Political Agency, Election Quality, and Corruption*

Online Appendix (not intended for publication)

Miguel R. Rueda† Nelson A. Ruiz‡

March 1, 2018

*We thank participants and audiences at the PIM colloquium at Emory University, Midwest 2017 Panel on Political Economy of Corruption, PSPE Doctoral Workshop at the London School of Economics, Graduate Studies Wallis Conference at the University of Rochester, South East Latin American Political Behaviour Mini-Conference, and in particular Guillermo Rosas for their helpful comments and advice. We also thank Luis R. Martinez and the staff of Transparency International Colombia for kindly making their data available and for their assistance at different stages of the project. All errors are our own.

†Assistant Professor. Department of Political Science, Emory University. Email: miguel.rueda@emory.edu.

‡Affiliate. International Development Department, London School of Economics and Political Science and Public Policy Group E-TH Zürich. Email: n.a.ruiz-guarin@lse.ac.uk.

A Proofs and Competitive Manipulation

Proof of Proposition ??

We first derive an expression for the probability of the incumbent winning, which allows us to describe her maximization problem. We proceed with comparative statics.

Let r_1^e be the expected rents by the voters at the time of the election when manipulation does not influence the elections. This value does not depend on η, since they do not know it at that point. An estimate of the incumbent's η is then $\frac{g_1}{(R-r_1^e)}$. Voters vote for the incumbent whenever

$$\frac{g_1}{(R-r_1^e)} \geq 1,$$

which, using the budget constraint, happens with probability

$$P\left(\eta \geq \frac{R-r_1^e}{R-r_1}\right) = \frac{1}{2} + \xi\left(1 - \frac{R-r_1^e}{R-r_1}\right).$$

Let $\theta \equiv E + \overline{r}$. The maximization problem of the incumbent is

$$\max_{0\leq r\leq \overline{r}, 0\leq m\leq 1} \; r - c(m) + \left[m\left(\frac{1}{2}+\chi\right) + (1-m)\left(\frac{1}{2} + \xi\left(1 - \frac{R-r_1^e}{R-r_1}\right)\right)\right]\theta.$$

The first order conditions in an interior equilibrium are

$$r^* = R - (1-m^*)\theta\xi,$$

and

$$c'(m^*) = \theta\chi.$$

The second order sufficient condition for a maximum is $2(1-m^*)c''(m^*) > \theta\xi$. The left hand side of this inequality does not depend on ξ and is positive for typical strictly convex cost functions in an interior solution. Small enough values of ξ would satisfy the condition.

To prove statements 2 and 3, apply the Implicit Function Theorem to the second first order condition to see that

$$\frac{\partial m^*}{\partial \theta} = \frac{\chi}{c''(m^*)}$$

and

$$\frac{\partial m^*}{\partial \chi} = \frac{\theta}{c''(m^*)},$$

which are both positive. As for the rents,

$$\frac{\partial r^*}{\partial \theta} = \xi \left(-1 + m^* + \frac{c'(m^*)}{c''(m^*)} \right).$$

Note that if $c'(1) < \theta\chi$ there is no interior solution, and $m^* = 1$ and $r = \overline{r}$ in equilibrium. In this case, the level of rents is not affected by higher values of office. For interior solutions, there is a value of office, $\overline{\theta}$, such that $m^* = 1$. Given that $\frac{\partial r^*}{\partial \theta}$ is a continuous function of θ, and that $\lim_{\theta \to \overline{\theta}^-} \frac{\partial r^*}{\partial \theta} > 0$, the second statement is proven. □

Competitive manipulation

We now consider a setting in which both the challenger and the incumbent are allowed to engage in electoral manipulation at the beginning of the first period. We denote the manipulation level chosen by the challenger by m_C and that of the incumbent by m_I. The probability of election results being influenced by manipulation is $m_I + m_c$ whenever this fraction does not go above unity or 1 otherwise. Lastly, we assume that when manipulation influences the results, the challenger will win whenever

$$u(m_I) - u(m_c) \geq \delta,$$

where δ is a shock that is distributed uniformly in $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$ and u is a twice continuously differentiable function with $u'(.) > 0$ and $u''(.) < 0$. The term δ captures the uncertainty regarding the relative effectiveness of manipulation between parties. Everything else remains as in the original model's setting with only the incumbent choosing the level of rents and both parties having the same cost of manipulation function. We focus on characterizing an interior symmetric equilibrium of this game.

