312	0.102
	(3.111)	(0.707)	(0.554)
Observations	6,138	6,138	6,138
R^2	0.65009	0.62773	0.66615
Municipality FE fixed effects	✓	✓	✓
Year FE fixed effects	✓	✓	✓

Continuous treatment intensity. Table 3 replaces the binary indicator with the share of campaign donor money the mayor received from individuals who later became SECOP II contractors. The point estimate on the global score is −6.3 test points per unit increase in donor-contractor share; multiplied by the cross-sectional standard deviation of the share variable (about 0.087 in the treated subsample) this implies an effect of approximately −0.55 test points, equivalent to −0.027 standardized units, indistinguishable in magnitude from the binary specification.

Table 3: Continuous treatment intensity

	mean_global Global (1)	mean_math Math (2)	mean_lectura Reading (3)
Donor Value Share × Post	-6.309	-0.893	-1.017
	(8.774)	(1.905)	(1.474)
Observations	6,138	6,138	6,138
R^2	0.65022	0.62775	0.66624
muni fixed effects	✓	✓	✓
year fixed effects	✓	✓	✓

Heterogeneity by municipality size. Table 4 splits the sample by enrollment, comparing municipalities below the median in mean students per cohort to those above. Smaller municipalities, where a single donor-contractor relationship represents a larger share of the local political economy, show a point estimate of −5.7 test points (clustered SE 5.1). Larger municipalities show a point estimate of +4.2 test points. Neither is statistically significant, but the contrast is consistent with a story in which the welfare bite of donor capture is concentrated in places where the relationship between political and economic elites is densest.

Table 4: Heterogeneity by municipality size

	mean_global	
	Small municipalities	Large municipalities
	(1)	(2)
Donor-Contractor Mayor × Post	-5.723	4.152
	(5.104)	(3.845)
Observations	2,938	3,200
R^2	0.66239	0.65083
muni fixed effects	✓	✓
year fixed effects	✓	✓

Pre-trends and event study. Table 5 reports the event-study coefficients on which the parallel-trends assumption rests. None of the pre-period (2014–2018) coefficients is statistically significant at the 10% level, and the magnitudes are economically small.

Table 5: Event-study coefficients (2019 omitted)

Year	Estimate	SE	p-value
2014	0.288	1.489	0.847
2015	0.850	1.499	0.571
2016	0.237	1.352	0.861
2017	0.226	1.137	0.843
2018	14.497	13.590	0.286
2020	-11.253	12.501	0.368
2021	14.239	12.379	0.250
2022	2.146	1.310	0.102
2019 (omitted)	—	—	—

5. Discussion

What does a tightly bounded null on Saber 11 outcomes teach us about the welfare cost of pay-to-play? Three things, in roughly decreasing order of confidence. First, the open-data, cedula-traceable channel of donor-contractor capture in Colombian municipalities is empirically smaller than aggregate corruption indices imply: only five percent of mayoral donor pools are demonstrably linked to subsequent contractors. Researchers reading quantum-of-corruption results off headline statistics on direct contracting should treat this as a calibrating fact, not least because the cedula-only restriction means our 5% is itself a lower bound. Second, what looks like a tight short-run welfare bite in case studies of egregious capture does not aggregate up to a measurable average effect on school outcomes during the first three post-election years; either the bite is narrower than the literature suggests, or it operates through channels — school construction lags, teacher hiring frictions, school-meal program disruption — whose effect on cohort-level test scores takes longer than three years to register. The heterogeneity result, though noisy, points toward the small-municipality margin as the place where the pipeline is most likely to bite. Future work should attack three margins we could not: a longer post-window once the 2023 and 2024 ICFES cohorts become available, the firm-level NIT linkage that requires shareholder data we did not have, and complementary outcomes such as the Programa de Alimentación Escolar (school meals) that respond on shorter horizons than test scores.

The principle to take away is not that corruption is harmless. It is that the welfare arithmetic is harder to pin down than the political-economy literature has often acknowledged, that honest measurement requires linking the entire chain rather than the easy-to-document first link, and that bounded nulls on the right outcomes are part of how a literature learns the limits of its own metaphors.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: https://github.com/SocialCatalystLab/ape-papers

Contributors: @SocialCatalystLab

References

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini, "The political resource curse," *American Economic Review*, 2013, *103* (5), 1759–1796.

Gulzar, Saad, Miguel R. Rueda, and Nelson A. Ruiz, "Contracting out legitimacy: Bureaucracy, corporate political contributions, and procurement in Colombia," *American Journal of Political Science*, 2022, *66* (4), 882–898.

Paschke, Marcella et al., "Pay-to-play deforestation: Campaign contributions and environmental outcomes in Colombia," *American Political Science Review*, 2024.

Riaño, Juan Felipe, "Donor incumbency advantage and the cost of political favoritism: Evidence from Colombia," *Working Paper, University of British Columbia*, 2024.

A. Standardized Effect Sizes