Political Agency, Election Quality, and Corruption

Online Appendix

Miguel R. Rueda* Nelson A. Ruiz†

March 8, 2019

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

†Departmental Lecturer in Comparative Politics. Department of Politics and International Relations. University of Oxford. Email: nelson.ruiz@politics.ox.ac.uk

A Proofs

Proof of Proposition 1

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 the voters 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 $\theta \equiv E + \overline{r}$. When ξ is small enough, 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 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. All other features of the baseline model are kept. 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 in 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, one can show 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.

Publicly funded electoral manipulation

In this section, we allow public resources to be used for rents, public good provision, and financing electoral manipulation. That is, the budget constraint is $g_t = \eta(R - m - r_t)$. All other features of the model are as in the baseline model. Let m^e be the expected manipulation levels by the voters at the time of the election. This value does not depend on η, since voters do not know it at that point. An estimate of the incumbent's η is then $\frac{g_1}{(R - r_1^e - m^e)}$. Voters vote for the incumbent whenever

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

which happens with probability

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

The maximization problem of the incumbent is

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

The first order conditions are

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

and

$$c'(m^*) = \theta\chi - 1.$$

The second order sufficient condition for a maximum is still $2(1 - m^*)c''(m^*) > \theta\xi$. This holds when there is enough uncertainty about the ability of the incumbent to provide public goods.

There are three things to note with this extension. First, the level of manipulation is lower than that of the baseline setting as can be seen with the manipulation first order condition. This is intuitive, as larger manipulation levels reduce the amount of public goods that voters observe when the manipulation does not determine the election. This, in turn, negatively affects the incumbent's perceived ability to provide public goods.

The second is that the relationship between manipulation and rents is weaker than that of the baseline model. This is because, on the one hand, greater manipulation makes it more likely for the elections not to be determined by voters who care about public good provision, but on the other, manipulation directly reduces the perceived ability of the incumbent if elections are not influenced by it. As long as the value of office is large enough ($\theta\xi > 1$), the relationship between electoral manipulation and rents will still be positive.

Finally, the relationship between the value of office and rents is positive, but weaker, for high office values. To see this note that

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

Taking the limit of this partial derivative as θ approaches $\overline{\theta}$ (the value that makes $m^* = 1$), we get $\frac{\chi(\overline{\theta}\xi - 1)}{c''(1)}$, which is positive if $\overline{\theta}\xi > 1$.

Horizontal accountability

Here we extend the original analysis to allow for an incumbent who has been elected and who has engaged in electoral manipulation to be investigated and punished for such actions. We examined two plausible settings to study horizontal accountability. In the first one, we assume that only when manipulation has influenced the results of the election, the judiciary will punish the illegitimately re-elected incumbent by removing her from office with some probability. In particular, as in the baseline model, we assume that, with probability m, the electoral results are influenced by manipulation, in which case, this gives an electoral advantage to the incumbent (her probability of winning is $1/2+\chi$, instead of $1/2$ which would be the equilibrium probability of winning with no manipulation). When there are electoral irregularities and the incumbent wins, however, the judiciary will remove from office the incumbent with probability γ. Everything else remains as in the baseline model.

The first order condition for the optimal level of rents, as well as the second order conditions of the incumbent's maximization problem, are the same as in the baseline model. We find, however, that a necessary condition to have positive levels of manipulation in equilibrium is $(\chi(1-\gamma)-\gamma/2)>0$. Intuitively, strong judicial institutions (high γ), and small electoral advantages of successful manipulation (small χ), make this condition less likely to hold. We still observe a positive relationship between equilibrium levels of manipulation and rents, as it is the case that more manipulation reduces the chances of the electorate to freely vote on the bases of observed public goods. Finally, higher values of office induce larger rent accumulation for sufficiently large values of office. The strength of this association is weaker than in the baseline model, as the impact of value of office in manipulation is weaker. Although, a more valuable office induces the incumbent to engage in manipulation, the prospects of being caught and punished by the judiciary make this increase smaller.

