COMBATING POLITICAL CORRUPTION WITH POLICY BUNDLES

Frederico Finan
Maurizio Mazzocco

ABSTRACT

In this paper, we develop a dynamic model of politicians who can engage in corruption. The model offers important insights into what determines corruption and how to design policy to combat it. We estimate the model using data from Brazil to measure voters' willingness to pay for various commonly-proposed anti-corruption policies, such as increasing audit probabilities, increasing politicians' wages, and extending term limits. We document that while voters have a high willingness to pay for audit policies, due to their effectiveness in reducing corruption, a multi-pronged approach that bundles certain policies can achieve similar welfare gains at a fraction of the cost.

Frederico Finan
Department of Economics
University of California
508-1 Evans Hall #3880
Berkeley, CA 94720-3880
and NBER
ffinan@econ.berkeley.edu

Maurizio Mazzocco
Department of Economics
University of California, Los Angeles
Bunche Hall 8377
Los Angeles, CA 90095
mmazzocc@econ.ucla.edu

1   Introduction

The abuse of entrusted power by politicians through rent-seeking and corruption is a serious concern in much of the developing world. There have been countless examples both across country and over time of political elites diverting funds intended for basic public services such as in education, health, and infrastructure (Rose-Ackerman and Palifka (2016); Fisman and Golden (2017)). Not surprisingly, corruption is widely considered to be a major obstacle for economic and social development, and several studies have documented a strong negative relationship between corruption and various measures of economic development such as investment and growth (Mauro (1995); Bai et al. (2017); Colonnelli and Prem (2019)). Therefore, designing policies effective at reducing political corruption is of first-order importance.

Policymakers and academics have proposed and evaluated several policies to combat corruption. The most common approaches include government audits, extending political time horizons, or increasing politicians' wages. Importantly, several studies have found empirical support for such policies in various settings. Nevertheless, the existing literature on anti-corruption policies remains limited in three important ways. First, it is difficult to compare across policies when they have been evaluated during different periods and settings. For example, we have seen that audit policies can be effective in reducing corruption in Brazil, China, and Indonesia, but how do they compare to extending term limits in Mexico? Second, there are strengths and weaknesses associated with any single policy. Perhaps, we can do better by combining policies that minimize each individual policy's limitations. Third, the evidence on the effects of anti-corruption policies comes mostly in form of reduced-form findings. Politicians, however, are forward-looking actors who make dynamic decisions and anti-corruption policies are likely to affect not only their current choices, but their future ones as well. It is difficult to capture these future margins of adjustment in the reduced-form.

To address these gaps in the literature, we need to better understand why politicians engage in corruption over the course of their life. Specifically, we need to understand not only the main incentives and constraints politicians face, but also how their current decisions affect their future choices. Herein lies the main contributions of our paper. We develop and estimate a dynamic model of an incumbent politician's decision to, among other things, engage in corruption. By

simulating the estimated model, we are then able to compare within the same setting the effects of various anti-corruption policies, including the combination of policies, i.e. policy bundles.

We develop a model in which local incumbent politicians decide how much to steal versus how much to invest in the production of public goods. Politicians are heterogeneous in their (unobserved) ability to produce these goods. The decision to steal in a given period affects future outcomes and decisions, including the decision to run for office, future wages and fines, and reelection chances. Consistent with our data and previous studies, voters care about public consumption and will punish politicians who are found to be corrupt.[1] Voters in our model also care about a politician's electoral appeal, which provides another source of unobserved heterogeneity across politicians.

Our model is quite general, as it is applicable to various settings. But to estimate it, we rely on data from local governments in Brazil. Local governments in Brazil provide an ideal institutional setting to study corruption for at least four reasons. First, mayors receive millions of dollars each year from the federal government to provide local public goods, including education, health, and sanitation. With the large influx of federal funds and limited federal oversight, local corruption in Brazil has been a serious concern. According to our data, corruption was discovered in 73 percent of all municipalities, where on average 8.2 percent of these federal funds were diverted. This translates into losses of approximately $600 million in local governments per year. Second, in 2003 the Brazilian government introduced an anti-corruption program that randomly audited municipal governments for their use of federal funds. These audits provide an objective measure of corruption that together with the program's randomization are crucial for the identification and estimation of the model's parameters. Third, in 1997 Brazil allowed mayors to hold office for two consecutive terms. This variation is important for identifying the effect of re-election incentives on corruption. Finally, besides the data on corruption, we also have detailed information about all candidates that ran for mayor since 2000, including their age, education, wealth, and future wages in the formal sector.

Our estimated model matches several important features of the data, including ones that are not targeted in the estimation. For example, we can match well the difference in stealing between mayors who are in their first term versus those who are in their second and final term. This

---

[1] Several studies have found evidence of voters punishing corrupt politicians at the polls. See Olken and Pande (2012) for a review of the literature.

comparison is important because it reflects the combination of two different forces resulting from elections. First, it reflects a dynamic decision by mayors to forgo stealing in their first term in order to get reelected to a second term. Second, it captures the selection effects created by elections. We find that second-term mayors are on average more positively selected than first-term mayors both in their observed and unobserved characteristics. An important feature of our model is that we can compute how much of this difference is due to selection. We find that selection accounts for 32% of the difference in stealing between first and second term mayors.

Given these insights, we then use our estimated model to quantify the reduction in stealing and subsequent welfare gains of five commonly proposed anti-corruption policies. These include increasing the probability of a federal audit, extending the number of terms mayors can serve, auditing mayors who have been charged but not yet convicted of corruption, banning corrupt politicians from running in future elections – an actual policy in Brazil referred to as the "Clean Record" Act –, and doubling mayor's wages. Among these five policies, we find that increasing the probability of an audit is the most effective at reducing corruption. An increase in the probability from 5%, the value at the beginning of our sample period, to 16.8%, the value at the end of our sample period, decreases corruption by 41 percent and municipal residents are willing to pay on average 1.23% of their annual income to adopt this policy during the course of their lifetime. For comparison, adding an additional term to the number of consecutive terms a mayor can hold office is about half as effective. We estimate that this policy would reduce corruption by about 23 percent, with residents willing to pay on average 0.6% of their annual income.

The most disappointing policy is the one that increases the politicians' wages. Residents must be compensated with 0.14% of their annual income to accept such a policy. The reason why is because the wage policy only incentivizes politicians who can and choose to run for reelection. For these politicians, higher wages does lead to less corruption. But it has the opposite effect on politicians who cannot or choose not to run for reelection, with an increase in the fraction stolen that outweighs the decline for the first group. Our results therefore indicate that this policy has limited or even adverse effects if applied to politicians that do not have or can avoid electoral incentives.

The fact that we can compare across different policies all within the same setting is an important contribution of our paper. Another contribution is that we also simulate the effects of

combined policies. This feature is important because, as we demonstrate with our model, each policy has its strengths and weaknesses. For example, the term-limit policy reduces corruption because it strengthens an incumbent's electoral incentives for one additional term. But it has limited effects on mayors who are in their last term or have electoral incentives even without the reform. The "Clean Record" Act has the same weakness: it only affects politicians that plan to run for reelection. Despite these limitations, we can increase the efficacy of the two policies by just combining them. The restriction that corrupt politicians cannot run for reelection is more effective because they can participate in the elections for one additional term. The term-limit policy has a larger impact because it also affects earlier terms through the no-run restriction. By just combining the two policies, the average willingness to pay more than doubles from 0.6% to 1.3% of annual income.

The efficacy of the combined policy is still limited by the lack of an effect on last term mayors. We can remedy this drawback by simply increasing the probability that a last-term mayor is audited, as a way to compensate for the absence of electoral incentives. If we increase the audit probability in the last term to the level used in the audit policy, the average willingness to pay increases further to 1.36%. In addition, having to audit only mayors in their last term offers another advantage. Audit policies, although effective, are expensive because they necessitates a direct financial disbursement to pay for the increase in number of audits. Thus, it saves to have audits that are targeted, which is what this last policy bundle also achieves.