The probability of the incumbent winning conditional on manipulation influencing the outcome is

$$\frac{1}{2} + \phi\left(u(m_I) - u(m_C)\right).$$

Solving the maximization problem of the incumbent and the challenger, we find that the equilibrium level of manipulation, m^*, chosen by both parties satisfies

$$-c'(m^*) + \theta\phi 2m^* u'(m^*) = 0$$

and the rents are

$$r^* = R - \theta\xi(1 - 2m^*).$$

A sufficient condition for these solution to be a maximum is

$$(1 - 2m^*)\left[c''(m^*) - \theta\phi(2u'(m^*) + 2m^* u''(m^*))\right] > \theta\xi.$$

For commonly used strictly convex cost functions and strictly concave utility functions (power, exponential, and logarithmic) the left-hand side is positive and a small ξ would satisfy the inequality.

Using the Implicit Function Theorem we see that

$$\frac{\partial m^*}{\partial \theta} = -\frac{\phi 2m^* u'(m^*)}{-c''(m^*) + \theta\phi(2u'(m^*) + 2m^* u''(m^*))},$$

which is positive for an interior maximum. As for the rents,

$$\frac{\partial r^*}{\partial \theta} = \xi \left(-1 + 2m^* + 2\theta \frac{\partial m^*}{\partial \theta} \right).$$

Note that if $-c'(1/2) + \theta\phi u'(1/2) > 0$, there is no interior solution. Let $H(m) = \frac{c'(m)}{2mu'(m)}$. If u and c are such that m^* is strictly increasing on θ, there is a $\overline{\theta}$, such that $m^* = 1/2$. Given that $\frac{\partial r^*}{\partial \theta}$ is a continuous function of θ, and that $\lim_{\theta \to \overline{\theta}^-} \frac{\partial r^*}{\partial \theta} > 0$, we conclude that the derivative is positive for large enough values of office.

Finally, it is easy to see that the conditional probability of the incumbent winning when manipulation is effective is 1/2 in equilibrium, which is the same as the probability of the incumbent winning conditional on manipulation not being successful. □

B Variable Definition

Table 1: Variable Definitions and Sources

Variable	Description
Armed actor	Dummy that takes the value of 1 if there was combat in which either guerrillas or paramilitary forces were involved, or if there was a unilateral military action taken by any of these groups. Source: CERAC.
Own revenues	Revenues from the local government as a share of the municipalities' total revenues. Source: National Planning Department.
Margin	Average of all margins of victory in races in a given year weighted by valid votes in each race in a municipality. Margins for plurality elections are calculated as the gap between the winner's and the runner-up's votes. For proportional representation races, margins are the gap between the electoral quotient of the party winning the final seat and the electoral quotient of the closest loser. Source: National Registrar's Office and authors' calculations.
Polling station size (Actual)	Population 20 years or older per polling place in the municipality. Source: DANE, National Registrar's Office, and authors' calculations.
Rural Population	Fraction of the population living in a rural area in the municipality. Source: University of los Andes CEDE municipal panel.
Underperforming schools	Share of schools in the municipality classified below 'average performance' by the Instituto Colombiano para la Evaluación de la Educación (ICFES). Source: University of los Andes CEDE municipal panel.
Total population	Total population. Source: DANE.