The previous setting assumes that only when the electoral manipulation influenced the results and benefited the reelected incumbent, will she be punished with some probability. This assumption is motivated by the fact that prosecutors could have less interest in pursuing a case in which the manipulation efforts were not successful and did not end up benefiting the winner of the election. In a second setting we assume the judiciary punishes a reelected incumbent who engaged in manipulation with some probability, regardless of whether the manipulation ended up influencing the election outcome.

We assume that if elected, the incumbent gains the full value of being in office θ with probability $1-\gamma$. With probability γ, the reelected incumbent is punished by the judiciary for the previous electoral manipulation and receives as a payoff $\theta-m$. We want to capture with this assumption the fact that while minimum manipulation can reduce the legitimacy of an elected official, it might not represent a substantial loss for the reelected incumbent. If manipulation is large, however, the loses for the candidate increase. For example, while a few votes bought are hardly going to affect an elected official who is caught buying votes, an election where thousands of voters were bribed, possibly affecting the outcome of the election, can end the political career of the incumbent if prosecuted. Finally, we assume that the expected prize of winning is larger than that of losing.

The derivations in this case closely follow the ones in the baseline model. In an interior equilibrium

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

and

$$c'(m^*) = \chi(\theta - \gamma m^*) - \gamma(m^*\chi + 1/2).$$

As before, a second order condition for a maximum would be satisfied if there is enough uncertainty about the incumbent's type (ξ is small enough).

The equilibrium level of manipulation is lower than in the baseline model and is decreasing in the strength of the judicial institutions. It is still the case that rents are increasing in manipulation. To see this, note that in order to have an equilibrium with positive manipulation, $\chi\theta - \gamma/2$ must be strictly positive, or otherwise there is no solution for the second equation above in $(0, 1]$. The partial derivative of r^* with respect to m^* is $\xi(\theta + \gamma(1 - 2m^*))$. If it is negative, this would imply that $\frac{\gamma}{2\chi} < \theta < (2m^* - 1)\gamma$, but, since $\chi \in (0, 1/2)$, there is no θ that can satisfy these inequalities. A similar reasoning shows that the value of office has a positive effect on rent accumulation for large values of office.

Also note that if γ is low and θ is high (public office is highly valued and the judiciary is not able to prosecute and punish most violations of the law), which are conditions that characterize poor developing democracies, equilibrium manipulation will be high enough to make the relationship between rents and value of office and manipulation and value of office weaker than in the baseline model.

Vote buying as a form of manipulation

The baseline model captures how electoral accountability is affected by any form of electoral manipulation that has a non-deterministic effect on the outcome of the election, that is costly, and that, if effective, gives an advantage to the candidate that engages in it. The empirical application focuses on one type of electoral manipulation, vote buying. In this section we show that the baseline model results are compatible with a modified model that takes one element of vote buying that was not considered in the baseline model and previous extensions. This feature of vote buying is that there is a mass of voters that, when targeted, receive a material benefit that pushes them to vote for the candidate providing these benefits.

Let m be the effort incurred in a vote buying campaign. This effort is spent improving the offers to voters who are bribed. In particular, we assume that bribed voters receive a payment of $p(m)$ with $p(0) = 0$, $p' > 0$, and $p'' \leq 0$. The function captures the fact that larger offers have decreasing marginal benefits. Also, as more effort is spent in the vote buying campaign, more voters are targeted with bribes. A fraction m of voters is targeted, while $1-m$ are not. Those who are not targeted, vote on the basis of perceived incumbent's ability based on the observed provision of public goods. Bribed voters also receive an idiosyncratic shock favoring the incumbent $\iota \sim U\left[-\frac{1}{2}, \frac{1}{2}\right]$ and a general popularity shock $\delta \sim U\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$. All else remains as in the baseline model.

The incumbent will beat the randomly selected challenger if

$$m(p(m) + \delta) + (1-m)\xi\left[1 - \frac{R - r_1^e}{R - r_1}\right] > 0$$

which happens with probability $F\left(p(m) + \frac{1-m}{m}\xi\left[1 - \frac{R-r_1^e}{R-r_1}\right]\right)$ with F being the cumulative density function of a uniform with support in $\left[-\frac{1}{2\phi}, \frac{1}{2\phi}\right]$.

