CSAE Working Paper  WPS/2016-18

Turning a blind eye: a Regression Discontinuity Design Analysis of Party-Based Support for Corruption in Brazil

Louis Graham

University of Oxford

Abstract

While corruption has long been conceptualised using the Principal-Agent framework, recent academic literature has proposed that in many countries, corruption is used as a political tool by ‘unprincipled principals’. Using data from a corruption audit programme in Brazil and an RDD design, I study whether a state governor and municipal mayor being of the same party increases corruption linked to the mayor. I find evidence that when the governor’s party wins the municipal mayoral election, corruption declines by 60-80% of the mean corruption level, and by 100-120% for larger municipalities. This is consistent with a model in which governors use corruption to gain control over mayors through potential blackmail, and in which governors want to prevent their party being associated with corruption to protect their re-election chances. I find further evidence consistent with the first argument. Much weaker evidence is found consistent with the second.

JEL Classification: C310, D720, D730

Key words: Corruption; Political Networks; Regression Discontinuity Design

1 louisgraham333@gmail.com
2 This paper was made possible by funding from the Economic and Social Research Council (ESRC). I am very grateful for the constant support given by Stefano Caria and the Oxford Department for International Development (ODID), without which this paper would not have been realised. Further, I am thankful for the support given by Damian Clarke and Simon Quinn when writing this paper. A version of this paper was submitted as an extended essay for the MSc in Economics for Development at Oxford
Introduction

Corruption, or the abuse of political office for private gain (Svensson 2005), is a particularly damaging phenomenon in developing countries today. It has been shown to lower a country’s GDP growth, increase its Gini index and poverty rate, and lower private sector productivity and profits (Ugur & Dasgupta 2011; DFID 2015; Fisman and Svensson 2007).

However, to date anti-corruption efforts have had “rather disappointing effects” (Kolstad et al. 2008: 59). One reason behind this is that the determinants of corruption are not clearly understood. In particular, the economic literature is limited by the continued analysis of corruption as a principal-agent problem following Becker & Stigler’s (1974) model of law enforcement (Persson et al. 2013). This analysis sees corruption as originating from self interested agents (politicians or bureaucrats) being entrusted to provide public services by publically-motivated principals (the public or more senior politicians) who cannot fully monitor their actions. While this analytical structure is useful in some circumstances, it relies on the existence of ‘principled principals’, whose only objective is to provide public services. Should this assumption fail, anti-corruption interventions based on this approach are likely to be ineffective. In this paper, I explore a different paradigm recently introduced to the economic literature; whether, and for what end, ‘unprincipled principals’ allow or encourage corruption among agents.

Using a Regression Discontinuity Design, and a data-set of municipal corruption audit results from Brazil collected by Brollo (2011; 2013), I test whether a state governor and municipal mayor being of the same party increases corruption linked to the mayor. Comparing municipalities in which the state governor’s party narrowly won the municipal mayoral election to those in which they narrowly lost, I find evidence that mayoral corruption is greater when the seat is held by parties in opposition to the governor. This increase is substantial (60-80% of the mean level of corruption), though it is not significant in all specifications. For a sub-sample of the top half of municipalities in terms of population, I find more robust evidence of this negative effect. For these municipalities, the increase in corruption is about 100-120% of the mean corruption level. This suggests that for large municipalities in particular, state
governors allow more corruption among opposition parties than their own party at the municipal level.

These results are then consistent with ‘unprincipled principals’ using corruption as a political tool. The direction of the results is surprising, as intuitively one might expect governors to allow more corruption among mayors with closer ties to themselves. However, they are consistent with a model I outline, in which ‘unprincipled’ state governors seek to gain control over opposition mayors with the threat of a corruption scandal, and in which governors fear their re-election chances may be harmed by the public associating their party with corruption.

Further, I directly test each of these pathways. I explore whether mayoral re-election chances are greater for mayors in the same party as the governor than those in opposition, conditional on their exhibiting the same amount of corruption. Using controls and a placebo test, I find that mayors in the same party as the governor are significantly more likely to be re-elected than opposition mayors with the same level of corruption. This is consistent with governors exposing corruption among opposition mayors for political gain. Weaker evidence is found to support the theory that governors prevent corruption in their own party to prevent the association of the party with corruption.

This paper contributes empirically and theoretically to the sparse economic literature on corruption as a political tool, rather than a principal-agent problem. To the best of my knowledge, it provides the first empirical test of the ‘principled principals’ assumption, and provides theory for what end principals use corruption. The political science literature on corruption is more advanced in this direction, suggesting that corruption is a political tool in countries with Neo-Patrimonial, or Clientelistic, political systems. In these systems, politicians rule through a system of personal connections and patronage, under a veneer of a modern, bureaucratic system (Von Soest 2013; Medard 1982). These systems are pervasive in several developing countries, including Brazil (Bach 2011). Here, corruption is a method of rule used by political leaders to maintain stability (Van Soest 2013) as a “particular strategy of rule or power consolidation” (Baland et al. 2010: 4615). Senior politicians allow or encourage corruption among juniors to buy off threats, reward supporters and blackmail opposition (Kolstad et al. 2008).
This approach has been slow to assimilate into the economic literature (DfID, 2015; Kolstad et al. 2008). Using quantitative data from Kenya and Uganda, Persson et al. (2013) detail the pressure to be corrupt placed on politicians and bureaucrats from various unprincipled principals (senior politicians, and the general public). They then argue that corruption is consequently better seen as a collective action problem. While everyone may be better off, and may even believe that they will be better off, without corruption in public life, everyone has an incentive to keep engaging in corruption personally. However, little empirical and theoretical work has been carried out beyond this.

This paper thus contributes to the understanding of the environment in which corruption takes place, and hence to knowledge of how best to structure interventions to tackle it. Evidence-based anti-corruption efforts have lacked understanding of the environment in which corruption takes place, and have broadly been ineffective (Persson et al. 2013). By focusing on the principal-agent framework, interventions have largely been based on changing the agent’s constraints; examples include creating formal monitoring institutions, and increasing wages for public servants. However, without tackling the underlying power dynamics behind corruption, auditing and other monitoring systems may be undermined or misused by unprincipled principals for their own benefit. Further, wage rises can be pocketed or passed on to bosses with no negative effect on corruption, as observed in Ghana (Foltz and Opoku-Agyemang 2015).

In addition, this paper continues the current movement to using objective data-sets for corruption, rather than corruption perception indices, which may not reflect actual corruption particularly well (Olken 2007; Banerjee et al. 2012; Baland et al. 2010).

This paper proceeds as follows: Section II gives background details into Brazil’s political system and history of corruption. Section III presents the theory of how corruption may be used as a political tool by ‘unprincipled principals, and presents a mathematical model to clarify the argument. Section IV presents the econometric strategy for the main results, section V the main results, section VII explores key mechanisms for the main results, and section VIII concludes.

---

3 Informal monitoring institutions, such as the Black Monday campaign in Uganda, are also used. This campaign monitors corruption in the Ugandan government, and encourages people to wear black every Monday to raise awareness of corruption.
II) Background

Brazilians Political System

Brazil’s political system is highly decentralised, with three separate systems of government; the federal government, 27 states (including the federal capital district), and over 5,500 municipalities. Municipalities, which range from a few hundred inhabitants to a few hundred-thousand, are responsible for a wide range of public services, using federal grants for healthcare, education and more (Brollo 2013). Each layer of government has an elected executive (the President of Brazil, the State Governor, and the Municipal Mayor), an elected legislature (National Congress, State Legislative Assembly and Municipal Council), and judiciary bodies (Power & Taylor 2011).