Our paper contributes to several strands of the literature. First, we contribute to the literature on corruption and in particular, on anti-corruption policies.[2] Several papers have evaluated anti-corruption policies, with government audits or crackdowns being the most common approach. For example, Olken (2007) conducts a field experiment in Indonesia that increases the probability of a government audit from 4% to 100%. He finds that this intervention reduced corruption in road projects by 8 percentage points. Bobonis, Camara Fuertes, and Schwabe (2016) studies Puerto Rico anticorruption program. They find that the disclosure of information about corruption in a municipality does reduce corruption level, but only in the short run. In subsequent terms, municipal corruption levels increased, especially among those who refrained from corruption prior to the first audit. Avis, Ferraz, and Finan (2018) analyze Brazil's anti-corruption program and

---

[2]For excellent surveys on corruption, see Olken and Pande (2012); Rose-Ackerman and Palifka (2016).

exploit the fact that municipalities have been audited multiple times at random. They find that there were 8 percent fewer acts of corruption in municipalities that had been audited in the past compared to those that had never been audited. Chen and Kung (2018) show that China's recent anti-corruption crackdowns reduced corruption by 42.6% in the provinces targeted by the central inspection teams.

Some studies have provided evidence to suggest that extending political time horizons might also reduce corruption.[3] For example, using the same data presented in this paper, Ferraz and Finan (2011) have shown that second-term mayors, who are no longer eligible for reelection are significantly more corrupt than mayors with reelection incentives. Lopez-Videla (2020) studies a recent reform in Mexico that allowed mayors, who had been limited to a single term, to run for re-election for an additional three-year term. Using the staggered implementation of the law, he shows that mayors with longer time horizons steal less and provide more public goods.

Another frequently-proposed anti-corruption policy has been to increase politicians' wages. As Becker and Stigler (1974) originally pointed out, by increasing the value of a job, the employee will refrain from stealing as long as there exists a realistic threat of punishment. We have seen some evidence for this behavior, not among politicians, but for bureaucrats. For example, Tella and Schargrodsky (2003) analyze a corruption crack down on hospitals' input prices in Argentina. They find that the association between wages and input prices (i.e. their measure of corruption) varied according to the audit intensity. Niehaus and Sukhtankar (2013) use panel data on corruption in India's National Rural Employment Guarantee Scheme and find that higher daily wages lead to lower theft from piece-rate projects.

Our study contributes to these aforementioned strands of the literature in several ways. First, while many of these reduced-form studies have provided important causal estimates of the effects of an anti-corruption policy, their estimates shed only limited insights into the mechanisms that produce these effects and, hence, on the strength and limitations of the policies. In contrast, the model we estimate captures many of the mechanisms by which politicians choose to engage in corruption, which enables us to assess empirically their relative importance and the advantages and disadvantages the policies can present. It is the understanding of these mechanisms that is critical for the design of policy as a redress for corruption. Second, our approach enables us to

---

[3]See Ashworth (2012) for an excellent review of the literature.

simultaneously evaluate several policies in the same setting and establish which one is the most effective at reducing corruption and why. Third, our model allows us to estimate not only the effects on corruption, but also welfare, which is arguably what we ultimately care about.

Our paper relates also to a growing literature that estimates structural models of political decisions to study how reforms to institutions, including term limits, can affect politicians' behavior (e.g. Diermeier, Keane, and Merlo (2005), Stromberg (2008), Lim (2013), Aruoba, Drazen, and Vlaicu (2015), and Sieg and Yoon (2017), Finan and Mazzocco (2020)), regulators' decisions (e.g. Kang and Silveira (2020)), and the return from lobbying (e.g. Kang (2015)). In this paper, we estimate to our knowledge the first structural model of the decision to engage in corruption over the lifetime of a politician.

Finally, our paper is part of a growing literature that uses randomized variation for structural estimation. See for instance, Todd and Wolpin (2006), Kaboski and Townsend (2011), Attanasio, Meghir, and Santiago (2012), and Meghir et al. (2019).

2 Political Corruption and Politics in Brazil

The model we develop below is quite general and applicable to various settings. But to estimate it, we will use data from local governments in Brazil. In this section, we describe the data and present some key reduced-form findings that motivate our modeling and estimation choices. In particular, we investigate six questions: 1) How is corruption distributed across mayors? 2) How do reelection incentives affect corruption? 3) How much does corruption affect public consumption? 4) Does being found to be corrupt affect the decision to run for re-election? 5) Do voters punish corrupt politicians? 6) Does being found to be corrupt affect a mayor's future wages? The patterns we present below are not necessarily specific to Brazil; other studies have documented similar findings in other settings.

Public Funds and Corruption. In 2003, Brazil's federal government started a national program to audit municipalities for their use of federal funds. Our data on corruption come from the audit reports generated by this anti-corruption program. In particular, we use the corruption measure created by Ferraz and Finan (2011) – the total amount of resources related to corrupt

activities as a share of the total amount of resources audited.[4] These data, which span the period 2001-2003, document that municipal corruption is a serious concern in Brazil. As we can see from Table 1, mayors received on average R$2,038,274 of federal transfers from the federal government to provide local public goods, including public education, health, and sanitation, and stole 6.3 percent of them. There is also a considerable amount of heterogeneity. The $25th$, $50th$, and $75th$ percentiles of the fund distribution are equal to $R\$806,372$, $R\$1,184,342$, and $R\$2,051,654$, which indicates that the distribution is skewed right (skewness = 10.81). The distribution of the fraction stolen has similar features. It is skewed right, with a skewness of 2.85, and its $25th$, $50th$, and $75th$ percentiles are at 0, 2.1, and 7.6 percent, respectively. In addition, about 26% of audited mayors were not found to be corrupt.

As originally documented by Ferraz and Finan (2011), term limits affect corruption levels. In Brazil, mayors can only serve two consecutive terms. We find that mayors who were in their first term steal on average 5.6 percent of the allocated funds, whereas second-term mayors divert 7.3 percent, a 30 percent increase. In column 1 of Table 2, we show that this difference is also robust to controlling for various mayor and municipal characteristics.[5] The fraction of mayors caught stealing is also significantly different between the two terms: 71% of first term mayors were found to be corrupt compared to 76% of second-term mayors.

These results are consistent with a broader literature showing how politicians with shorter time horizons are often associated with worse outcomes. In addition to the studies described above, Coviello and Gagliarducci (2017) document for the case of Italian mayors that, on average, costs of public work were significantly higher in municipalities with a term-limited mayor relative to municipalities with a first-term mayor. Also, having the same mayor in power for an additional second term increased the likelihood that the mayor awarded the public contract to a local firm or to the same firm repeatedly, which they argue is suggestive of corruption. Gamboa-Cavazos and Schneider (2007) use firm-level data from Mexico on extra official payments made to public authorities and document that these payments, which the authors interpret as bribes, are a function of how long the politician has been in office.

---

[4]Ferraz and Finan (2011) define political corruption as any irregularity associated with fraud in procurements, diversion of public funds, and over-invoicing. See Ferraz and Finan (2008) and Ferraz and Finan (2011) for a description of the anti-corruption program and details on the construction of the data.

[5]See Ferraz and Finan (2011) for additional tests of robustness.

Public Consumption. Mayors affect the welfare of its citizens mostly through the provision of public goods, such as education and health. They do this directly by funding their production, but also indirectly by setting policies that can affect the economy more generally. To capture the various ways in which mayors can affect their citizens' welfare, we use average per-capita GDP of the municipality over the term as a proxy for the public consumption provided by the mayor. These data are constructed by Brazil's national statistical institute (IBGE) and are available yearly since 2000.[6] One potential issue with using this measure as a proxy for government value-added is that it also contains activity from the private sector. Thus, when we structurally estimate the effect of a mayor's decisions on the provision of public goods, we control for an index of private inputs that we constructed using Brazil's employer-employee matched data (Relação Anual de Informações Sociais (RAIS)). Specifically, we control for the first principal component of a factor analysis that includes the number of firms in a municipality, average private sector wages, and the rate of employment, all measured in 2000.

We investigate the relationship between corruption and public consumption in column 2 of Table 2. We regress the log of per-capita public consumption on the log of federal funds, the log of the amount of funds diverted, a dummy for being in the second term, the log of population, literacy rate, and GDP at the beginning of the term. We also allow the effects of literacy rate and GDP to vary according to whether the mayor is in his second term. We find that per-capita public consumption is positively associated with federal transfers, but negatively associated with corruption. The coefficient on log corruption implies that a 10% increase in the amount diverted is associated with a 2.8% reduction in per-capita public consumption. We also find a positive coefficient on the indicator for being in a second term, but the point estimate is imprecisely measured.