The first order conditions of the incumbent's optimization problem in an interior solution when there is enough uncertainty regarding the popularity shock are

$$r_1^* = R - \frac{1-m^*}{m^*}\xi\phi\theta$$

and

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

Having sufficient uncertainty about the ability of the incumbent guarantees that the second order conditions for a maximum hold. If $\theta\phi p'(1) > c'(1)$, there would be a corner solution with $m^* = 1$ and $r = \overline{r}$. If not, there is a θ, $\overline{\theta}$ such that $m^* = 1$. If we take the limit of $\frac{\partial r^*}{\partial \theta}$ as θ approaches $\overline{\theta}$, we can verify that rents are increasing in the value of office for high enough values of office.

Asymmetric learning, accountability, and electoral manipulation

In this section we explore an alternative model in which the types of politicians are known before the incumbent chooses policy and manipulation levels. We follow Besley (2007) by focusing on a simpler setting with discrete types and a discrete action space. As before, there are two periods (indexed by t) where there is an election at the end of period one. Politicians' type is denoted by $\eta \in \{0, 1\}$. If $\eta = 1$, the politician always chooses the policy that voters value the most. We call this type of politician a *public minded* politician. When $\eta = 0$, the politician might deviate from the policy most preferred by the voters and we call these type of politician *opportunistic*. At the beginning of the game, nature chooses the type of the incumbent, which will be public minded with probability π. The incumbent's type is her private information.

Once the incumbent has learned her type, she chooses between two policies, $e_t \in \{0, 1\}$, and whether to engage in electoral manipulation $m \in \{0, 1\}$ in the upcoming election (where $m = 1$ indicates the incumbent will engage in electoral manipulation). When the incumbent chooses policy $e_t = 1$, voters receive utility of $\Delta > 0$. If the incumbent chooses $e_t = 0$, voters receive a utility of 0. We can think of this last policy choice as capturing a situation in which the incumbent engages in administrative corruption. We assume voters have information on the policy chosen before the election, but they do not observe the choice of manipulation nor the incumbent's type. At the end of the first period, voters decide whether to vote for the incumbent or a challenger.

As for the payoffs of the incumbent, whenever she chooses $e_t = 0$, she receives a payoff of r_t, which is the realization of a random variable with support on $[0, \infty)$, cumulative distribution function, G, and expected value μ. The value of r_t is drawn by nature at the beginning of each period. In addition, a politician in office earns a term E, which captures other office rents.

If the incumbent chooses to manipulate the election, a fraction ω of the population will vote for the incumbent regardless of her policy choices. The rest of voters, a fraction of $1 - \omega$, will vote on the bases of the incumbent's policy choice. When the incumbent chooses to manipulate, she incurs in a cost $C > 0$. If the incumbent chooses not to manipulate the election, all voters are free to vote on the bases of observed policy outcomes.

Our goal is to find the conditions under which an opportunistic incumbent will choose not to engage in administrative corruption (choosing $e_1 = 0$). The equilibrium concept used is Perfect Bayesian Equilibrium, which in this case requires that the politician in office will behave optimally given the reelection rule of those voters who have freedom to choose in the two periods. Moreover, voters must use Baye's rule to update their beliefs regarding the incumbent's type.

An opportunistic politician in office in the second period will choose the socially sub-optimal policy $e_2 = 0$ while the public minded politician will choose $e_2 = 1$. Therefore, voters are better off if they reelect a public minded incumbent or choose to vote for the challenger if they believe they are dealing with an opportunistic incumbent. In period one, if voters receive Δ, the probability that they are dealing with a public minded politician is

$$\frac{\pi}{\pi + (1-\pi)\lambda},$$

where λ is the probability of an opportunistic type choosing the good policy. Such probability is larger or equal than π. Because of this, we focus in an equilibrium in which voters reelect the incumbent if and only if they receive Δ in the first period.

Under such a reelection rule, it is not profitable for an opportunistic incumbent to choose to manipulate while selecting the optimal policy. If she does that, she wins the election since everyone votes for her (the ω voters who are forced to vote for the incumbent and the rest who are following the reelection rule), but she incurs a cost C. Without manipulation, she wins the election outright as all voters follow the reelection rule and she does not incur the cost of manipulation.