The executive in each is very powerful, and executive figures are particularly able to misuse state resources corruptly. For example, municipal mayors detail, propose and then implement out municipal budgets, with the legislature only in a position to veto items (Ferraz and Finan 2011). Consequently, this paper concentrates on the role of the executive in corruption. Further, governance in Brazil is traditionally dominated by coalitions; the plethora of parties means that ruling as a single party has usually proved impossible. As a result, this paper primarily compares mayors who are in the same party as the state governor, or have a coalition partner of the same party, against those who have neither.

Throughout its recent history, Brazil has suffered from regular corruption scandals at all layers of government. At the time of writing, the federal government is embroiled in several the ‘Petrolão’ or ‘Car Wash’ scandal. This involves the state owned oil company Petrobras, whose executives allegedly accepted bribes from sub-contractors, which were then channelled to politicians for vote buying and other purposes (Guardian 2015). A police investigation into this scandal has implicated senior politicians, including former Brazilian President Lula da Silva, and current President of the Lower House Eduardo Cunha, while 352 out of 594 members of Congress face accusations of criminality (Guardian 2015; Economist 2016). Corruption perceptions data backs up this picture of corruption, for example revealing that 70% of Brazilians believe corruption to be a ‘very serious problem’, while Brazilians rate the ethical standards of their politicians as worse than in 135 other countries (Carson and Prado 2014).
III) Theory and Model

As has been argued, corruption may in many circumstances be a political tool rather than simply the result of a principal-agent problem. This section first asks whether governors have the power to affect mayoral corruption, and then discusses whether they would want to allow more or less corruption among mayors from their own party. It then sets out a mathematical model to capture several of the key points discussed.

Can Governors Affect Mayors’ Corruption Decisions?

Of course, whether or not senior politicians benefit from corruption among junior politicians is irrelevant if they cannot affect their decisions. Qualitative data from Brazil certainly suggests that governors are able to alter the environment in which mayors make corruption decisions. Macauley argues that state governors hold power over both state and municipal police institutions (which investigate most municipal corruption), and state and municipal judicial institutions (which prosecute municipal corruption) (Power and Taylor 2011). Through formal and informal methods, it is likely that governors can place some mayors in a position in which they have nothing to fear from acting corruptly, and some in a more precarious position.

Will Corruption Be Higher or Lower with Political Ties?

While the hypothesis that governors may want to affect mayors’ corruption decisions is clear in the literature, whether they will want to allow more or less corruption among mayors of their own party is less so.

Initial intuition, and much of the political literature, suggests that state governors should allow or encourage mayors from their own party to be more corrupt than those from other parties. The essence of Neo-Patrimonialism and Clientelism suggests that governors would be more likely to favour mayors in their own party, as clientelistic and party networks are likely to have significant overlap. Hence governors are in effect mayors’ patrons, and trade ‘gifts’ for mayoral support in “bringing in a particular sectoral or territorial vote” (Power and Taylor 2011: 224). Further, Macauley argues that within Brazilian parties, politicians support each other as they ascend the political ladder in exchange for rewards from those who attain political seniority.

---

4 In particular the political science literature
However, I propose two factors which suggest the opposite effect. As will be tested empirically, these two reasons for decreased corruption in mayors of the governor’s party may dominate those more common theories which suggest an increase.

Firstly, whilst allowing mayoral corruption can be a way to reward supporters for their support, it can also be used as a tool to control mayors, and prevent their opposing the governors’ private or political agendas. As Baland et al. state, “bureaucrats in patrimonial regimes are even encouraged to be corrupt and steal, perhaps because this gives patrons move leverage over them” (Baland et al. 2010: 4613). As governors have direct control over those in their own party and coalition partners (using internal party disciplinary measures), but not over opposition parties, they may thus wish to allow opposition mayors to be more corrupt in order to gain better control over them. Importantly, one might expect that this effect will be particularly strong for municipalities with larger populations, as their mayors are likely to have more political clout.

Secondly, it is well established that the electorate expresses a distaste for corruption in Brazil (Ferraz and Finan 2008) and elsewhere. Should voters come to associate a party with corruption, then that party’s gubernatorial candidates may be harmed. Hence governors may have an incentive to prevent mayors in their own parties from acting corruptly, and to allow greater corruption among the opposition.

Model

This section presents a simple variation on the principal-agent model, which extends the analysis to include unprincipled principals with a vested interest in the agent’s corruption. The objective of this model is to capture the above arguments, and the ambiguous prediction that they give.

In this one period model, agents (mayors) are each paid a wage ($\bar{w}$) at the end of the period, which may differ from mayor to mayor, and are in charge of a budget ($\bar{F}$), from which they can extract some amount as corruption ($c$, where $0 \leq c \leq \bar{F}$). However, they have a probability of being caught which increases with the amount they steal ($p(c, \alpha)$), where $p'(c) > 0$, in which case they will be fired, and will receive a reservation wage ($\bar{R}$, assumed to equal 0 for simplicity) and no corruption spoils. Importantly, the principal can affect this probability function through choosing $\alpha$ ($p'(\alpha) > 0$), and hence
\( \alpha \) reflects the strength of the monitoring put in place by the principal. A mathematically simple function for the probability of being caught is as follows:

\[
(\text{I}) \quad p(c, \alpha) = \frac{\alpha c}{\bar{F}}, \text{ where } 0 \leq \alpha \leq 1
\]

It follows that if alpha is set equal to 1 (maximum monitoring), the mayor would have the same probability of being caught as the proportion of funds they appropriate. This provides the useful properties of mathematical simplicity and not giving probabilities greater than 1, and provides the intuition important for this model. The agent then chooses corruption to maximise their expected (CRRA) utility:

\[
(\text{II}) \quad \max_c E[U(c, \bar{w}, \bar{F})] = \frac{ac}{\bar{F}} (0)^\rho + (1 - \frac{ac}{\bar{F}}) (\bar{w} + c)^\rho
\]

Where \( u(x) = x^\rho \) and \( 0 < \rho < 1 \) (the agent is risk averse). Solving for the optimal value of \( c \), one can then easily prove the following conditions:

\[
(\text{III}) \quad c^* = \frac{\rho \bar{F}}{(\rho + 1)\alpha} - \frac{\bar{w}}{\rho + 1}
\]

\[
(\text{IV}) \quad \frac{dc^*}{d\alpha} = -\frac{\rho \bar{F}}{(\rho + 1)\alpha^2} < 0, \quad \frac{dc^*}{d\bar{w}} = -\frac{1}{\rho + 1} < 0, \quad \frac{dc^*}{d\rho} = \frac{\bar{F} + \alpha \bar{w}}{\alpha(\rho + 1)^2} > 0,
\]

The first differential in (IV) is most important here; it shows that as monitoring (\( \alpha \)) increases, corruption falls. The model also demonstrates the other usual properties of a principal-agent model of corruption; corruption decreases with the wage, and increases with risk aversion (the lower \( \rho \) is, the more risk averse the agent is).

The principal/governor can therefore influence mayors’ corruption levels through the parameter \( \alpha \). Were they ‘principled’, they would face a trade-off of stopping corruption against the costs of monitoring. However, in this model principals instead trade-off the benefits that they achieve from mayoral corruption against the costs that corruption imposes on them (the cost of monitoring is assumed to be zero for simplicity). I assume that the principal has the following additive utility function:

\[
(\text{V}) \quad V(\bar{D}, I(c), c) = U(\bar{D} + I(c)) - (1 + \beta \bar{D})c, \quad \beta \bar{D} > -1
\]
Where $\bar{D}$ represents direct, formal control over the mayor, and $I(c)$ informal control gained from knowledge of corruption, which is increasing in corruption ($I'(c) > 0$). It is assumed that direct control ($\bar{D}$) is larger over mayors of the same party (using the formal party apparatus). Principals also lose utility from the corruption of mayors below them, though the degree of this effect depends on the parameter $\beta$, which may be positive or negative, capturing the joint effect of the desire to reward/bribe supporters (implying a negative $\beta$), and the principal’s desire to protect their own re-election chances (implying a positive $\beta$).