C Measurement Errors, Linear Models, and IVs

Consider the population model $y = x\beta + \varepsilon$. We have data measured with error $\tilde{y} = y + u$ and $\tilde{x} = x + v$. Further, suppose that $cov(x, v) = cov(x, \varepsilon) = cov(v, \varepsilon) = 0$. The

OLS estimate of β is

$$\hat{\beta}^{OLS} = \frac{cov(\tilde{y}, \tilde{x})}{var(\tilde{x})} = \frac{cov(x\beta + \varepsilon + u, x + \upsilon)}{var(x + \upsilon)}$$

and

$$plim\ \hat{\beta}^{OLS} = \frac{\beta var(x) + cov(x, u) + cov(u, \upsilon)}{var(x) + var(\upsilon)}.$$ [1]

A higher rate of underreporting of corruption where vote buying is common implies that $cov(x, u) < 0$. On the other hand, general underreporting of both vote buying and disciplinary sanctions for lack of institutional trust or poor enforcement of laws against any type of corruption in some municipalities implies $cov(u, \upsilon) > 0$. Even if measurement errors are orthogonal to y, x and ε (classical measurement error case), the estimate would still be attenuated.

An instrumental variable regression that uses a valid instrument z ($cov(z, x) \neq 0$ and $cov(z, \varepsilon) = 0$), that is also uncorrelated with υ and u ($cov(z, \upsilon) = cov(z, u) = 0$) gives a consistent estimate of the effect of interest.

$$\hat{\beta}^{IV} = \frac{cov(\tilde{y}, z)}{cov(\tilde{x}, z)} = \frac{cov(x\beta + \varepsilon + u, z)}{cov(x + \upsilon, z)}$$

and

$$plim\ \hat{\beta}^{IV} = \beta\frac{cov(x, z)}{cov(x, z)} = \beta.$$

[1] Bound et al. (1994) present a general framework to study the linear model with variables with additive errors. This derivation is a particular case of their analysis.

D Fuzzy RD Assumption Checks

Since we have multiple discontinuity points, we carry out the sorting tests focusing on the distribution of municipalities in the sample according to their distance (in number of registered voters) from the discontinuities. The null hypothesis in these tests is that the density is continuous at the cutoff. The first test we carry out is proposed by Cattaneo, Jansson and Ma (2017).[2] Figure 1 shows that there is no statistically significant discontinuity in the density at zero. Moreover, we do not see a greater concentration of municipalities right above the cutoff as we would expect if politicians were trying to exploit the rule that determines the number of polling stations to their advantage. The test statistic is -0.77 with a p-value of 0.43. Similar results were found using the McCrary test (McCrary 2008). In that case, the log difference in the height of the density before and after the cutoff is -0.062 with a standard error of 0.152.

Table 2 explores whether there are discontinuities in the controls at the thresholds that determine additional polling stations. To test for discontinuities, we estimate the effect of having an additional polling station on all variables used as controls in the analysis. We see that none of the estimated effects are significant at conventional levels.

[2]Their proposed test uses a local polynomial approximation to the density that avoids estimation problems at boundary points when using standard kernel estimators.