Decision to Run and Electoral Outcomes. Elections for mayors take place every four years. We use data from the 2000 and 2004 elections. Besides the election results, these data provide information on various demographic characteristics, including each candidate's gender, age, years of schooling, and self-reported wealth.[7] We summarize these characteristics for the mayors in our

---

[6]We downloaded them at site: www.ipeadata.gov.br. See the online appendix for a description of the databases and their corresponding variables used in the analysis.

[7]We downloaded the election data from Brazil's Electoral Commission (Tribunal Electoral Superior). www.tse.gov.br.

sample in Table 1.

In the 2004 elections, 72% of mayors ran for reelection. In column 3 of Table 2, we investigate whether Brazil's anti-corruption program affected this decision. Specifically, we estimate the probability of running for reelection on whether the mayor was audited, whether the mayor was caught stealing as a result of the audit, and log of per-capita public consumption during the term. We also control for log of private per-capita GDP, mayor's age, log of population, and literacy rate. We find that having been caught diverting some public resources reduced the probability of running by 12.3 percentage points compared to mayors who were audited but not found to be corrupt. We find a positive correlation between public consumption and the decision to run: a 10% increase in public consumption is associated with a 13.5 percentage point increase in the likelihood of running. Our results also indicate that older mayors are less likely to run for reelection, with one additional year of age being associated with a 0.033% decline in the probability.

Among the mayors who ran for reelection, 57% were reelected. In column 4 of Table 2, we investigate whether some of the same factors that affect the decision to run also affect reelection rates. Mayors who were caught stealing have a reelection probability that is 15 percentage points lower than mayors who were audited but not found to be corrupt. This finding is not only consistent with the results originally found in Ferraz and Finan (2008), but also with those found in other settings as well; for example Bobonis, Camara Fuertes, and Schwabe (2016) in the case of Puerto Rico, Costas-Perez, Sole-Olle, and Sorribas-Navarro (2012) for Spain, and Chong et al. (2015) for Mexico.

As opposed to corruption, an increase in per-capita public consumption has a positive association with reelection rates. We estimate a coefficient of 0.204, which implies that a 10% increase in public consumption during the term is associated with a 2 percentage point increase in the probability of winning. This correlation is consistent with a large empirical literature showing that incumbent politicians are more likely to be reelected when growth rates and public good provision are higher. For example, using a sample of 74 countries over the period 1960-2003, Brender and Drazen (2005) show that higher growth rates in GDP per capita are associated with higher reelection rates in lesser developed countries and newer democracies.

Earnings of Mayors and Ex-Mayors. Given the negative correlation between corruption and electoral success, one might also wonder whether being found to be corrupt from an audit affects the mayor's future wages. To shed light on this question, we match our politicians who are no longer in office to the RAIS data over the period of 2005 to 2013. We then compute their average wage in the private sector over this period.[8] Altogether, we are able match 68 percent of politicians to at least one post-office wage.[9]

In column 5 of Table 2, we report the estimated coefficients of a regression of wages of ex-mayors on their education, age, age squared, population size, a dummy for being audited, and a dummy equal to one if the mayor was caught stealing. While all the variables commonly included in wage regressions have the expected sign and are statistically significant, we find no evidence that having been identified as a corrupt politician has an effect on future earning: the coefficients on both the audit and corruption dummies are small and statistically insignificant.

To estimate the model, we also need a measure of how much mayors earn while in office. In principle, mayors can set their own salaries and, while there is no readily accessible dataset that contains this variable, this information is publicly available on most municipality's websites. To collect these data, we randomly sampled 10% of municipalities stratified by three population thresholds. We then downloaded the mayor's wage from the mayors' office website. The average monthly earnings paid to mayors was equal to R$3,233 for municipalities with a population less than 10,000 residents, R$4,268 for municipalities with population between 10,000 and 50,000 residents, and R$5,077 for municipalities with a population above 50,000.

Summary of Main Empirical Findings In this section, we have highlighted six empirical patterns that motivate our model's choices below. First, our model should be able to account for the possibility that the fraction stolen is higher among politicians serving in their last term. Second, there is substantial heterogeneity in the amount politicians steal, with the stealing distribution skewed to the right. Third, the amount of public good produced by the incumbent depends

---

[8]The results are similar if we instead use their maximum wage over this period.

[9]There are two principal reasons why we were unable to match the remaining 32 percent. First, if a mayor become self-employed and did not hire any employees over the 8-year period, then they would not appear in the RAIS. Second, if the mayor decided to either retire or work exclusively in the informal sector, they would not appear in the RAIS. We find that a mayor's education level is the primary predictor of whether or not they appear in the RAIS. Importantly, whether the mayor was audited and found to be corrupt does not predict their likelihood of appearing in the RAIS.

on the share of funds invested in its production. Fourth, it is important to model the decision to run because a significant fraction of politicians choose to forgo reelection and the actions a mayor takes while in office, such as choosing to steal, may affect this decision. Fifth, whether voters choose to vote for the incumbent depends on the politician's actions while in office, such as the amount of funds the mayor diverts and invests in the production of the public good and, hence, the actual amount of public good produced. Lastly, the wages of ex-mayors who were found to be corrupt in the past are not adversely affected.

3 Model

In this section, we develop a finite-horizon model of an incumbent politician's decision to engage in corruption over the course of their lifetime. Each term, politicians decide how much to steal, how much of their resources to save, and whether to run for reelection. Once stealing decisions are made, public goods are determined, and voters must decide whether or not to reelect the incumbent. Politicians are heterogeneous both in their ability to produce public goods and their electoral appeal.

3.1 Preferences and Technology

Residents care about the amount of public good they receive (e.g. schools, police force, parks, and roads). Local governments produce these public goods using public funds. Thus, all else equal, residents will consume less public goods and experience lower levels of welfare, the more funds mayors divert.

Preferences. We consider a municipality m populated by n individuals living for T periods, all of whom are potential politicians and have a common discount factor β. Individuals have preferences over a private good c and public consumption Q, which the municipality produces. Not all goods provided by the local government are pure public goods, as some have a degree of rivalry. For example, individuals who live in more populated areas may enjoy parks less because of overcrowding. To account for this, individuals derive utility from adjusted per-capita public consumption $\bar{Q} = \frac{Q}{d^\eta}$, where d represents the population size (*density*) of the municipality and the

11

parameter $\eta \in [0, 1]$ measures the degree of rivalry, with $\eta = 0$ indicating no rivalry and $\eta = 1$, full rivalry.

Individuals who run for election must pay a utility cost κ, but if elected, derive a utility ρ from being in power. Thus, we can characterize an individual's preferences with the following utility function:

$$u^i \left(c_t^i, \bar{Q}_t \right) + \rho - \kappa.$$

In the estimation, we will assume that

$$u^i \left(c_t^i, \bar{Q}_t \right) = \frac{\left(c_t^i \right)^{1-\delta}}{1 - \delta} + \theta \ln \bar{Q}_t,$$

where θ represents the relative taste for public consumption.

Technology.   Mayors affect the production of per-capita public consumption in two ways. First, they choose how much of the public funds f_t^{pu} to invest in its production, z_t^{pu}, and how much to divert, s_t^i. Second, mayors are heterogenous in their ability to produce the public good, which is comprised of two parts: an observable part, which is the mayor's education level, e_i; and an unobservable part, a_i, which is drawn from a log-normal distribution with mean μ_a and standard deviation σ_a. Because the amount of public goods vary by population size and are produced with the help of firms operating in the municipality, we also allow the production of per-capita public consumption to depend on private sector's characteristics, z_t^{pr}, and population size, d_t. In the estimation, we assume that the public consumption production function has the following form:

$$\frac{Q_t}{d_t} = \left(\frac{z_t^{pu}}{d_t} \right)^{\alpha_1} (z_t^{pr})^{\alpha_2} \exp \left\{ \alpha_3 + \alpha_4 e_i + \sum_{j=1}^{N_d} \alpha_{4+j} (d_t)^j \right\} a_i.$$

We specify the production function in terms of per-capita public consumption and per-capita inputs and we control in a flexible way for population using a polynomial of order N_d to account for differences in production functions across municipalities of different sizes. We also impose $\alpha_1 < 1$ so that the production function is concave in public inputs.