Whether or not manipulating the election is optimal for an opportunistic incumbent while choosing $e_1 = 0$ depends on how costly and effective such manipulation is. We assume that $\omega \geq 1/2$, which allows manipulation to be used in equilibrium. There are two relevant cases to consider: when $C < (E + \mu)$ and when $C \geq (E + \mu)$. In the first case, manipulating while choosing $e_1 = 0$ is optimal as the benefits of being reelected by cheating in the election outweigh the manipulation costs. In the second case, manipulating the election is not worthwhile, even if it guarantees reelection.

A question that is left is whether winning the election by choosing the good policy and not cheating is better than winning by cheating and choosing the policy that voters dislike. This is the case whenever, $r \leq \min\{C, (E + \mu)\}$ and so $\lambda = G(\min\{C, (E + \mu)\})$.

The previous remarks indicate that, just as with the baseline model in the paper, there is not a strictly increasing relationship between the probability of politicians choosing a beneficial policy and the value of office. For low values of office ($(E + \mu) < C$), the relationship is increasing, as manipulation is still too costly to be used, but if reelection is valued enough ($(E+\mu) \geq C$), politicians will want to manipulate to get into office and higher values of office will not give more incentives to the politician to do what voters want. It is also easy to see that in the absence of electoral manipulation, λ will be strictly increasing in the value of office for all possible values of office, which indicates that we should be more likely to observe administrative corruption when there is electoral manipulation that is not too costly to implement. Finally, it is clear that in this model larger values of office are associated to electoral manipulation.

Fuzzy RD, selection effects, and relationship between costs of electoral manipulation and politicians' types

Note that only the opportunistic politicians engage in electoral manipulation in the asymmetric learning model. This is important for our empirical results as it raises the possibility that the observed positive association between corruption and vote buying uncovered in the OLS regressions is driven mainly by selection effects. However, what our fuzzy regression discontinuity results show is that, among opportunistic politicians, those who can engage in manipulation easily (those running in municipalities with small polling station on average) will be more likely to take the suboptimal policy. This requires that the proportion of opportunistic types is the same on both sides of the thresholds of registered voters that determine the discontinuous changes in polling station size. Given the difficulty to ex-ante assess the sizes of polling stations, and that our tests show no significant differences of mayors' characteristics around the thresholds, this seems a plausible assumption. That is, our fuzzy RD guarantees that the observed differences in corruption on both sides of the discontinuity are not driven by selection, but by differences in incentives to take the right policies determined by differences in costs of manipulation.

If there is a large positive correlation between C and η in the data (the opportunistic types tend to have some expertise in electoral manipulation), this would generate a strong positive correlation between administrative corruption and vote buying in the OLS results. If the correlation between C and η is weaker, the relationship between manipulation and vote buying would not be as strong.

To see this, it helps to think of the plane where the x-axis is manipulation (with $m \in \{0,1\}$) and the y-axis is the corrupt activity (with $e_1 \in \{0,1\}$). First, note that according to the model, there would not be municipalities located in $(1,0)$ as there are no politicians who manipulate and don't engage in corruption. Second, with a high positive correlation between C and η, most observations are going to be concentrated in $(0,0)$ and $(1,1)$, with a few in $(0,1)$. All the municipalities with public minded politicians are in $(0,0)$, as well as those where there are opportunistic types with high costs of manipulation whose random draw r was low. In $(0,1)$ we have municipalities where there are opportunistic types with a high cost of manipulation that got a large r. In $(1,1)$ there would be the municipalities with opportunistic types and low costs of manipulation, which are more numerous. As the correlation between C and η becomes less negative, there would be fewer municipalities in (1,1) and more in (0,1), making the regression line have smaller positive slope (the larger number of points in (0,1) pushes the y-intercept of the regression line to go upwards).

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 $\tilde{x} = x + \upsilon$. Further, suppose that $cov(x, \upsilon) = cov(x, \varepsilon) = cov(\upsilon, \varepsilon) = 0$. 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 actions 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, as shown here:

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

[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.