Assuming that $U(.)$ takes a similar CRRA form to the agent’s:

$$ U(x) = x^\tau $$

Where $0 < \tau < 1$, the principal then has concave gains in corruption, and linear losses in corruption. This is justifiable; it is intuitive that the costs of corruption will be linear or even convex (small amounts of corruption may be unnoticed by the public or unappreciated by mayors), while the benefits of control will have diminishing returns. By making a simplifying assumption that $I(c) = c$, and substituting (III) into (VI), one can solve for the optimal $\alpha$:

$$ \alpha^* = \frac{\rho F^\theta}{(\rho + 1)((1 + \beta \bar{D})^\theta - \bar{D}^\theta + \bar{\omega})^{\theta^2}} \quad \text{where} \quad \theta = \frac{1}{\tau - 1} \quad \text{and so} \quad \theta < -1 $$

From this formula, one can then see how the optimal choice of monitoring changes with the level of direct control over the mayor:

$$ \frac{d\alpha^*}{d\bar{D}} = \frac{(-\rho F^\theta)}{(\rho + 1)((1 + \beta \bar{D})^\theta - \bar{D}^\theta + \bar{\omega})^{\theta^2}} \times \left( \theta \beta (\rho + 1)(1 + \beta \bar{D})^{\theta - 1} - (\rho + 1)\tau^\theta \right) > 0 $$

The first term of this formula is clearly negative. If $\beta > 0$ (and so protecting re-election outweighs rewarding supporters) then the second term is also negative, and so as direct control over a mayor rises, the governor will increase monitoring, and thus decrease corruption in that municipality. However, if $\beta < 0$ (so rewarding supporters outweighs protecting re-election), then the second term, and hence the overall result, is ambiguous.

---

5 In this model, for simplicity, indirect control does not enter the utility function of the agent, i.e. they do not internalise this externality. One might expect that it should enter the agent’s utility function negatively.

6 Remembering $\theta < -1$
as the benefits of increasing control over the opposition is weighed against rewarding one’s supporters.

A simple extension to this model introduces the intuition that having control over a mayor will be more important for the governor if the mayor is head of a larger municipality (they will likely have bigger political clout). To include this, the principal’s utility function (VI) changes to:

\[
V(\bar{H}, \bar{D}, I(c), c) = U(\bar{H}(\bar{D} + I(c))) - (1 + \beta \bar{D})c
\]

Where \(\bar{H}\) is the size of the municipality. Following the same logic as before, one can show that\(^7\):

\[
\frac{d\alpha^*}{d\bar{H}} < 0
\]

This then says that the larger the municipality, the more a governor will be willing to allow corruption in order to gain control over a mayor.

**IV) Data**

This section discusses the datasets used for the empirical analysis, and their sources.

*Corruption Data*

This paper uses a dataset on corruption at the municipal level in Brazil for its dependent variables, collected by Brollo (2011; 2013). This is an expanded version of that originally collected by Ferraz and Finan (2008; 2011). This dataset collates the results of a random auditing process for municipalities initialised by the government of President Lula da Silva in 2003, and run through the Federal Controladoria Geral da União (CGU). Several auditing rounds took place every year from 2003, each of which selected a random sample of municipalities with less than 450,000 inhabitants\(^8\). Those selected received an in-depth, on the ground audit into the use of funds transferred from the federal to municipal government\(^9\) from 2001 (the beginning of the 2001-2004

---

\(^7\) One can show that \(\alpha^* = \frac{1}{(\rho + 1)(\rho + 1)\beta^2} \frac{\bar{H} + \bar{D} - \bar{D}}{\bar{F} + \bar{F}}\), where \(\gamma = \frac{r}{r-1} < 0\). As \(\bar{H}\) increases, \(\bar{H}^r\) thus decreases, increasing the denominator and decreasing \(\alpha^*\).

\(^8\) I.e. around 99% of Brazil’s 5,600 municipalities

\(^9\) This constitutes the vast majority of the municipal government’s revenue (Brollo, 2013, calculated that for municipalities with under 50,000 inhabitants only 6% of the budget is raised by local taxes. This pattern is broadly similar for larger municipalities)
mayoral and councillor term) until the date of audit. This dataset calculates the proportion of transferred resources involved in corrupt activities in each mayoral term. This paper uses only data on the 2001-2004 term. Two definitions of corruption are used: narrow corruption and broad corruption, defined respectively as “severe irregularities” and those plus “irregularities that could also be interpreted as bad administration” (Brollo 2013: 23). This dataset was collected for the first 29 rounds (carried out from 2003 until 2009), and contains data for 1309 municipalities.

As argued by Ferraz and Finan (2008; 2011), while this data is subject to certain issues, it provides an excellent objective measure of corruption. Municipalities are chosen by a transparent (carried out in front of the media) and random lottery process, and hence it is unlikely that municipalities with political ties to the federal or state government were more likely to avoid selection (Ferraz and Finan 2011). Further, while it is plausible that corrupt municipal politicians could bribe auditors for a favourable report, auditors were highly paid federal employees with little to gain from collusion, and there is no empirical evidence of this taking place (Ferraz and Finan 2011).

Further corruption variables are binary indicators for if there is any broad of narrow corruption at all, and the total amount of resources that are involved in each type of corruption. I calculate the latter using the fraction of total resources involved in each type of corruption and the total federal transfers to each municipality (sourced from IPEA, as described later).

Elections and Party Data

This paper uses data collected by the Tribunal Superior Eleitoral (TSE), the senior judicial body in Brazil, to compose key independent variables regarding party affiliation and margins of victory. This dataset contains the results for each candidate in the 2000 municipal mayoral elections (those for the 2001-2004 term), plus their party and their affiliated coalition. Following Dell (2015), only municipalities in which the governor’s party came first or second in the municipal mayoral election (directly, or as part of the mayor’s coalition) were included. For these, a dummy variable was created with the value 1 if the governor’s party won, and zero otherwise. A continuous variable (the running variable for the below RDD analysis) was created to capture the win margin of the governor’s party (with a positive value if they won, and a negative value if they lost). Similar variables were created only looking at a direct party link (and so counting
mayors whose coalition partners were of the governor’s party as being in the opposition). However, I focus on the former variable for two reasons; firstly, restricting the treatment to only those with a direct party link reduces the sample size drastically to only 345, which prevents high-powered analysis, and secondly, coalitions form an important part of the political process in Brazil, and it is likely that ties within coalitions are of similar importance as those within parties.

Finally, as gubernatorial and mayoral terms are staggered, municipalities in which the state governor was not re-elected in 2002, midway through the mayoral term, were dropped from the dataset.

Further Data

The majority of the other variables used by this paper are collected from the Instituto de Pesquisa Econômica Aplicada (IPEA), a government-led organisation, or data collected by Ferraz and Finan (2008, 2011).

V) Econometric Strategy

This section first sets out the Regression Discontinuity Design strategy, using the framework of the Rubin Causal Model (Rubin 1974). It then discusses practicalities for this econometric technique, and lays out this paper’s primary analysis strategy. It concludes with a discussion of methods used for clustering standard errors at the state level, given the small number of states in the dataset.