Figure 1: Test of manipulation of the number of registered voters

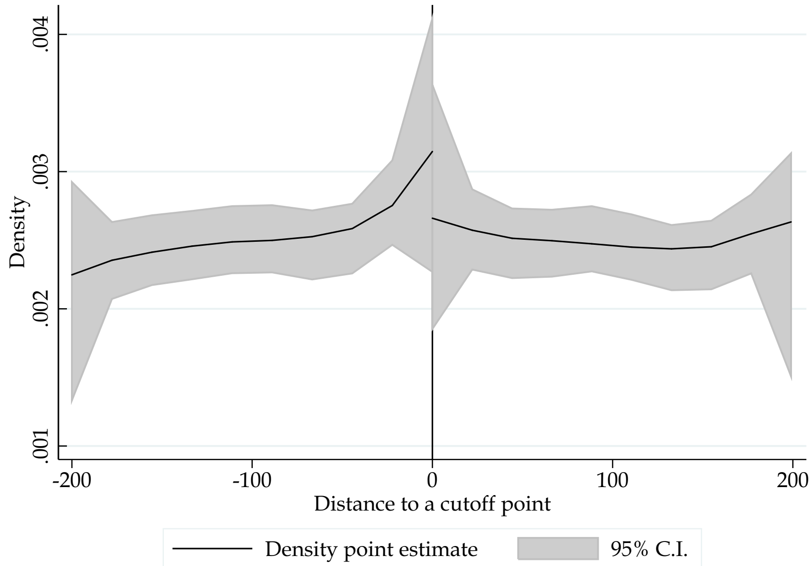


Table 2: Testing for discontinuities in controls

Dep. Variable:	Mean	Std. Dev	Coef. (RDD)	Std. Error.	Observations	Bandwidth	P-value
Panel A. Fiscal covariates							
Discretionary revenue	16,579.19	192,601	-41,866.52	30,433.374	479	41.367	.169
Local revenue (t-1)	11.916	12.025	-4.149	2.597	562	50.922	.11
Mayor's maximum salary	6.675	2.424	-.08	.45	710	65.62	.859
Panel B. Socieconomic variables							
Average Margin of victory	0.090	0.061	-.008	.013	630	53.935	.537
Armed group $(t-1)$	0.393	0.489	.08	.119	555	47.578	.505
Population $(t-1)$	37,760.6	229,473.4	-50,213.129	42,382.271	529	45.325	.236
Rural population $(t-1)$	0.593	0.240	-.014	.043	766	66.381	.741
Underperforming schools $(t-1)$	0.487	0.398	.1	.082	546	48.861	.222

Coef. (RDD) denotes estimates of the effect of adding one additional polling station. The results use Calonico, Cattaneo and Titiunik (2014) optimal bandwidths, bias correction, and robust standard errors, with linear local polynomials and triangular kernels.

E Transparency Index Results

The transparency index is formed by three main components. The first, which we'll call the *visibility* component, captures the degree to which the municipality administration facilitates citizen the access to information regarding the administration of public resources. The second, which we will call the *norms* component, measures the extent to which general budgeting norms and procedures are being followed by the municipality. The third component captures whether citizens are actively participating in the municipality budget design and planning and whether that participation is promoted by local officials.

Consistent with the theory, results in Table 3 show there is a negative association between vote buying and the index of transparency. Moreover, this association is driven by the visibility component, suggesting that in places where vote buying is common, public officials make it more difficult for citizens to monitor public finances. The coefficients on vote buying in the norms and participation indices models are also negative but not precisely estimated. An increase of one standard deviation in the number of vote buying reports is associated with a reduction in the visibility index of 3.4 units (a fifth of a standard deviation of the index). Although the coefficient is small, it is important to note that more transparency in public administration can push people to report more vote buying cases, and therefore, the estimates can be considered a lower bound of the true effect.[3]

Figure 2 presents the estimated relationship between discretionary revenue and the index of transparency and its components using Robinson's semi-parametric estimator. Baseline controls are included in all models. We see that the slopes are positive for low and intermediate levels of our value of office proxy, but for high office values the pattern is less

[3]For this cross section, the average size of polling stations is not a strong instrument for vote buying and the instrumental variables strategy does not give us reliable estimates of the effect of vote buying on the transparency indices.