3.2 Mayors' and Voters' Decisions and Characteristics

Mayors. In our model, mayors make three decisions. Given the transfers they exogenously receive from the central government f_t^{pu}, mayors first decide the amount to divert s_t and hence how much to invest, $z_t^{pu} = f_t^{pu} - s_t$, in the production of public consumption. Mayors then decide how much to consume c_t versus save b_t of their resources, including any money they have stolen. Finally, mayors choose whether to run for reelection, provided they are not in their last term.

The central government audits municipalities at random with probability p_t. Mayors take this into account when deciding how much to steal, how much to consume, and whether to run for reelection. If a municipality is audited, which we will denote by $\delta_{au} = 1$, and public resources have been diverted, the mayor is caught with probability 1 and the amount stolen becomes public knowledge. The mayor will incur a fine that is increasing in the amount stolen, $g(s_t)$. As in other countries, Brazil does not set the size of the fine ex-ante. Judges determine the fines on a case-by-case basis and are in addition to returning the amount stolen. We model this heterogeneity in fine size by drawing a multiple of the amount stolen that the mayor must pay, τ, from a log-normal distribution with mean μ_τ and variance σ_τ^2. The fine schedule represents one of the main potential deterrents of corruption.

Each individual in a municipality owns $\bar{h}$ units of labor, supplied inelastically in return for a wage w that depends whether the person is currently a mayor or an ex-mayor. Mayors receive a deterministic wage $\bar{w}$ set by the municipality based on population size. Ex-mayors receive wages drawn from the distribution $p_w(w|Z)$, where Z denotes a vector of individual and municipal characteristics that determine wages of ex-mayors. The data suggest that ex-mayors do not experience lower wage offers if they were caught diverting public funds. We therefore obviate this consideration and model their wage process using a specification that depends only on education, a second order polynomial in age, and indicators for whether the person resides in a medium or large municipality. Specifically, we let

$$\ln w_t^{pm} = \gamma_0 + \gamma_1 e_t + \gamma_2 age_t + \gamma_3 age_t^2 + \gamma_4 \delta_{mm} + \gamma_5 \delta_{lm} + \epsilon_t,$$

where δ_{mm} is a dummy equal to one if the population in the municipality is between 10,000 and 50,000 and δ_{lm} is equal to one if the municipality's population is above 50,000.

Individuals can save or borrow an amount b at an interest rate R. A mayor's wealth affects their corruption choices in two countervailing ways. On the one hand, given the concavity of the utility function, richer mayors should steal less because they already have sufficient resources to provide for their private consumption. On the other hand, they should steal more because the effect of the financial punishment is less of a deterrent for mayors who can easily afford to pay the fine.

In our model, mayors privately enjoy the gains from corruption. Politicians may also engage in corrupt activities to finance their own party. Because our data does not enable us to distinguish between these two motives, we abstract from the party component.

Voters. Residents vote for either the incumbent or a challenger on the basis of three factors.[10] The first is the incumbent's appeal during the election relative to the challengers, $\bar{\phi}_i$. We assume that mayors are endowed with different electoral appeal that voters observe, but we as econometricians do not. During the elections, residents are more likely to vote for politicians with a higher $\bar{\phi}_i$, all else equal. We assume that electoral appeal can take on two values, $\bar{\phi}_L$ and $\bar{\phi}_H$, and that the probability of high electoral appeal $\bar{\phi}_H$ is given by p_H.

Electoral appeal plays two important roles in our model. It allows us to capture the possibility that mayors with higher appeal steal more because their reelection probabilities are ex-ante higher. And, the differences in electoral appeal across mayors – along with the variation in ability and public funds – enable us to generate the observed heterogeneity in the fraction stolen by mayors.

The second factor that affects voters' decisions is the amount of adjusted per-capita public consumption $\bar{Q} = \frac{Q}{d^\eta}$ they expect the current mayor to produce in the next term relative to the challengers. To form these expectations, voters use all the information available at the time of the elections to predict the incumbent's ability level and, thus, the incumbent's future choices. The information available at the time of the elections includes the amount of public goods produced in the current term, whether the incumbent was audited in the current term and the fraction stolen if audited, the incumbent's characteristics, and the municipality's characteristics.[11] Formally, we model the amount of adjusted per-capita public consumption a voter expects to receive from the

---

[10]We do not model the decision to turnout because voting is mandatory in Brazil. On average, 85-90% of eligible voters vote during local elections.

[11]Since Brazil has a two-term limit for mayors, these politicians can run for reelection at most once. Consequently, voters have no information on the incumbent past the current term.

14

incumbent in the next term as follows:

$$E\left(\bar{Q}_{t+1} \,\middle|\, \bar{Q}_t, \delta_{au_t}, s_t, \bar{\phi}_i, X_{i,t}\right) = f\left(\bar{Q}_t, \delta_{au_t}, s_t, \bar{\phi}_i, X_{i,t}\right).$$

The third factor that affects voting is a voter's preference shock, $\varepsilon_{j,t}$, which we assume is normally distributed with mean zero and a normalized variance of one.

Given these three factors, resident j votes for incumbent i if the expected adjusted per-capita public consumption plus the electoral appeal and the voter's preference shock is above a threshold $\underline{Q}$, i.e.

$$f\left(\bar{Q}_t, \delta_{au_t}, s_t, \bar{\phi}_i, X_{i,t}\right) + \bar{\phi}_i + \varepsilon_{j,t} - \underline{Q} \ge 0,$$

or, equivalently,

$$\bar{f}\left(\bar{Q}_t, \delta_{au_t}, s_t, \bar{\phi}_i, X_{i,t}, \underline{Q}\right) \ge -\varepsilon_{j,t},$$

For the estimation, we assume that

$$\bar{f}\left(\bar{Q}_t, \delta_{au_t}, s_t, \bar{\phi}_i, X_{i,t}, \underline{Q}\right) = \bar{\xi}_1 + \bar{\xi}_2\bar{\phi}_i + \bar{\xi}_3\delta_{au_t} + \bar{\xi}_4\delta_{au_t}\delta_{\{s_t>0\}} + \bar{\xi}_5 \log\left(\frac{Q_t}{d_t^{\eta}}\right) + \bar{\xi}_6 age_t + \bar{\xi}_7\delta_{\{lm\}}$$

$$= \bar{\xi}_1 + \bar{\xi}_2\bar{\phi}_i + \bar{\xi}_3\delta_{au_t} + \bar{\xi}_4\delta_{au_t}\delta_{\{s_t>0\}} + \bar{\xi}_5 \log Q_t - \bar{\xi}_5\eta \log d_t + \bar{\xi}_6 age_t + \bar{\xi}_7\delta_{\{lm\}},$$

where $\delta_{\{s_t>0\}}$ is a dummy equal to one if the mayor has stolen a positive amount and δ_{lm} is a dummy for a municipality with more than 50,000 residents. To evaluate the anti-corruption policies, we only need to identify the parameter $\phi_i = \bar{\xi}_1 + \bar{\xi}_2\bar{\phi}_i$ for each i. We can therefore rewrite the function $\bar{f}$ as follows:

$$\bar{f}\left(\bar{Q}_t, \delta_{au_t}, s_t, \bar{\phi}_i, X_{i,t}, \underline{Q}\right) = \phi_i + \xi_1\delta_{au_t} + \xi_2\delta_{au_t}\delta_{\{s_t>0\}} + \xi_3 \log Q_t - \xi_3\eta \log d_t + \xi_4 age_t + \xi_5\delta_{\{lm\}}. \quad (1)$$

If individual i is elected we set the variable δ_{el} equal to 1.

This voting rule introduces the second main deterrent to corruption, which is also consistent with two of the empirical facts presented in Section 2: (i) mayors who are audited and caught stealing are less likely to be reelected and (ii) conditional on stealing, incumbents who produce more public consumption are more likely to win the election. The voters' choices will generate selection both on electoral appeal and on ability since high-appeal mayors are more likely to be

15

elected and more able mayors produce more public good, all else equal.