The Regression Discontinuity Design

Under the Rubin Causal Model, each municipality has two potential corruption outcomes (y), depending on whether the state governor’s party wins the mayoral election or not (the treatment, w). As each individual receives only one value of treatment in reality, only one potential outcome is seen:

\[ y_i = \begin{cases} y_{1i} & \text{if } w_i = 1 \\ y_{0i} & \text{if } w_i = 0 \end{cases} \]

The parameter of interest is the Average Treatment Effect (the ATE):

\[ E[y_{1i} - y_{0i}] \]
For a multivariate regression of corruption on a same party dummy and control variables to yield the ATE, it must be the case that the treatment (whether the governor’s party wins the mayoral election) is assigned independently of potential corruption outcomes, conditional on the control variables. This is called the conditional unconfoundedness assumption. The dataset available for this paper does not include several variables which may affect corruption outcomes (for example mayoral pay, which is not publically available, or the presence of local monitoring bodies). It may then be possible that municipalities whose mayor is in the same party as the governor have systematically different potential outcomes, biasing the estimate of the ATE.

The (strict\textsuperscript{11}) Regression Discontinuity Design technique allows for a relaxation of the conditional unconfoundedness assumption. It utilises situations in which treatment is assigned only to those units with more than a cut-off value of some covariate (the running variable). As Imbens and Lemieux note, while the running variable may have a relationship with the outcome variable, “this association is assumed to be smooth, and so any discontinuity of ... the outcome as a function of this covariate at the cut-off value is interpreted as evidence of a causal effect of the treatment” (Imbens and Lemieux 2007: 3). One only then requires the assumption of continuity in potential outcomes. Informally, this means that potential outcomes must not experience a discontinuity at the cut-off (or, less importantly, anywhere else). Formally (Imbens and Lemieux 2007: 5-6);

\begin{enumerate}
\item \(E[y_1|X = x]\) and \(E[y_0|X = x]\) are continuous in \(x\), and:
\item \(\text{Pr}(y_1 \leq y|X = x)\) and \(\text{Pr}(y_0 \leq y|X = x)\) are continuous in \(x\) for all \(y\).
\end{enumerate}

One can then calculate the Average Treatment Effect for units with the running variable approximately equal to the cut-off value:

\begin{equation}
ATE(\epsilon) = E[y_{1\epsilon} - y_{0\epsilon}|x = \epsilon] = \lim_{x \rightarrow \epsilon} E[y_1|x > \epsilon] - \lim_{x \rightarrow \epsilon} E[y_1|x < \epsilon]
\end{equation}

For this paper, the running variable is the margin by which the same the governor’s party wins the municipal election (a positive value if they win, and a negative value if

\textsuperscript{10} Two other assumptions are also required (SUTVA and overlap), though they are less relevant to this discussion

\textsuperscript{11} Strict RDDs, in which treatment=1 above a cut-off and 0 below, differ to fuzzy RDDs, in which the probability of treatment experiences a jump at some value of the running variable. Here a strict RDD is relevant, and the following empirical strategy is that relevant for this design.
they lose). This is similar to Dell (2015). Clearly, there is a cut-off at zero; municipalities with a running variable with value greater than zero by definition have a mayor of the same party as the governor (or with a coalition partner of the same party). Hence treatment jumps from zero to one at the cut-off. Further, it is likely that potential outcomes will not jump at the cut-off; other than the difference in elected party, the municipalities which just elected the same party as the governor can be expected to be very similar to those which just failed to elect them. By estimating the expected value of the outcome variable (the proportion of funds involved in narrow or broadly defined corruption) just either side of the cut-off, one can then estimate the effect of a mayor/ mayor’s coalition partner being in the same party as the governor on municipal corruption.

**RDD Practicalities**

Consequently, the aim of an RDD is to estimate the expected outcome value just either side of the cut-off, and calculate the difference between them. As there are no data-points within an infinitesimally small region around the cut-off, in practice this is calculated by running a separate regression for data on either side of the bandwidth. This then leads to two interrelated issues; functional form and bandwidth. One must choose the functional form of the regression on either side, and results can be highly sensitive to the choice made; for example, using a linear regression either side of the cut-off when the true relationship between the running variable and outcome variable is quadratic may cause a type 1 error. As Lee and Lemieux state, “The consequences of using an incorrect functional form are more serious in the case of RD designs [than other specifications]” (Lee and Lemieux 2010: 36). To help reduce the importance of functional form, one can estimate the relationship non-parametrically. At its most simple (the rectangular kernel method), this entails restricting data to that within a certain ‘bandwidth’ around the cut-off, and running ‘local linear’ or ‘local quadratic’ regressions either side. A slightly more complex version uses triangular kernels, which give more weight to observations closer to the cut-off. Local linear/ quadratic kernel regressions have become the dominant method for RDD analysis (see Dell 2015; Fafchamps & Labonne 2015; Brollo et al. 2013), and have been shown to be give an optimal rate of convergence, thus limiting bias in estimates (Porter 2003).
One must then choose a) The bandwidth, b) The local functional form choice, and c) the kernel method. Bandwidth selection involves a trade-off between bias through incorrectly chosen functional form (which increases with the bandwidth size) and variance of the estimates (which decreases with the bandwidth size due to an increased sample size). Imbens and Kalyanaraman (2010) provide a method which uses the data to make this trade-off optimally for a local linear specification (henceforth the I-K bandwidth)\(^\text{12}\), providing a bandwidth with ‘asymptotic no-regret’ optimality. Calonico, Cattaneo, and Titiunik (2014) provide a further method for a local quadratic specification (henceforth the CCT bandwidth)\(^\text{13}\). However, one should still be concerned if results are particularly sensitive to bandwidth choice, and hence robustness analysis requires running using other bandwidths (in practice, these are traditionally half and double the optimal bandwidth, and other intuitively sensible ones). Regarding the local functional form choice, Gelman and Imbens (2014) suggest using local linear or quadratic methods, and not those of higher order (as results based on higher order polynomials are sensitive to polynomial choice, and have poor inference characteristics). Regarding the kernel method (rectangular, triangle or other), if estimates are highly sensitive to kernel choice, they are likely highly sensitive to bandwidth choice, and hence estimates “are not very credible anyway” (Imbens and Lemieux 2007: 16). Consequently any can be used, with other methods used as robustness checks.

Primary Analysis Plan

For my headline results, I will estimate the following local linear (2) and local quadratic (3) regressions, limiting the data to that within the I-K and CCT bandwidths respectively:

\[
(2) \quad y_m = \beta_0 + \beta_1 \text{govpartywin}_m + \beta_2 \text{govpartywinmargin}_m + \beta_3 (\text{govpartywin}_m * \text{govpartywinmargin}_m) + \epsilon_m
\]

\(^{12}\) Embedded in the \texttt{rd} Stata command

\(^{13}\) Embedded in the \texttt{rdrobust} Stata command
Where govpartywin is a dummy equal to 1 if the winning mayoral candidate, or one of their coalition partners, is of the same party as the governor, and govpartywinmargin is the running variable. Standard errors will be clustered at the state level (as discussed in the following sub-section). This in effect runs separate linear or quadratic regressions on either side of the cut-off (the rectangular kernel technique), with $\beta_1$ and $\pi_1$ estimating the treatment effect. Each will be run with two separate outcome variables (the fraction of resources involved in broad corruption, and the fraction of resources involved in narrow corruption), with the optimal bandwidth calculated individually for each case. Further, I will report the same four results for municipalities in the bottom half of the population in the distribution, and for municipalities in the top half. This allows me to test for heterogeneity in treatment effects for different sized municipalities, as is predicted by the model. For comparison, and to demonstrate the motivation for the RDD technique, I will also present a naive bivariate OLS regression of both corruption variables on the samepartygov dummy.