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

D Fuzzy RD Assumption Checks

Since we have multiple discontinuity points, we carry out the sorting tests by 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 do 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

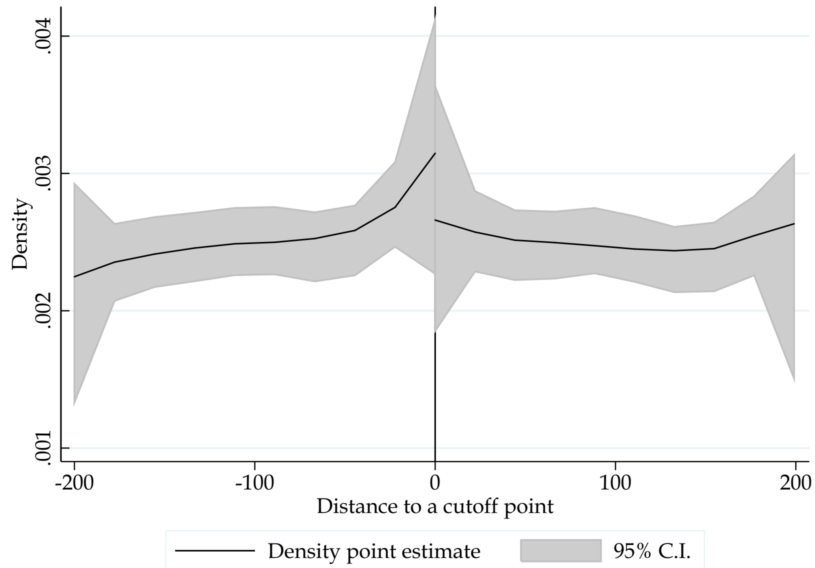


Table 2: Testing for discontinuities in controls

Dependent variable	Mean	Std. Dev.	Coef. (RDD)	Std. Error	Obs	Bandwidth	P-value
Panel A. Fiscal Covariates							
Discretionary revenue	16,579.194	192,601.014	-32,060.283	27,988.95	507	44.572	0.252
Local revenue (t-1)	11.916	12.025	-4.159	2.652	577	51.287	0.117
Mayor's salary	6.675	2.424	-0.079	0.459	722	66.172	0.864
Panel B. Socioeconomic Characteristics							
Avg. Margin of victory	0.090	0.061	-0.008	0.013	630	53.720	0.531
Armed group (t-1)	0.393	0.489	0.08	0.117	555	47.147	0.497
Population (t-1)	37,760.597	229,473.398	-38,015.403	41,643.496	581	49.478	0.361
Rural Population (t-1)	0.593	0.240	-0.014	0.043	766	66.242	0.748
Underperforming schools (t-1)	0.487	0.398	0.1	0.082	560	49.102	0.225
Panel C. Mayor characteristics							
Mayor held office before	0.178	0.383	-0.04	0.071	739	73.711	0.575
Mayor had sanctions before	0.120	0.325	-0.023	0.07	747	64.546	0.746
Mayor was involved in a lawsuit	0.026	0.159	0.033	0.049	543	53.150	0.509

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. Mayor was elected before refers to the same municipality.

E Future participation in elections

Table 3 presents marginal effects of vote buying on sanction indicators for models that include an interaction of vote buying with indicators of future political aspirations. The first of these indicators is a dummy variable that takes the value of 1 if the current mayor ends up running for the same municipality council or for mayor in the next possible election.[3] The second indicator takes the value of 1 if the current mayor ends up running for mayor of the same municipality at any point in the future, and finally, the third indicator is an index that is increasing in the overlap of the mayor's constituency and the constituency of mayoral, council, congress, department assemblies, and governorship races in which the mayor participated in the next possible election. In this last case, the index takes the population of the municipality of the mayor and divides it by the population of the district linked to the race where the mayor runs in the future. A value of zero indicates the candidate either does not run in any race in the future, or runs appealing to voters outside the municipality where she served as mayor. A value of 1 indicates the mayor runs in a race where the voters are those living in the same municipality where she served as mayor. Intermediate values indicate the mayor runs for a governorship, a seat in congress, or one in the department assembly.