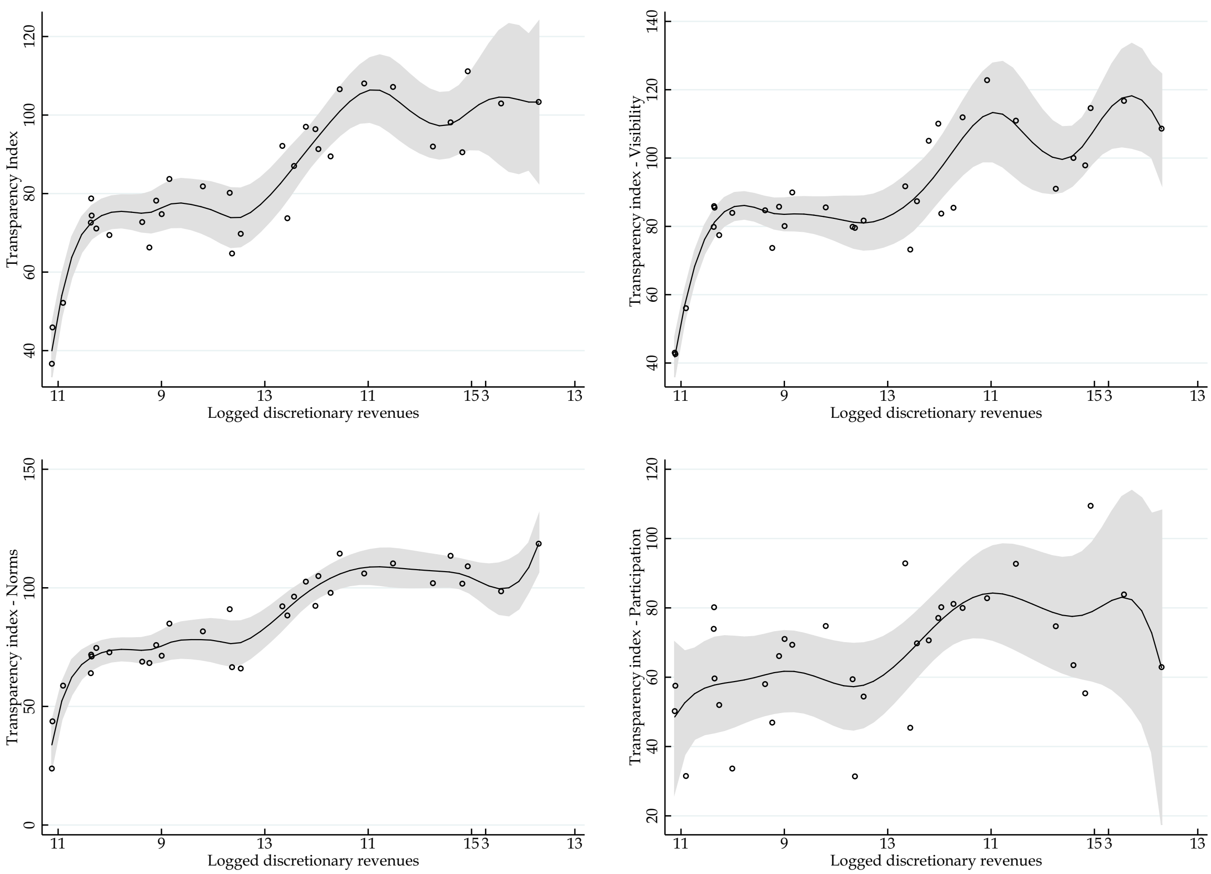
Table 3: Transparency and vote buying

Dep. Variable:	Visibility	Norms	Participation	Transparency
	(1)	(2)	(3)	(4)
Vote buying	-7.102***	-3.240	-4.034	-4.944**
	(2.270)	(2.290)	(2.993)	(1.962)
Observations	238	238	238	238

This table reports OLS coefficients. All models include baseline controls and an indicator of whether the mayor had previous disciplinary sanctions at the time of the election. Robust standard errors are in parentheses. *** p<0.01, **p<0.05, *p<0.1.

clear. In the norms case, where the prediction fits better the data the slope becomes negative. In this small cross section of municipalities for which there are reports of electoral manipulation, there does not seem to be a positive relationship between the value of office and transparency when the value of office is high.

Figure 2: Transparency and discretionary revenues



F Other Tables and Figures

Table 4: Summary statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Panel A: Variables of interest					
Prosecuted	2,072	0.242	0.429	0	1
Guilty	2,072	0.164	0.370	0	1
Removed	2,072	0.095	0.293	0	1
Transparency	252	56.54	14.32	17.59	88.15
Vote buying (reports per 1,000)	2,072	0.027	0.110	0	1.747
Discretionary revenue (number of minimum wages)	2,068	17,316	197,987	20	6,329,840
Mayor's salary (number of wages)	2,012	6.70	2.46	6	25
Panel B: Controls					
Armed actor	2,072	0.39	0.49	0	1
Education	2,072	0.47	0.4	0	1
Margin of victory	2,072	0.09	0.07	0.001	0.59
Own resources	2,072	12.06	12.09	0.01	78.86
Population	2,072	40,128	242,091	1,303	7,050,228
Polling station size (Rule)	2,072	387.84	13.13	303.25	400
Polling station size (Actual)	2,072	305.07	75.27	108.0455	940.6667
Registered voters	2,072	24,649	145,743	690	4'378,026
Rural population	2,072	0.58	0.24	0.002	0.98
Sanctions	2,072	0.13	0.33	0	1