Cluster Robust Standard Errors

The outcome variables in this paper (broad and narrow corruption fractions) are likely to be correlated at the state level, as municipalities within a state share the same governor, and under the hypothesis provided by this paper the governor can affect municipal corruption. As Cameron and Miller (2015) state, the usual technique to deal with this is to implement a variation on White’s heteroskedasticity –robust standard errors (White 1980), designed by Liang and Zeger (1986). However, using this approach with few clusters, and especially with unbalanced clusters can result in rejection rates which vastly exceed those assigned to critical values (Cameron et al. 2008). Two approaches can be used to tackle this issue; wild-bootstrapping (Cameron et al. 2008) and cluster-adjusted t-statistics (CATs). While the former has been more common, Esarey and Menger (2016) argue that the latter, developed by Ibragimov and Muller (2010; 2016), gives lower rates of false positives and better power, especially with unbalanced clusters. This approach involves running separate regressions within each cluster and using the distribution of the mean of the cluster coefficients to compute
a standard error and p-value. It requires sufficient variation in the dependent variable, and sufficient sample size, in each cluster. Should a cluster lack variation and sample size, it is dropped for the analysis. In this paper I will use the higher powered CAT approach wherever possible. However, when more than two clusters lack sufficient variation/ sample size, I will use the wild bootstrapping approach\textsuperscript{14}.

For each regression, I will report which method has been used, the coefficient, and the p value. Further, when using the CAT approach, I will also report another coefficient generated in the analysis, which is the mean of the coefficients from running regressions in each cluster. Emphasis is on the original coefficient, as it weights each municipality equally.

\textbf{VI) Results: Primary Analysis}

This section first outlines some summary statistics of interest, before presenting and describing preliminary checks for the RDD, graphic analysis of the primary results, and estimates and robustness checks for the primary analysis.

\textit{Summary statistics}

The final dataset contains 541 observations. The mean percentage of federally transferred funds involved in corruption was 4.97\% for the broad definition, and 2.48\% for the narrow definition. 292 municipalities engage in broadly defined corruption, while 170 engage in corruption defined narrowly. In 285 municipalities (just over half), the state governor’s party won the 2000 mayoral election. These numbers are similar to the few estimates of corruption prevalence in the literature; Auriol (2006) estimates bribery in the procurement of government goods and services (an element of the overall corruption calculated here) to value 3.5\% of global procurement spending, while Svensson (2003) finds firm bribes to amount to 8\% of total firm costs in Uganda.

\textit{Preliminary RDD Checks}

Before presenting the main results, I carry out three preliminary tests to check the suitability of the RDD approach, as suggested by Imbens and Lemieux (2007).

\textsuperscript{14} I use the Stata commands ‘clustse’ and ‘cgmwildboot’ for CAT and Wild Bootstrapping respectively.
The first test helps analyse the continuity in potential outcomes assumption described above, by testing for discontinuities in observable covariates at the cut-off. It may be the case that at the cut-off it is not only the treatment that experiences a jump (from zero to one), but other observable or unobservable factors which affect outcomes. Hence the effect attributed to the treatment may pick up the effect of other factors. To test this, I run specification (2) using 31 covariates as the outcome variable $y$, with the null hypothesis that $\beta_2 = 0$ representing no discontinuity in the covariate at the cut-off. Table 1 in the Appendix reports the coefficients and p values for samepartygov in each. As can be seen, only one covariate experiences a statistically significant jump (exports) at the 5% significance level. Given that one would expect to find significance at the 5% level in 1-2 covariates if all were actually balanced, this is not too worrying, and provides good evidence that potential outcomes are continuous across the cut-off, as is required for RDD analysis.

The second test is that designed by McCrary (2008), which tests for evidence of manipulation of the running variable. It is possible that mayoral candidates of the same party as the governor are able to commit sufficient electoral fraud to push them from just losing to just winning an election. If this were the case, then it “should provoke serious scepticism about the appropriateness of the RD design” (Lee and Lemieux 2010: 18). I use the Stata command DC density for this test, which uses specification (2) with the density of the running variable (govpartywinmargin) as the outcome variable. Graphical results from the McCrary test are given in Figure 1. As is evident, the null hypothesis that there is no jump in the density of the running variable cannot be rejected (the test provides a discontinuity estimate of -0.074, with a standard error of 0.165).
The third test is a placebo test, which tests for jumps at other values of the running variable. Specification (2) is run for a number of different cut-offs (the median either side of the original cutoff, and cut-offs of -0.25 to 0.25 in 0.05 intervals). Practically, this entails adding or subtracting the placebo cut-off value from the running variable, and creating a new samepartygov dummy which equals 1 when the new running variable is greater than 0. The basic assumption of this test is that there is no reason to expect a jump at any other value of the running variable (i.e. potential outcomes are smooth for all values of the running variable), and hence finding one suggests that an RDD could be inappropriate. Table 2 in the Appendix reports the results. As can be seen, the null hypothesis that there is no treatment effect of the placebo can be rejected at the 5% significance level for cut-offs at the median below, -0.1 and -0.2 for broad corruption, and at the 10% significance level for the median above for narrow corruption. This provides a challenge to the suitability of a regression discontinuity in this case, especially two of those cases come at the median on either side (when there is the most information on either side of the cut-off). However, I note that the rate of null hypothesis rejection is not much higher than would be expected, and that in themselves, discontinuities at other values of the running variable are not damning to the RDD, so long as they do not imply discontinuities in potential outcomes around the cut-off.

**Graphical Analysis**

This section inspects a graphical presentation of the RDD approach, using the Stata command rdplot. Figure 2 presents this analysis with the fraction of resources involved in narrow corruption as the dependent variable. This figure collects observations in bins of the running variable and calculates the mean value of narrowly defined corruption in each bin. For illustrative purposes, a parametric (order four polynomial) estimate of the relationship between the running and outcome variable is shown for either side of the cut-off. There is a clear jump (downwards) at the cut-off in this graph by around 0.02 (i.e. 2 percentage points less corruption). Further, given only 7 bins to the left of the cut-off, and 8 to the right, contain more than 10 observations, within this range there is no evidence of other large jumps, confirming the overall results from the third preliminary test.
Primary Analysis

Tables 3 and 4 present the main results from the primary analysis for broadly and narrowly defined corruption respectively. Specifications 2 and 3 are run for each type of corruption for the entire sample, and for population halves. As discussed in section V, wherever possible I use the high powered CAT approach to cluster robust standard errors with few clusters developed by Ibragimov and Muller (2010; 2016). In practice this is only for the full sample; lower sample size in the population sub-samples necessitates the wild bootstrapping approach.\footnote{See section V – Cluster Robust Standard Errors for a further explanation}

The RDD estimates, which provided the assumptions in section V are satisfied can be interpreted causally, are for the whole sample negative and statistically significant (mostly at the 10% level, with one at the 5% level). The coefficients using specifications (2) and (3) (RDD linear and RDD quadratic respectively) are relatively similar for both corruption variables. The CAT coefficients (an average of the coefficients from a regression in each cluster/state) are also negative. Hence these results suggest that a mayor being in the same party, or having a coalition partner in the same party, as the state governor decreases broadly defined corruption by 2.9 percentage points, and narrowly defined corruption by 1.9 percentage points. This constitutes a 60-80% reduction in corruption from the mean in the data-set. The naive
OLS estimates show a (statistically insignificant) smaller negative relationship between being in the same party as the governor and corruption, on both measures of corruption. However, due to likely unconfoundedness/selection on unobservables, these cannot be evaluated as causal.