3.3 The Individual Decision Process

We now describe the decision process of individual i in municipality m. Individual i chooses private consumption, savings and, if the current mayor, stealing. These choices are made to maximize lifetime utility

$$E\left[\sum_{t=1}^{T} \beta^t \left(u\left(c_t^i, \bar{Q}_t\right) + \rho - \kappa\right)\right],$$

subject to the constraint that expenditure on private consumption plus savings must equal the available resources in each period and state of nature ω:

$$c_t^i + b_t^i = w_t^i \bar{h} + 1_{\{\delta_{el,t}^i=1\}} s_t^i + R_t b_{t-1}^i - 1_{\{\delta_{el,t-1}^i=1,\delta_{au,t-1}^i=1\}} g\left(s_{t-1}^i\right) \qquad \text{for each } t \text{ and } \omega,$$

where $1_{\{\delta_{el}^i=1\}}$ is an indicator function equal to one if individual i is the elected mayor in period t and $1_{\{\delta_{el,t-1}^i=1,\delta_{au,t-1}^i=1\}}$ is equal to one if, in period $t-1$, individual i was the mayor and the municipality was audited.

As mayor, individual i's decisions must satisfy two additional constraints. First, the resources stolen plus the resources invested in the production of public consumption must equal public funds in each period and state of nature:

$$z_t^{pu} + s_t^i = f_t^{pu} \qquad \text{for each } t \text{ and } \omega.^{12}$$

Second, the production function determines the amount of per-capita public consumption,

$$\frac{Q_t}{d_t} = \left(\frac{z_t^{pu}}{d_t}\right)^{\alpha_1} (z_t^{pr})^{\alpha_2} \exp\left\{\alpha_3 + \alpha_4 e_i + \sum_{j=1}^{N_d} \alpha_{4+j} (d_t)^j\right\} a_i \qquad \text{for each } t \text{ and } \omega.$$

The current mayor also decides whether to run for reelection at the end of the term.

The sources of uncertainty faced by the residents of a municipality depend on whether they are a mayor. Mayors face uncertainty in the amount of funds the municipality will receive from the

---

[12]We do not model local taxes because in Brazil 85 percent of a municipality's funds are transfers from the central government.

central government, whether the municipality will be audited, and the voters' preference shocks in case they run for reelection. Non-mayors only face uncertainty in wages and the amount of public consumption produced by the current mayor.

Our corruption data likely contains measurement error. Thus, we introduce a measurement error ν_t^i in stealing that we draw from a normal distribution with mean 0 and standard deviation σ_ν. The stealing we observe in the data is: $s_t^{i,o} = s_t^i + \nu_t^i$.

3.4   Timing of the Model

The model's timing is as follows. At the beginning of term t, the federal government audits a fraction of municipalities at random, discloses the results publicly, and then collects fines based on the audit outcomes from the mayors that were caught stealing. The current mayor, who has a given electoral appeal and ability, then decides whether to run for reelection based on the quantity of public goods produced in the previous term and, if audited, on the fraction of public funds stolen. Elections take place. Incumbents who win the election will govern in term t. If an incumbent chooses not to run or loses the election, a new mayor is selected from the distribution $f_m(Z)$ of mayor' characteristics, where Z includes the mayor's age, education, savings, ability, and electoral appeal. Wages, public funds, and private inputs are then realized. Lastly, the elected mayors choose consumption, savings, the fraction of public funds to invest in public consumption, and the fraction to steal for their personal use.

3.5   Recursive Formulation

We solve and estimate the model using its recursive formulation. It requires computing two value functions: one for current mayors, V_M, and one for past mayors, V_{PM}, which is needed to compute the value function of current mayors. Also, to compute the value function of ex-mayors for term t, we must know the amount of per-capita public consumption produced by the current mayor. This requires that we compute the value functions of past mayors for each potential incumbent.

To derive the recursive formulation, let S_t^M and S_t^{PM} be the set of state variables at time t for current and ex-mayors. In the data, only 3 percent of past mayors have run for election after leaving office for at least one term. We therefore assume that individuals can be mayor only once

in their life. We can then write the decision problem of an ex-mayor for term t as:

$$V_{PM}^i\left(S_t^{PM},t\right) = \max_{c_t^i,b_t^i}\ u^i\left(c_t^i,\bar{Q}_t\right) + \beta E\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\right]$$

$$s.t.\ \ c_t^i + b_t^i = w_t^i\bar{h} + R_t b_{t-1}^i - 1_{\{\delta_{el,t-1}^i=1,\delta_{au,t-1}^i=1\}}g\left(s_{t-1}^i\right),$$

where the wage is drawn from the wage distribution of ex-mayors $p_w\left(w|Z\right)$ and the dummy $1_{\{\delta_{el,t-1}^i=1,\delta_{au,t-1}^i=1\}}$ indicates that ex-mayors can be fined at t only if they were in power in the previous term and were audited.

The decision problem of a current mayor is more complicated because it includes the choice to run for reelection at the end of the term. If the mayor decides to run, he wins with a probability $p\left(S_t^M\right)$. Let the value function of a winning incumbent be $V_{WM}^i\left(S_t^M,t\right)$. Then, we can write the corresponding decision problem as:

$$V_{WM}^i\left(S_t^M,t\right) = \max_{c_t^i,b_t^i,s_t^i}\ u^i\left(c_t^i,Q_t\right) + \rho - \kappa + \beta E\left[V_M^i\left(S_{t+1}^M,t+1\right)\right]$$

$$s.t.\ \ c_t^i + b_t^i = w_t^i\bar{h} + s_t^i + R_t b_{t-1}^i - 1_{\{\delta_{el,t-1}^i=1,\delta_{au,t-1}^i=1\}}g\left(s_{t-1}^i\right)$$

$$z_t^{pu} + s_t^i = f_t^{pu}$$

$$\frac{Q_t}{d_t} = \left(\frac{z_t^{pu}}{d_t}\right)^{\alpha_1}(z_t^{pr})^{\alpha_2}\exp\left\{\alpha_3 + \alpha_4 e_i + \sum_{j=1}^{N_d}\alpha_{4+j}\left(d_t\right)^j\right\}a_i,$$

where w_t^i is the mayor's wage. With probability $1 - p\left(S_t^M\right)$ the challenger wins the election, in which case the mayor's value function corresponds to the value function of an ex-mayor. We can therefore compute the value function of an incumbent that chooses to run for reelection, $V_{RM}^i\left(S_t^M,t\right)$, as:

$$V_{RM}^i\left(S_t^M,t\right) = p\left(S_t^M\right)V_{WM}^i\left(S_t^M,t\right) + \left(1 - p\left(S_t^M\right)\right)V_{PM}^i\left(S_t^{PM},t\right).$$

If the mayor decides to forgo reelection, the corresponding value function is equal to that of an

ex-mayor, $V_{PM}^i \left(S_t^{PM}, t \right)$. Altogether, a mayor chooses to run for reelection if

$$V_{RM}^i \left(S_t^M, t \right) \geq V_{PM}^i \left(S_t^{PM}, t \right).$$

To estimate the model, we assume that the decision to run is also affected by a shock $\epsilon_R \sim N(0, \sigma_R)$, where σ_R is normalized to 1. A mayor chooses to participate in the election if

$$V_{RM}^i \left(S_t^M, t \right) \geq V_{PM}^i \left(S_t^{PM}, t \right) + \epsilon_R.$$

The value function of the current mayor is:

$$V_M^i \left(S_t^M, t \right) = \max \left\{ V_{RM}^i \left(S_t^M, t \right), V_{PM}^i \left(S_t^M, t \right) + \epsilon_R \right\}.$$

The state variables S_t^M for current mayors include the number of terms the individual has been in power, the population size of the municipality, the mayor's age and education, the mayor's ability and electoral appeal, the amount of public goods produced and the amount stolen by the mayor in the previous period, the mayor's savings, the probability that the municipality will be audited, and the amount of public funds and private inputs the municipality receives. The state variables S_t^{PM} for ex-mayors include the population size of the municipality, their age and education, whether they were audited and stole in the past, their savings, and the amount of adjusted per-capita public good produced by the current mayors.

4   Identification and Estimation

In this section, we present formal arguments for the identification of the model's parameters. Specifically, we show that under our functional form assumptions, only one set of the model's parameters corresponds to the distribution of available data. We then discuss how the identification results and the data can be used to estimate the corresponding parameters. We estimate all of the model's parameters with the exception of the curvature of the utility function δ and the discount factor β. Following the literature, we set $\delta = 2$ (e.g. Attanasio and Weber (1995)) and $\beta = 0.98$ (e.g. Attanasio, Low, and Sánchez-Marcos (2008)).

Electoral Decision Parameters: ϕ_L, ϕ_H, $\xi_1, \ldots, \xi_5$, p_H.