Figure 3: Effect of vote buying transactions on sanctions at different bandwidths around discontinuities

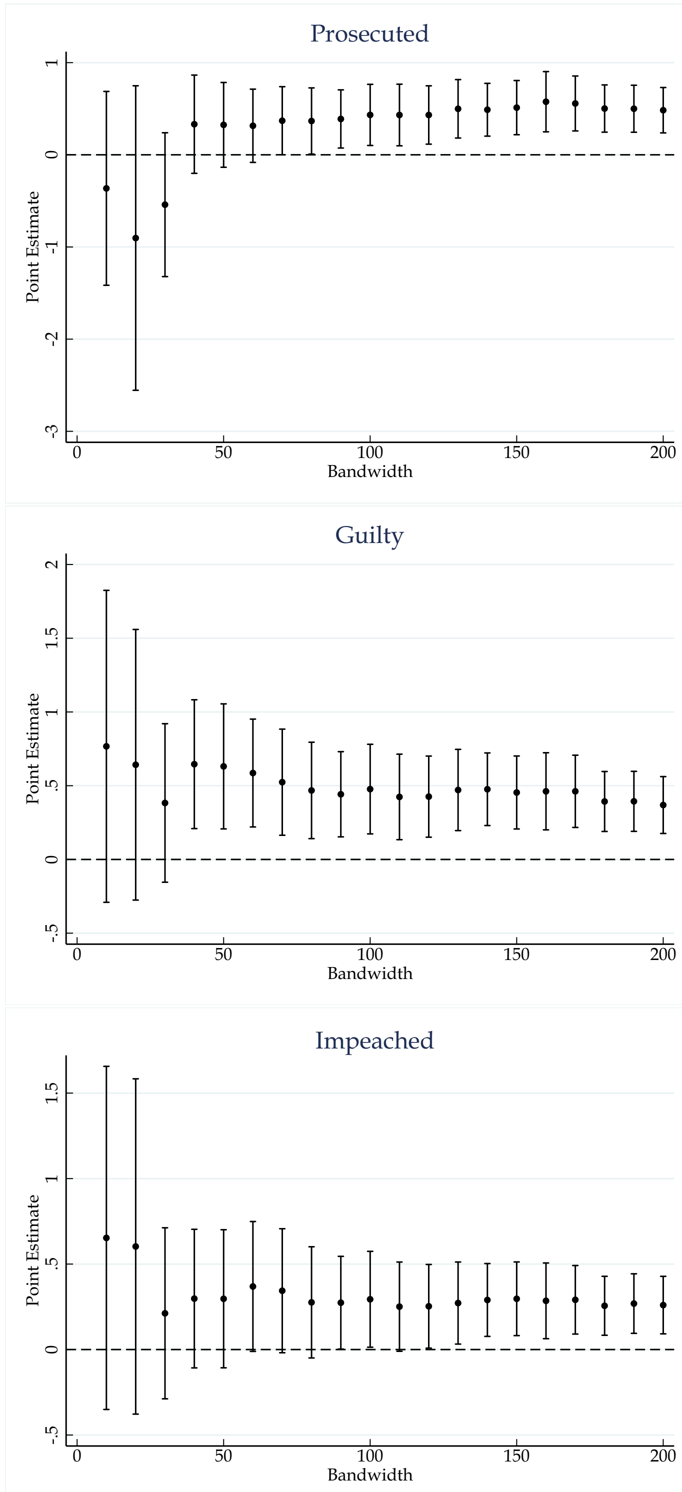


Figure 4: Disciplinary sanctions and value of office

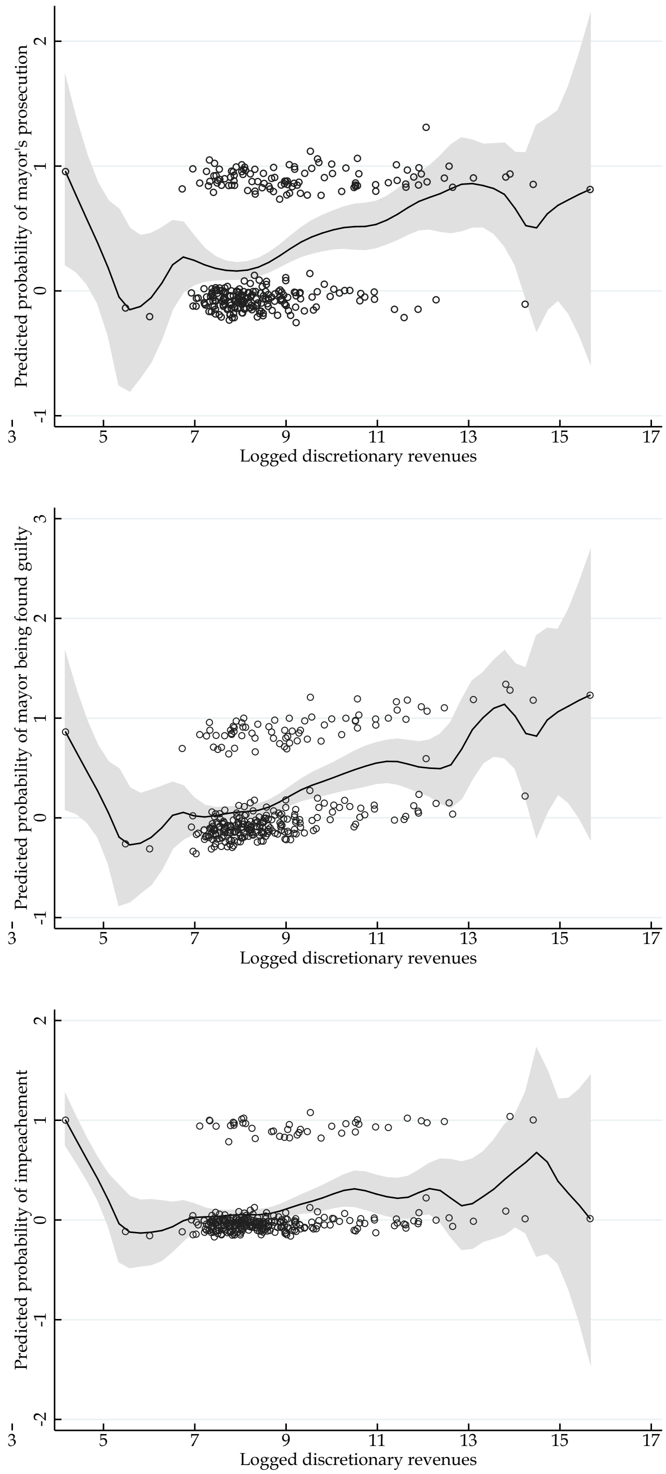


Table 5: Sanctions, vote buying, and value of office (non-linear models)

Dep. Variable:	Prosecuted	Guilty	Removed	Prosecuted	Guilty	Removed
	(1)	(2)	(3)	(4)	(5)	(6)
Vote buying	0.350**	0.344**	0.243			
	(0.145)	(0.150)	(0.176)			
Discretionary revenue				0.357	0.529	0.058
				(0.329)	(0.350)	(0.415)
Sample	Full	Full	Full	Vote buying	Vote buying	Vote buying
Observations	2,072	2,072	2,072	297	297	297
Municipalities	1,086	1,086	1,086	262	262	262

This table reports Logit coefficients. All models include baseline controls and an indicator of whether the mayor had previous disciplinary sanctions at the time of the election. The 'Vote buying' sample includes municipalities where there was at least one report of vote buying. Robust standard errors are in parentheses. *** p<0.01, **p<0.05, *p<0.1.

References

Bound, John, Charles Brown, Greg J. Duncan and Willard L. Rodgers. 1994. "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12(3):345–368.

Calonico, Sebastian, Matias D. Cattaneo and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs." *Econometrica* 82(6):2295 – 2326.

Cattaneo, Matias D., Michael Jansson and Xinwei Ma. 2017. "Simple Local Polynomial Density Estimators." University of Michigan, Working Paper.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics* 142(2):698 – 714.