These results also provide evidence that the effect is greatest for municipalities with larger populations. Those in the largest municipalities (the top 50% in the population distribution) have a greater reduction in corruption (5 and 3 percentage points for broad and narrow corruption respectively, or around 100-120% of the mean corruption level of each type), and the effect is highly significant (at the 1% level) for both linear specifications. However, while the coefficient remains similar for the quadratic specification, it is only significant for broadly defined corruption. Those in the smallest municipalities have smaller (and even positive) coefficients, which are uniformly highly insignificant.
In total, these tables provide reasonable evidence that a mayor being of the same party as the governor has a causal effect on corruption for the whole population. This evidence is stronger for larger municipalities. Further, these results provide strong evidence against the hypothesis that corruption will increase when mayors' coalition partners are of the same party as the governor.

Robustness Checks

Tables 3 and 4 in effect contain one robustness check; whether specifications (2) and (3) provide similar evidence. As can be seen, on the whole this test is passed. Table 5 (in the Appendix) presents further robustness checks. Firstly, it presents results for the same RDD specifications (2 and 3) with half and double the optimal bandwidths, as well as standard bandwidths of 5%, 10% and 15%. Secondly, it presents 2 and 3 using triangular kernels, instead of rectangular. Unfortunately, the Stata commands used for triangular local linear and triangular local quadratic regressions are not compatible with the command used for cluster-robust inference with few clusters, and hence this test is run with the usual cluster robust standard-errors. Thirdly, results are presented using the total amount of corruption, rather than the proportion of resources involved in corruption, as the dependent variables. Fourthly, the analysis is narrowed to look only at a direct party relationship between the municipal mayor and the state governor (i.e. govpartywin = 1 only if the mayor is of the same party as the governor, and not if just a member of their coalition is). Finally, the wild bootstrapping technique is used for the entire sample.

In the results for the entire population, while coefficients remain fairly stable and negative, their significance depends on the value of the bandwidth, the inclusion of mayoral allies in the definition of samepartygov, and to a lesser extent using the proportion rather than amount of corruption, using rectangular kernels, and using the CAT technique (significance is retained for linear specifications of broad corruption here). While this is somewhat expected (using less than the optimal bandwidth increases variability of estimates, removing mayoral allies reduces the sample size considerably, and wild bootstrapping is lower powered than CAT), this further demonstrates that the results found in Tables 3 and 4 provide strong evidence against the hypothesis that being in the same party as the governor increases corruption, and weaker evidence that there is an effect for the entire population, and that it is significant.
The results disaggregated by population are more robust; for the top half of the population distribution, the linear RDD specification is robust to all but double the optimal (I-K) bandwidth, while coefficients remain small and highly insignificant for the lower half of the population. However, the quadratic RDD specification, while maintaining a fairly stable negative coefficient in each test, is not robust to various bandwidths, to using the amount of corruption as the dependent variable, and to using direct party, rather than coalition, links. In total, the evidence is stronger for the hypothesis that a mayor being of the same party as the governor reduces corruption in larger municipalities.

**VII) Mechanisms**

This section gives more depth to the results in the primary analysis, and aims to find evidence for or against different potential pathways for the primary results.

*More/ less corrupt parties*

Rather than being driven by the causal effect on corruption of a mayor being in the same party as the governor, the results in Tables 3 and 4 could potentially be driven by parties exhibiting differential levels of corruption. The logic of this can be seen in the following extreme scenario: say there were only two parties (A and B), one of which (A) controlled all the gubernatorial seats. The mayoral seats are then shared between A and B. Every municipality to the right of the cut-off (the mayor being the same party as the governor) then has a mayor from party A, while every municipality to the left (the mayor being a different party to the governor) has a mayor from party B. Consequently, the jump could simply reflect party differences in corruption.

To test this, I first split the dataset into one for each party (i.e. one data-set includes only those with mayors from the PMDB party), and carry out specification (2) for each data-set to see if the results hold within each party. Clearly, the very small sample size here means the test is very underpowered, and evidence should be taken as suggestive. This is done only for the 8 parties with more than 20 observations (totalling 88% of observations). For 5 out of these 8, the same negative (but insignificant) relationship as seen in Tables 3 and 4 holds within the party (workings not reported here). For the other 3, a positive (but insignificant) relationship is seen.
For the same 8 parties, I then carry out specification (2) on the entire data-set, using a dummy for the mayoral party as the dependent variable. This tests if there is a significant difference in the probability of being in a certain party if a mayor is in opposition to the governor, verses if a mayor is in the same party. Hence this tests if parties are balanced either side of the cut-off. Of the 3 problematic parties, two are not balanced across the cut-off (the PTB and PSDB).

Table 6 presents the results from removing these parties, and running specification (2) on the remainder of the sample. As can be seen, coefficients remain similar to those in Tables 3 and 4 and are significant at the same levels. This then suggests that the results in Tables 3 and 4 are not driven by party differences in corruption.

**Corruption as a weapon**

The second piece of analysis examines potential evidence for the first theorised path; that governors allow the opposition to be more corrupt in order to ‘get them in their pocket’. It looks at whether mayoral candidates in the same party as the governor are more likely to be re-elected than those in a different party *with the same level of audited corruption*. The intuition behind this is that the governor may ‘unearth’ corruption scandals or feed stories to the media regarding opposition mayoral candidates, in order to harm their electoral chances. Hence mayors of the same party as the governor are hypothesised to be more likely to be re-elected than equally corrupt members of the opposition. To test this, the dataset is restricted to those mayors not in their second concurrent term (who are ineligible for re-election) and the following multivariate linear probability model is run:

\[
\text{reelect2004}_m = \beta_0 + \beta_1 \text{samepartygov}_m + \beta_2 \text{corruptionx}_m + \\
\pi X_m + \varepsilon_m
\]

Where reelect2004 is a dummy for whether the mayor is re-elected in 2004, and corruptionx is a variable for either the proportion of broad corruption or narrow corruption. X is a vector of control variables which may be correlated with both samepartygov and the dependent variable. These include variables for improvements/failings in economic and social indicators in the mayoral term\(^\text{16}\), characteristics of the

\(^{16}\text{Changes in \{tax revenue, current transfers from the federal government, industry value added, services value added, GDP, homicide rate; all per capita\}}\)
electorate\textsuperscript{17}, characteristics of the candidate\textsuperscript{18}, and characteristics of democratic practice in that municipality\textsuperscript{19}. However, there may still be an issue of omitted variable bias here, as access to control variables is limited, and one might expect there to be several reasons why re-election rates should be lower among those in opposition to the governor. To help deal with this I use a placebo test: the regression is repeated only for those municipalities with no incidences of corruption, and hence those in which corruption scandals cannot be used as a weapon. This then helps capture the difference in re-election chances without the corruption threat; for example the benefits in terms of campaign finance and political support for those mayors of the same party as the governor. Comparing the two results helps isolate the use of corruption as a weapon.

Unfortunately, data on mayoral re-election in 2004 could only be collected for those municipalities audited in 2003 and 2004, and hence the sample size is reduced dramatically. However, as the audit process was random, this sample should be representative of the larger data-set.

Table 7 presents the results.