Identification.   Under the assumption that the mayor's electoral appeal takes on two values and that the electoral shocks are normally distributed, the probability of being reelected conditional on running has a mixed normal distribution, with the two mixing distributions having different means that depend on the electoral appeal. Specifically, let p_H^c be the probability that a mayor has high appeal conditional on running for reelection and let Φ denote the standard normal cumulative density function. Given the electoral rule described in Section 3, the likelihood function for the electoral outcomes has the following form:

$$
\begin{aligned}
L &= \Pi_{i=1}^N \left[(1 - p_H^c) \, \Phi \left(\phi_L + \xi_1 \delta_{au_t^i} + \xi_2 \delta_{au_t^i} \delta_{\{s_t^i > 0\}} + \xi_3 \log Q_t^i - \xi_3 \eta \log d_t^i + \xi_4 age_t^i + \xi_5 \delta_{\{lm\}}^i \right) \right. \\
&\quad + \left. p_H^c \Phi \left(\phi_H + \xi_1^i \delta_{au_t^i} + \xi_2 \delta_{au_t^i} \delta_{\{s_t^i > 0\}} + \xi_3 \log Q_t^i - \xi_3 \eta \log d_t^i + \xi_4 age_t^i + \xi_5 \delta_{\{lm\}}^i \right) \right],
\end{aligned} \tag{2}
$$

where N denotes the number of mayors. The maximum likelihood estimator for a normal mixture model with equal variances is consistent (Kiefer and Wolfowitz (1956), Redner (1981), and Basford and McLachlan (1985)). The parameters defining the reelection probability in (1), ϕ_L, ϕ_H, $\xi_1, \ldots, \xi_5$, and the probability of high electoral appeal conditional on running are therefore identified. This result relies on the fact that we can observe which mayors were audited and the amount they diverted.

We require an additional step to identify the unconditional probability of high electoral appeal p_H. This is the focus of the next proposition. It shows that the conditional probability of high electoral appeal is a strictly increasing function of p_H. Thus, because the conditional probability is known, the parameter p_H is also identified.

Proposition 1 *The unconditional probability that a mayor has high electoral appeal p_H is identified if the probability of high electoral appeal conditional on running is observed.*

Proof. In the online appendix. ∎

Estimation.   Following the identification discussion, we estimate all the electoral appeal parameters, except the unconditional probability of high electoral appeal, by maximizing the likelihood function in (2) using data on electoral outcomes, audits, amount stolen, public consumption,

population size, and age. We estimate the unconditional probability of high electoral appeal by Simulated Method of Moments (SMM), jointly with the other parameters discussed below, using as a moment the conditional probability of high electoral appeal.

Production Function and Ability Parameters: $\alpha_1, \ldots, \alpha_{4+N_d}$, μ_a, σ_a.

Identification. To evaluate the policies, we only need to identify the sum of the expected value of ability μ_a and the constant in the production function α_3. We thus normalize $\alpha_3 = 0$. Under this normalization, we will show that we can identify the production function parameters and the parameters of the ability distribution, as long as we observe the variables in the production function for first-term mayors.

The log of the production function of first-term mayors has the following form:

$$\log \frac{Q_t}{d_t} = \alpha_1 \log \frac{z_t^{pu}}{d_t} + \alpha_2 z_t^{pr} + \alpha_4 e_i + \sum_{j=1}^{N_d} \alpha_{4+j} (d_t)^j + \log a_i, \tag{3}$$

where the public inputs z_t^{pu} depend on the mayor's stealing and thus ability. Let z_t be a vector of variables that are mean independent of and, hence, uncorrelated with a_i, but are correlated with z_t^{pu}, and let $Z_t = [z_t^{pr}, e_i, d_t, z_t]$. Taking the expectation conditional on Z_t of equation (3), we obtain:

$$E\left[\log \frac{Q_t}{d_t} \middle| Z_t\right] = \alpha_1 E\left[\log \frac{z_t^{pu}}{d_t} \middle| Z_t\right] + \alpha_2 z_t^{pr} + \alpha_4 e_i + \sum_{j=1}^{N_d} \alpha_{4+j} (d_t)^j + E\left[\log a_i \middle| Z_t\right].$$

Since Z_t is mean independent of a_i, we have that $E\left[\log a_i \middle| Z_t\right] = E\left[\log a_i\right] = \mu_a$, and the previous equation becomes:

$$E\left[\log \frac{Q_t}{d_t} \middle| Z_t\right] = \alpha_1 E\left[\log \frac{z_t^{pu}}{d_t} \middle| Z_t\right] + \alpha_2 z_t^{pr} + \alpha_4 e_i + \sum_{j=1}^{N_d} \alpha_{4+j} (d_t)^j + \mu_a.$$

Thus, the parameters α_1, α_2, $\alpha_4, \ldots, \alpha_{4+N_d}$, and μ_a are identified if the variables $\log \frac{Q_t}{d_t}$, $\log \frac{z_t^{pu}}{d_t}$,

z_t^{pr}, e_i, d_t, and z_t are observed. Lastly,

$$\log \frac{Q_t}{d_t} - E \left[\log \frac{Q_t}{d_t} \middle| Z_t \right] = a_i - \mu_a.$$

The variance of a_i can therefore be identified as $E \left[\left(\log \frac{Q_t}{d_t} - E \left[\log \frac{Q_t}{d_t} \middle| Z_t \right] \right)^2 \right]$. Analogously to the electoral rule parameters, this identification result requires knowledge of the amount stolen by politicians.

Estimation.   We estimate the parameters α_1, α_2, $\alpha_4, \ldots, \alpha_{4+N_d}$, and μ_a by running a two-stage least square regression of $\log \frac{Q_t}{d_t}$ on $\log \frac{z_t^{pu}}{d_t}$, z_t^{pr}, e_i, and d_t for first term mayors. As the identification argument indicates, the estimation of the production function parameters requires knowledge of the amount diverted by mayors. We therefore restrict the sample to mayors that were audited. This sample selection does not invalidate the identification result because municipalities were randomly audited during lotteries that were nationally televised. Without the randomization, the estimates of the production function parameters would only apply to the selected sample.

We instrument the endogenous variable $\log \frac{z_t^{pu}}{d_t}$ using population thresholds that correspond to discrete changes in the amount of funds municipalities receive from the federal program called the *Fundo de Participação dos Municípios* (FPM). The FPM program is an automatic, formula-based transfer scheme that accounts for almost 80% of federal transfers. Because the amount of federal funds municipalities receive from this program varies discontinuously according to a municipality's population, we can use the population thresholds specified by the FPM formula to identify the causal effects of public inputs using a fuzzy regression discontinuity approach. Other studies have also used this identification strategy as an exogenous source of public spending (e.g. Brollo et al. (2013), Corbi, Papaioannou, and Surico (2018)). We use the residuals of this regression to estimate the variance of ability.

Wages of Past Mayors Parameters: $\gamma_0, \ldots, \gamma_5$ and σ_{pm}.

Identification.   The wage process of past mayors is assumed to be linear and, based on the empirical evidence, is independent of past stealing. The parameters $\gamma_0, \ldots, \gamma_5$ and σ_{pm} are identified if we observe wages, experience, age, and municipality size.

Estimation. We estimate these parameters by OLS.

Fine Parameters: μ_τ, σ_τ.

Identification. We have defined the fine variable, τ, as a multiple of the amount stolen and assumed that it is distributed as $\log \tau \sim N(\mu_\tau, \sigma_\tau)$. Thus, if the actual fine and the amount stolen are observed, the variable τ is also observed. We can identify the parameters μ_τ and σ_τ using the mean and standard deviation of the log of the observed variable τ.

Estimation. We estimate the parameters μ_τ and σ_τ using the sample mean and standard deviation of the observed $\log \tau$.

Distribution of Mayor's Characteristics: $f_m(Z)$.

Identification. The vector of mayor's characteristics Z includes the mayor's age, education, savings, ability, and electoral appeal. We assume that the mayor's ability and electoral appeal, whose distributions we have shown are identified, are independent of the remaining characteristics. We can then identify non-parametrically the joint distribution of age, education, and savings, as we observe those variables.

Estimation. We use a bin estimator to estimate the joint distribution of age, education, and savings.

Relative Taste for Public Consumption Parameter: θ.