We can see that the marginal effect of vote buying on sanctions is positive, but for those mayors that end up appealing to the voters they serve, the effect is larger and precisely estimated with all indicators. These patterns suggest the baseline findings are not driven by mayors who just get elected to extract rents and who do not expect to have a future political career.

[3]The next possible election is either: 1) the next regional election for non-mayoral races, 2) the next mayor election in other municipality, 3) the next national election, and if the mayor did not run in any of them, 4) the second future election of mayor in the same municipality.

Table 3: Marginal effects of vote buying on sanctions (future participation in elections)

Dep. Variable:		Prosecuted	Guilty	Removed
Runs in the same municipality	no	0.062	0.052	0.029
for council or mayor		(0.05)	(0.046)	(0.039)
in next possible election	yes	0.137*	0.166***	0.172***
		(0.079)	(0.064)	(0.058)
Runs in same municipality	no	0.061	0.057	0.034
for mayor		(0.05)	(0.047)	(0.04)
	yes	0.135*	0.14**	0.145**
		(0.076)	(0.062)	(0.056)
Index of constituency	0	0.072	0.062	0.03
overlap (all races)		(0.049)	(0.046)	(0.04)
in next possible election	1	0.097	0.123*	0.162***
		(0.084)	(0.069)	(0.057)

This table reports marginal effects of vote buying on sanctions. All models include baseline controls, an interaction of the variable of future participation in elections with vote buying and the variable of future participation in elections. Clustered standard errors at the municipality level are in parentheses. *** p<0.01, **p<0.05, *p<0.1.

F 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 access to information regarding the administration of public resources. The second component, 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 our theory, results in Table 4 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 public officials make it more difficult for citizens to monitor public finances in places where vote buying is common. 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.[4]

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

[4]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.