<table>
<thead>
<tr>
<th>Sub-Sample</th>
<th>All</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description of test</td>
<td>broad</td>
<td>narrow</td>
</tr>
<tr>
<td>Coefficient (samepartygov)</td>
<td>All Controls</td>
<td>Minus Candidate Controls</td>
</tr>
<tr>
<td>0.187</td>
<td>0.2</td>
<td>0.18</td>
</tr>
<tr>
<td>0.01***</td>
<td>0***</td>
<td>0.025**</td>
</tr>
<tr>
<td>Observations</td>
<td>89</td>
<td>89</td>
</tr>
</tbody>
</table>

\textsuperscript{17} Proportion with secondary education
\textsuperscript{18} Previous legislature or mayoral experience and education
\textsuperscript{19} Number of parties in the 2000 mayoral election, Proportion of councillors of the same party as the mayor, and Presence of a local radio station

As can be seen, for corrupt mayors, mayors of a different party to the governor are 17-20 percentage points less likely to be re-elected in 2004 than those of the same party with the same level of corruption. This effect is significant at the 1% or 5% level and is robust to the inclusion of different control variables. Further, for non-corrupt mayors, the difference is smaller (7-12 percentage points), and statistically insignificant. This then provides evidence that knowledge of corruption is used as a weapon by state
governors against opposition mayors\footnote{Of course, one might then expect that opposition mayors internalise this effect, and hence have an incentive to be less corrupt than they would otherwise be. This general equilibrium framework is absent from the mathematical model in section III for reasons of simplicity}, though one should remain aware of the small sample size, and potential for omitted variable bias.

**Mayoral corruption harms governors**

The third piece of analysis looks at the second theorised path for the primary results; the electorate blaming the national party for corruption at the mayoral level. If the governor is of the same party, then their re-election chances are hindered. To test this, I run the following linear probability model specification:

\begin{equation}
\text{reelect}_{x_s} = \beta_0 + \beta_1 \text{sameparty}_{\text{corruption}}_{x_s} + \beta_2 \text{diffparty}_{\text{corruption}}_{x_s} + \pi X_s + \varepsilon_s
\end{equation}

Each observation is now at the state level. The dependent variable is either a dummy for whether the governor is re-elected in 2002 or for whether they are re-elected in 2006 (dropping those not re-elected in 2002). The variable \text{sameparty}_{\text{corruption}} represents the average level of corruption among mayors in the same party (or with a coalition partner in the same party) as the governor, while \text{diffparty}_{\text{corruption}} is the same for those not in the same party. \(X\) is vector of state level control variables with the same categories as those for (4), but more limited\footnote{These are the change in current transfers, industry value added, services value added, GDP and homicide rate, all per capita, the education of the governor, and the number of parties in the relevant gubernatorial election.}. The econometric technique is particularly weak for this analysis; there are few control variables, and hence omitted variable bias is likely under this simple cross-sectional specification, and the sample size is very small.

Table 8 presents the results. While the coefficients are the same direction as expected, only one is significant at the 10\% level, and hence no compelling evidence is fund to support this pathway. Hence these results should be taken as no more than suggestive, and motivation for further research.
Discussion of Pathways

This section has presented fairly strong evidence against the results in section V being driven by particular parties, strong evidence of its being driven by corruption being used as a political weapon, and very weak evidence of senior politicians being hurt electorally by corruption among junior politicians in their party. These results are consistent with the model I outlined. However, there are other arguments which bear consideration, but cannot be tested here.

Firstly, it is possible that auditors colluded with state governors, and gave better audit reports for mayors in the governors’ parties. However, as described before, this is unlikely for several reasons. Secondly, the results in Tables 3 and 4 could also be consistent with a purely principal-agent model in which governors have better monitoring capacity over mayors of their own party. However, this is unlikely for two reasons. As previously noted, mayors are powerful and independent figures within municipalities, and are not subject to many formal checks and balances. Hence it is unlikely that their party would have any extra information on activities that they choose to hide. Further, the fact that the effect seen in Tables 3 and 4 is stronger for larger municipalities is inconsistent with this explanation. A principled governor would likely target larger municipalities for their anti-corruption efforts, using an array of monitoring techniques. This suggests that any party-based monitoring capacity would be relatively less important in larger municipalities, and hence the effect of party ties should be less, not more, pronounced here.

Thirdly, it may be the case that parties who do not hold the governor position in a state lack access to campaign funds (because the governor has access to more legitimate or...
illegitimate funds). Consequently, they may look to the municipal level for illicit fund-raising. This is not implausible; much of the funding for the federal level bribes in the 2005 Brazilian Mensalão scandal came from municipal level corruption (Power and Taylor 2011), but does not explain the re-election results found in this section. This is a good ground for future research.

**VIII) Conclusion**

The empirical evidence presented suggests that in Brazil, state governors will allow or encourage opposition mayors to be embezzle or divert more resources than those in their own party. On average, opposition mayors engage in extra corruption amounting to 60-80% of the mean level of corruption. In large municipalities, the effect increases to 100%, and the results are more robust. However, there are two caveats: results for the whole sample are not always robust to changes in bandwidth or measurement technique, and there is weak evidence of jumps at points of the distribution of the running variable other than the cut-off, which could undermine the Regression Discontinuity Design approach. Overall, the results are consistent with a model where the state governor tolerates corruption mainly in order to gain control over mayors and to protect their own electoral chances. It is not consistent with a model where the state governor tolerates corruption mainly to reward its supporters.

Further, this paper has found further evidence consistent with the interpretation that corruption is used as a political tool, as mayors with party links to the governor have higher re-election rates (17 to 20 percentage points) than those in a different party with the same level of corruption (controlling for other factors which may influence re-election). Weaker evidence has been found for the theory that governors suppress corruption by mayors of their own party to boost their own electoral chances.

All together, this provides evidence that corruption is not simply a principal-agent everywhere and anywhere. In Brazil, a state with neo-Patrimonial characteristics (Bach 2011), there is evidence of unprincipled principals who use corruption to further their own ends. This then provides a strong motivation to continue unpacking the political factors which drive corruption. A potential extension could explore the political determinants of corruption beyond the governor-mayor linkage. Each may be subject to pressures to be corrupt from further agents, for example federal politicians, and the electorate. My results also provide policy implications. If corruption is used as a tool for
political stability, those at the top play a large part in sustaining it for their own interest. Consequently, designing strategies in which stability can be maintained through formal, less socially harmful practices may give anti-corruption programmes the political support they need to succeed.
Bibliography


• **McCrary, J.** (2008), ‘Manipulation of the running variable in the regression discontinuity design: A density Test’ Journal of Econometrics 142 (2), 698-714

• **Medard** (1982), ‘The Underdeveloped State in Tropical Africa: Political Clientelism or Neo-Patrimonialism’ in Clapham (ed.) Private Patronage and Public Power (London, Francis Pinter, 1982)


• **Treisman, D.** (2007). ‘What have we learned about the causes of corruption from ten years of cross-national empirical research?’ *Annual Review of Political Science*, 10, 211-244.