Identification. We can identify the taste for public consumption parameter using the stealing decisions of second-term mayors, as we establish in the following Proposition.

Proposition 2 *The relative taste for public consumption θ is identified by the expected value of stealing for second-term mayors.*

Proof. In the online appendix. ∎

To provide the intuition behind the previous Proposition, note that three factors affect a second-term mayor's decision to engage in corruption: (i) the productivity of public inputs in

the production of public consumption, (ii) the severity of the fine schedule, and (iii) the relative taste for public consumption θ. We have shown that the parameters determining the first two factors are separately identified. Given those parameters, Proposition 2 establishes that stealing of second-term mayors decreases monotonically with the taste for public consumption: as mayors care more about public consumption, they reduce the amount of funds they divert.[13]

Estimation. We estimate the parameter θ by SMM using as a moment the expected value of stealing by second-term mayors.

Utility Cost of Running and from Being in Power Parameters: κ, ρ.

Identification. We can identify the utility cost of running parameter, κ, and the utility from being in power parameter, ρ, using the decision to run for two types of mayors: mayors who were audited and not found to be corrupt; and mayors who were audited and found to be corrupt. We formalize this result in the next Proposition.

Proposition 3 *The parameters κ and ρ are identified by the decision to run for reelection of audited incumbents who were not found to be corrupt and audited incumbents who were found to be corrupt.*

Proof. In the online appendix. ∎

For the intuition behind Proposition 3, note that incumbents running for reelection pay the cost of running κ regardless of the electoral outcome. This is not the case for the utility from being in power, as only elected mayors enjoy ρ. Thus, incumbents who have a higher probability of winning are more likely to run and more likely to experience ρ. This is the case for audited incumbents who were not found to be corrupt. All else equal, they have a higher probabilities of winning and thus running than audited mayor who were found to be corrupt. Therefore, a larger

---

[13]In our model, corruption differences between first-term and second-term mayors are generated by electoral incentives, selection, and optimal decisions (see Section 5.3 for a decomposition of the three components). The differences could also be generated by learning how to engage in corruption over the course of the first term. Although we can add learning to our model, we cannot separately identify it from electoral incentives given the available data. Moreover, Ferraz and Finan (2008) do not find any evidence of learning in their reduced-form comparison of first versus second-term mayors. Therefore, we have decided to abstract from it. The implicit assumption is that politicians learn quickly over the first few months of the first term and fully engage in corruption for the rest of the term. If this assumption is violated, the effect of learning will be combined with electoral incentives.

ρ increases the difference in the likelihood of running between these two groups of incumbents, independent of κ. We can therefore identify ρ by the difference in running probability between audited incumbents caught and not caught stealing. Given ρ, the probability of running for any incumbent declines monotonically with κ. We can therefore identify this parameter by the probability of running of audited incumbents that were not found to be corrupt.

Estimation. We estimate the parameters ρ and κ by SMM using as moments the observed probability of running for reelection for audited mayors who were not found to be corrupt and the difference between this probability and the corresponding probability for audited incumbents who were found to be corrupt.

Standard Deviation of the Measurement Error: σ_ν.

Identification The following Proposition establishes that we can identify σ_ν using the probability that a mayor chooses not to divert public funds.

Proposition 4 *The variance of the measurement error in stealing σ_ν is identified by the probability of observing stealing equal to zero.*

Proof. In the online appendix. ∎

The idea behind the Proposition is that once all the other parameters have been identified, the probability of observing no stealing for an audited mayor is an increasing function of the parameter σ_ν. Thus, there exists a unique value of σ_ν for a given value of the probability of observing zero stealing.

Estimation. We estimate the parameter σ_ν by SMM using as a moment the probability of observing zero stealing.

Estimation by Simulated Method of Moments

We estimate the parameters p_H, θ, κ, ρ, and σ_ν jointly using dynamic programming and the SMM (Gourieroux and Monfort (1996)). We do this in two steps. For a given set of model parameters, we simulate the individual decisions. We then compute in the data and in the simulations the

moments used in the estimation of the parameters and calculate the distance between them. The estimated parameters are obtained by minimizing the distance function $(m_d - m_s)' \Sigma (m_d - m_s)$, where m_d is the vector of data moments, m_s the vector of simulated moments, and Σ the inverse of the variance-covariance matrix of the moments.

In the simulations, we compute the value functions for each individual starting from the last period and proceeding backwards in two steps following Keane and Wolpin (1994). In the first step, we discretize the state space and then compute the expected value functions $E[V|S]$ for each period and point of the state space in the grid. In the second step, we approximate the expected value functions for each point of the state space using non-parametric methods. In practice, we regress the values of $E[V|S]$ obtained for each point in the grid on a polynomial of the discretized state variables. We then use the corresponding coefficients to construct the expected value functions for each period and value of the state space. Once the expected value functions are known, we can simulate the decisions of individuals in the municipalities observed in the data for different values of the model parameters. We repeat these steps until we have minimized the distances between the data and simulated moments.

5   Results

In this section, we present our estimation results, the model's fit, and a comparison of the effects of different anti-corruption policies.

5.1   Parameter Estimates

We begin with the parameter estimates of the electoral rule, presented in Table 3. We document that the effect of an audit on the reelection probability of an incumbent mayor depends on what it reveals. For mayors who are found to be corrupt, the audit decreases the probability of winning by 16.1 percentage points (coefficient=-0.554, s.e.= 0.274) relative to those who were audited and not found to be corrupt. Compared to the mayors who were not audited, the mayors who were not found to be corruption also enjoy a potential electoral advantage. An audit for these mayors increases their probability of winning by approximate 10 percentage points, although the effect is not measured with much precision (coefficient = 0.328; s.e.=0.239).

Voters value public consumption and reward incumbents accordingly. Our estimates suggest that an increase in public consumption of 10% from the mean produces a 16 percentage points increase in reelection probabilities. Unobserved heterogeneity is also important in our model. We estimate that 64 percent of mayors have high electoral appeal (Table 6) and that, conditional on running, these mayors are 65.8 percentage point more likely to be reelected than those with low electoral appeal. For the mean and variance of the fine distribution, we estimate coefficients of 0.094 and 0.284, respectively. These estimates imply that corrupt mayors, on average, have to repay the original amount and pay a fine equal to the amount stolen.

Table 4 reports the estimated coefficients for the production function. We estimate that a 10% increase in public inputs increases public consumption by 1.5%; whereas, a 10% increase in private inputs increases public consumption by 4.3%. Having a mayor with a college degree has a positive effect on production, but the coefficient is statistically insignificant. Another source of unobserved heterogeneity is the mayor's ability for producing public goods. We estimate that the mean of log-ability is -0.294 in the first term and -0.283 in the second term, suggesting that second-term mayors are more able than first-term mayors on average. Although these point estimates are not measured with much precision, their magnitudes are economically meaningful: a 0.1 standard deviation increase in mean ability increases public consumption by 4.1 percent. The difference in coefficients between first and second-term mayors implies an increase in per-capita public consumption of 1.1 percent. The standard deviation of log-ability equals 0.442.

In Table 5 we report the estimated coefficients for the wage process of past mayors. As expected, wages are positively associated with years of schooling and exhibit an inverted u-shape with respect to age. Past mayors also have higher wages in municipalities with larger populations.

In Table 6, we report our estimates of the preference parameters. With a relative taste for public consumption equal to 0.161, we estimate that individuals value the utility from private consumption about 7 times more than the utility from public consumption. We estimate a relatively low cost of running for reelection at 0.027, which corresponds to 2.7 percent of the average utility a mayor enjoys in one term. Given the large fraction of mayors who run for reelection, we estimate the utility from being in power to be quite large at 0.591, which corresponds to 59.1 percent of the average utility a mayor experiences from private and public consumption. We only need a small standard deviation for the measurement errors (0.063) to explain the corruption data.

5.2 Model Fit

Fit of Moments Used in the SMM Estimation. In the SMM estimation, we match six moments: (i) the average fraction of funds diverted by first-term mayors; (ii) the average fraction of funds diverted by second-term mayors; (iii) the fraction of audited mayors who did not steal and chose to run for reelection; (iv) the fraction of audited mayors who stole and chose to run for reelection; (v) the fraction of incumbents with high electoral appeal who ran for reelection; and (vi) the fraction of audited mayors who stole.