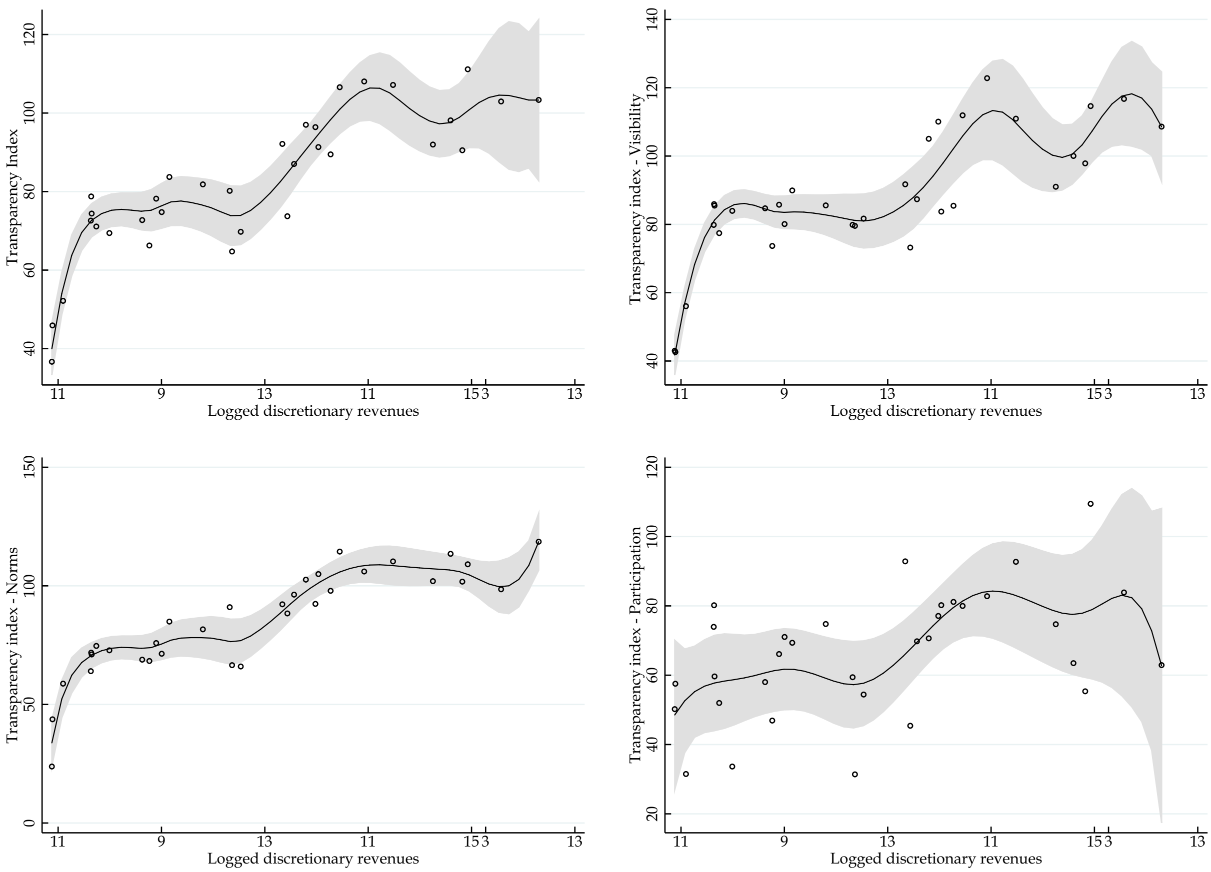
Table 4: Transparency and vote buying

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 better fits 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