### Appendix

#### Table 1 - Covariate Discontinuities

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Coefficient</th>
<th>P-Value</th>
<th>CAT/Wild Bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lottery Number</td>
<td>0.178</td>
<td>0.331</td>
<td>CAT</td>
</tr>
<tr>
<td>Population</td>
<td>-10520.62</td>
<td>0.694</td>
<td>CAT</td>
</tr>
<tr>
<td>Geographic Size</td>
<td>1081.878</td>
<td>0.22</td>
<td>Wild</td>
</tr>
<tr>
<td>Exports</td>
<td>-14159.256</td>
<td>0.04**</td>
<td>Wild</td>
</tr>
<tr>
<td>Tax Revenue</td>
<td>-15759.567</td>
<td>0.406</td>
<td>CAT</td>
</tr>
<tr>
<td>Current Transfers</td>
<td>-382604.56</td>
<td>0.435</td>
<td>CAT</td>
</tr>
<tr>
<td>Distance to the Federal Capital</td>
<td>139.543</td>
<td>0.372</td>
<td>CAT</td>
</tr>
<tr>
<td>Distance to the State Capital</td>
<td>-16.301</td>
<td>0.329</td>
<td>CAT</td>
</tr>
<tr>
<td>Agriculture Area</td>
<td>1995.855</td>
<td>0.193</td>
<td>Wild</td>
</tr>
<tr>
<td>Agriculture Employees</td>
<td>1362.826</td>
<td>0.33</td>
<td>CAT</td>
</tr>
<tr>
<td>Industry Value Added</td>
<td>-22539.424</td>
<td>0.266</td>
<td>CAT</td>
</tr>
<tr>
<td>Services Value Added</td>
<td>-40779.288</td>
<td>0.861</td>
<td>CAT</td>
</tr>
<tr>
<td>GDP</td>
<td>-77018.957</td>
<td>0.488</td>
<td>CAT</td>
</tr>
<tr>
<td>GDP Per Capita</td>
<td>95.28</td>
<td>0.217</td>
<td>CAT</td>
</tr>
<tr>
<td>Judicial District</td>
<td>-0.232</td>
<td>0.41</td>
<td>Wild</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>-0.006</td>
<td>0.7</td>
<td>Wild</td>
</tr>
<tr>
<td>Homicide Rate</td>
<td>-0.536</td>
<td>0.012</td>
<td>CAT</td>
</tr>
<tr>
<td>% Public With Secondary Education</td>
<td>0.001</td>
<td>0.845</td>
<td>CAT</td>
</tr>
<tr>
<td>Participatory Budgeting</td>
<td>0.096</td>
<td>0.24</td>
<td>CAT</td>
</tr>
<tr>
<td>Number of Councillors</td>
<td>0.094</td>
<td>0.347</td>
<td>CAT</td>
</tr>
<tr>
<td>% Councillors from Same Party as the Mayor</td>
<td>0.045</td>
<td>0.168</td>
<td>Wild</td>
</tr>
<tr>
<td>Number of Parties in Mayoral Election</td>
<td>0.861</td>
<td>0.296</td>
<td>CAT</td>
</tr>
<tr>
<td>Number of Parties in Councillor Election</td>
<td>0.25</td>
<td>0.342</td>
<td>CAT</td>
</tr>
<tr>
<td>Mayor Gender (Male=2, Female=4)</td>
<td>-0.153</td>
<td>0.229</td>
<td>CAT</td>
</tr>
<tr>
<td>Mayor Education</td>
<td>-0.225</td>
<td>0.885</td>
<td>CAT</td>
</tr>
<tr>
<td>Mayor Marital Status(Single=1, Married=3)</td>
<td>-0.681</td>
<td>0.112</td>
<td>CAT</td>
</tr>
<tr>
<td>Mayor Age</td>
<td>2.617</td>
<td>0.303</td>
<td>Wild</td>
</tr>
<tr>
<td>First Term Mayor</td>
<td>0.194</td>
<td>0.635</td>
<td>CAT</td>
</tr>
<tr>
<td>Second Term Mayor</td>
<td>-0.194</td>
<td>0.635</td>
<td>CAT</td>
</tr>
<tr>
<td>Mayor a Legislator in 1996 Term</td>
<td>0.056</td>
<td>0.52</td>
<td>Wild</td>
</tr>
<tr>
<td>Mayor Previously (Non-Concurrently)</td>
<td>0.088</td>
<td>0.8</td>
<td>Wild</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

Coefficient gives the coefficient from a standard application of specification (3) with the covariate as as the dependent variable. The associated p value is given using either the wild bootstrapping or CAT technique for cluster robust standard errors with few clusters.

#### Table 2 - Placebo Tests

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Median Above (0.1374)</th>
<th>Median Below (-0.104)</th>
<th>fraction_broad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.25</td>
<td>0.2</td>
<td>0.15</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>CAT Coefficient</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Median Above (0.1374)</th>
<th>Median Below (-0.104)</th>
<th>fraction_narrow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.25</td>
<td>0.2</td>
<td>0.15</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>CAT Coefficient</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

Coefficient gives the coefficient from a standard application of specification (2) with the relevant dependent variable. The associated p value is given using either the wild bootstrapping or CAT technique for cluster robust standard errors with few clusters, using the stata commands ggmwboot and cluster respectively. The second coefficient calculated in the CAT procedure for the cluster average is also reported. Each test uses rectangular kernels for local linear regression using the I-K optimal bandwidth. The difference in observations reflects different optimal bandwidths.
**Table 5 - Robustness Checks**

<table>
<thead>
<tr>
<th>Dependent Variable Sample Description of test</th>
<th>fraction_broad</th>
<th>fraction_narrow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RDD - Linear</td>
<td>RDD - Quadratic</td>
</tr>
<tr>
<td></td>
<td>Top 50% Population</td>
<td>Bottom 50% Population</td>
</tr>
<tr>
<td>Original Results Coefficient (samepartygov)</td>
<td>-0.029</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.024**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.02**</td>
</tr>
<tr>
<td>Half Optimal B-W Coefficient (samepartygov)</td>
<td>-0.024</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.02**</td>
</tr>
<tr>
<td>Double Optimal B-W Coefficient (samepartygov)</td>
<td>-0.011</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.168</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.155</td>
</tr>
<tr>
<td>B-W of 0.15 Coefficient (samepartygov)</td>
<td>-0.028</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.026**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.035**</td>
</tr>
<tr>
<td>B-W of 0.1 Coefficient (samepartygov)</td>
<td>-0.014</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.073*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.005***</td>
</tr>
<tr>
<td>B-W of 0.05 Coefficient (samepartygov)</td>
<td>-0.013</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.466</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.075*</td>
</tr>
<tr>
<td>Triangular kernels Coefficient (samepartygov)</td>
<td>-0.024</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.016**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.035**</td>
</tr>
<tr>
<td>Amount Corruption Coefficient (samepartygov)</td>
<td>-736883.5</td>
<td>-869483.2</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>-0.032**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.04**</td>
</tr>
<tr>
<td>Direct Party Coefficient (samepartygov)</td>
<td>-0.034</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.18</td>
</tr>
<tr>
<td>Wild Bootstrap Coefficient (samepartygov)</td>
<td>-0.029</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>P-Value (samepartygov)</td>
<td>0.07**</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

As in the original results, columns (2), (3), (9) and (10) are carried out using the CAT clustering procedure, except for the 'Wild Bootstrap' row which uses wild bootstrapping. The remainder of the columns use the wild bootstrapping clustering procedure, again as in the main results. The row 'Triangular Kernels' uses standard cluster robust standard errors. Unless otherwise specified, each test uses rectangular kernels for local linear regression using the I-K or CCT optimal bandwidth.
<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>broad (1)</th>
<th>narrow (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient (samepartygov)</td>
<td>-0.023</td>
<td>-0.018</td>
</tr>
<tr>
<td>P-Value (samepartygov)</td>
<td>0.019**</td>
<td>0.1*</td>
</tr>
<tr>
<td>Observations</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Cluster Method</td>
<td>CAT</td>
<td>CAT</td>
</tr>
<tr>
<td>CAT Coefficient (samepartygov)</td>
<td>-0.059</td>
<td>-0.036</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1

Coefficient gives the coefficient from a standard application of specification (2). The associated p value is given using either the wild bootstrapping or CAT technique for cluster robust standard errors with few clusters. The second coefficient calculated in the CAT procedure for the cluster.