G Other Tables and Figures

Table 5: 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

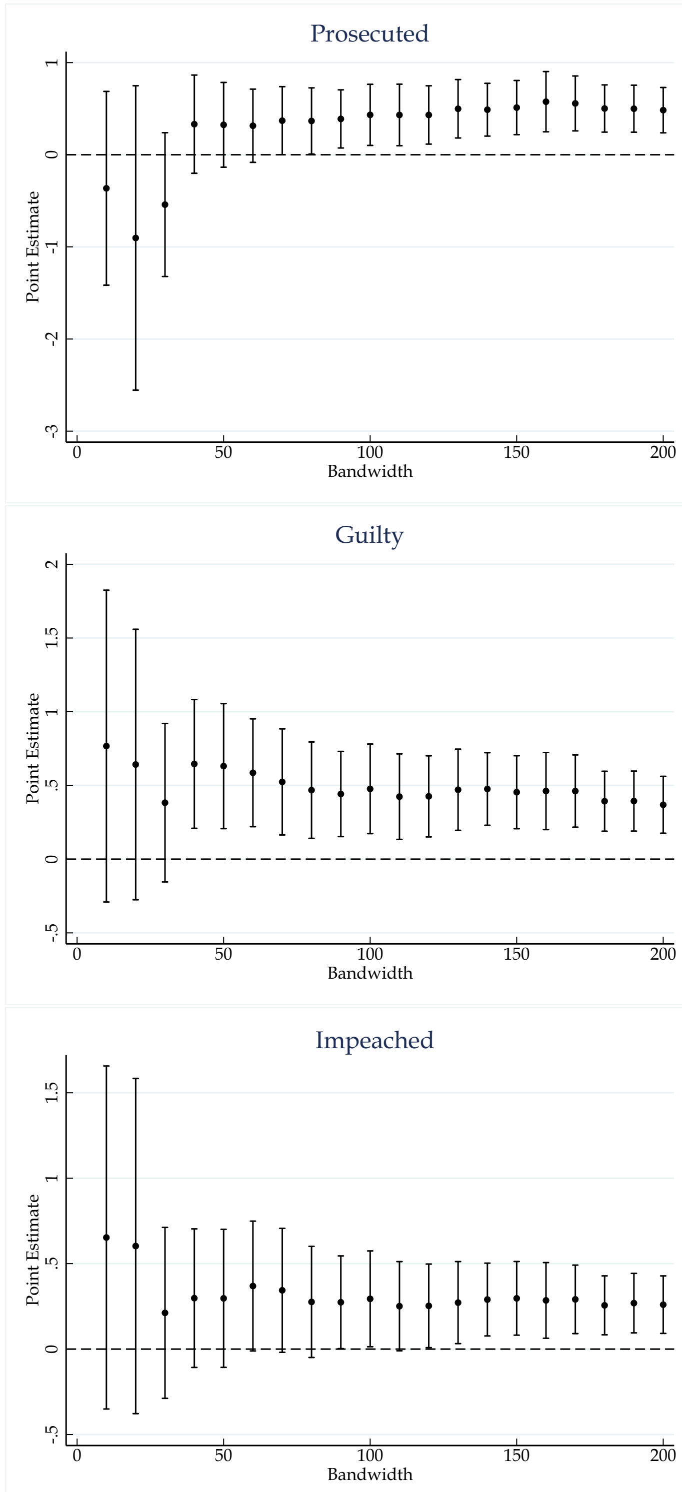


Figure 4: Disciplinary sanctions and value of office

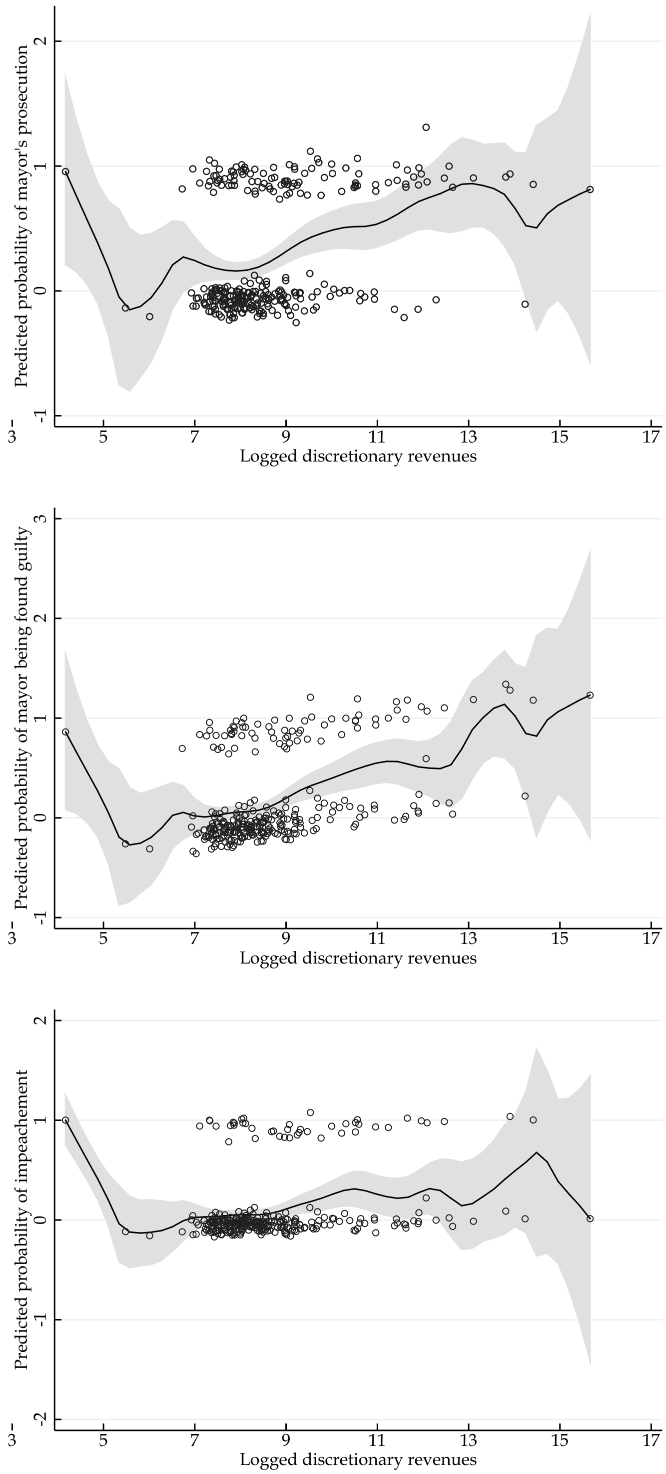


Table 6: Sanctions, vote buying, and value of office (non-linear models)

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

Besley, Timothy. 2007. *Principled Agents? The Political Economy of Good Government*. New York: NY: Oxford University Press.

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.