In panel A of Table 7, we compare the actual and simulated values of these moments. Our model matches these moments well. It can reproduce exactly the actual fraction stolen by both first-term mayors (5.59%) and second-term mayors (7.31%). This is important because the second moment identifies the relative taste for public consumption. We can also match well the fraction of audited incumbents who did not steal and chose to run – 75.60% in the model versus 77.92% in the data – and the corresponding fraction for audited incumbents who stole – 69.39% in the model versus 68.39% in the data. This is noteworthy because the second moment and the difference between these two moments allow us to identify the utility cost of running for reelection and the utility from being in power. We identify the probability of a high electoral appeal mayor using the observed fraction of incumbents who participate in the election with high appeal and our model matches this moment well – 90.54% in the model and 89.90% in the data. The last moment – the fraction of audited mayors who stole – is included in the estimation to identify the standard deviation of the measurement errors. It is therefore reassuring that the simulated moment (73.14%) matches the data moment (73.32%).

Fit of Moments Not Used in the Estimation. In panel B of Table 7 we assess how well our model matches moments not used in the estimation. We compare 8 moments and we generally do quite well. For example, we are able to match closely the share of incumbents who forgo reelection (29.37% versus 28.18%), the probability of winning reelection (58.89% versus 57.26%), and the probability of winning conditional on having been audited and caught stealing (51.73% versus 51.11%).

We also test the ability of our model to match mean log-ability for second term mayors. In the model, mean log-ability of first-term mayors is set equal to mean log-ability in the data by

28

normalizing the constant in the production function to be equal to zero, as discussed in Section 4. By contrast, mean log-ability for second-term mayors is a combination of mean-log ability for first-term mayors and the selection generated by the electoral decisions. We compute it by calculating mean log-ability in the simulations for incumbents that win a second term. In the data, mean log-ability for second-term mayors is measured by the constant in the estimation of the production function for second-term mayors. Despite the fact that we compute ability differently in the model versus in the data, we find that the model does relatively well in matching this moment, with a simulated mean of -0.245 versus a mean in the data of -0.283.

In Section 2, we documented the right skewness of the stealing distribution. The last part of Table 7 shows that our model, despite its parsimony, can generate this skewness. This is because of the interaction between a mayor's decision to steal and the production function for public consumption. In our production function, the return of one dollar invested in the production of per-capita public consumption depends on the amount of public funds received by the municipality, with a lower (higher) return for a larger (smaller) amount. All else equal, mayors will therefore divert more resources when the municipality receives more public funds, which is why the distribution of corruption generated by the model inherits the right skewness we see in public funds. Note that the data do display a slightly higher degree of right skewness. Increasing the degree of skewness in the model would require additional heterogeneity in the relative taste for public consumption. We decided against introducing this because it would make the identification of the model parameters less transparent.

5.3   Model Simulations

Elections can play two important roles in promoting voters' welfare. First, if voters care about public goods, then politicians can potentially improve their reelection chances by refraining from stealing and providing more public goods. This is often referred to as a *reelection incentive effect*. Second, elections allow voters to select more able politicians, which is referred to as a *selection effect*. As the literature has emphasized, these two effects can explain why in the data second-term mayors steal more than first-term mayors. But in our model, politicians can also save and if second-term mayors accumulate more wealth than first-term mayors, this can provide yet a third reason for the difference in stealing between first and second-term mayors. A key feature of our

model is the ability to separately identify the effects of all three mechanisms.

In Table 8, we simulate our baseline model and compute the average of various variables of interest, distinguishing between first and second-term mayors. For these baseline simulations, we use an audit probability of 5% in all periods in order to have a clean comparison across terms. In the estimation, we had used the audit probability observed in the data, namely 5% until 2001 and 16.8% afterward.

We find that second term mayors steal about 4 percentage points more than first-term mayors (12.9% versus 8.8%).[14] Public consumption is higher among second-term mayors, despite the fact that they steal more. This is a strong indication of the positive selection that elections can induce, and we can see this more clearly in the subsequent rows. Second-term mayors are positively selected along various observable and unobservable characteristics. For example, second term mayors tend to be wealthier, younger, and more educated. They are also significantly more able and have higher electoral appeal.

We can attribute the differences in stealing and public consumption between first and second-term mayors to reelection incentives, selection effects, and a potential savings effects. To identify the effect of each channel, we simulate the model under an environment in which re-election is not possible. We can identify the effects of *reelection incentives* by comparing the decisions of first term mayors under an environment with and without the possibility of reelection. To measure *selection effects*, we compare the decisions of second-term mayors to those of first-term mayors under a no-reelection regime, while also holding savings decisions constant. Finally, to identify *savings effects*, we again compare second-term mayors to first-term mayors under a no-reelection regime, but holding constant their observable and unobservable characteristics.

We decompose these effects for both stealing and per-capita public consumption in Table 9. We find that reelection incentives explains 42.8% of the difference in stealing between first and second-term mayors. Selection effects account for a negative 32.3%, in the sense that had the elections not induced a positive selection of politicians, corruption would have been 32.3% higher. The remaining difference between second and first term mayors can be attributed to savings decisions. Because corruption enables mayors to accumulate wealth over time – both for reasons of consumption smoothing and precautionary savings in case they are caught – second-

---

[14]These numbers are higher than those reported for the model's estimation because of the different audit probabilities.

term mayors are on average wealthier than first-term mayors. Moreover, the effectiveness of the financial punishment associated with being caught stealing in the future is lower, the wealthier the politician. Because our model captures these various saving dynamics, it can decompose the difference in corruption between first and second-term mayors due to savings, which at 89.5% turns out to be a non-trivial fraction.

On the other side of the coin, we have per-capita public consumption. We see that reelection incentives accounts for a negative 5%. In other words, without reelection incentives the difference in consumption between first and second-term mayors would be 5% larger. The majority of the difference can be attributed to a selection effect at 74.7%. This result makes sense given both the importance of observable and unobservable characteristics for public consumption and how positively selected second-term mayors are relative first-term mayors. Savings decisions account for 30.3% of the difference.

5.4   Policy Evaluation

An important contribution of our model is that it allows us to assess several anti-corruption policies all within the same setting. We are able to not only compare across different policies, but also combine them. This is important because, as we will show, each individual policy has its weaknesses, and by combining certain policies, we can mitigate these limitations to reduce corruption further.

In this section, we use our estimated model to evaluate five anti-corruption policies: 1) an increase in the audit probability; 2) a clean record policy; 3) an increase in the audit probability for mayors who have been charged (but not convicted) with corruption in the past; 4) allowing mayors to be elected for a third term; and 5) a doubling of mayors' wages. We introduce each of these policies to the base case, which sets the audit probability to 5% in all periods. In the subsequent section, we then show how some these policies can be combined to achieve greater reductions in corruption.

Increase in the audit probability (audit policy).   In 2003, Brazil's Federal Government introduced a program with the explicit goal of reducing corruption in local governments. As we discussed in Section 2, the program began to audit municipalities at random for their use of federal

funds. This program increased the probability that mayors would be audited within their term to approximately 16.8 percent. We can evaluate the effects of this program on subsequent corruption by simulating this increase in the audit probability in all periods.

In Figure 1, we see that an increase in the probability of being audited reduces corruption substantially. Our simulations suggest that by tripling the audit probability a mayor faces, the program has reduced corruption by about 4.3 percentage points, which represents a 41.4% reduction from our base case. The reduction in stealing is large for both first and second-term mayors alike (see second panel). Even though second-term mayors still steal more than first-term mayors, their relative reduction was higher (5.96 pp versus 3.30 pp). These results suggest that a higher audit probability can partially substitute for the disciplining effects that reelection incentives play.

While a higher audit probability can be effective at reducing corruption, it also happens to be quite expensive. Over the past 12 years, the program has audited approximately 180 municipalities per year at an annual budget of R$80 million. This implies an average costs of R$445,000 per audit.

Clean Record Policy. In 2010, Brazil established the Clean Record Act (*Lei da Ficha Limpa*), which prohibits individuals who have been convicted of corruption from holding public office for eight years. The objective of the law is to raise the implicit cost of corruption and thus to incentivize political candidates to abstain from stealing. In our model, we can simulate the effects of this policy by prohibiting incumbents who have been caught stealing from participating in future elections.

In Figure 1, we see that the Clean Record Act (CRA) reduces corruption by 12.2 percent. The policy affects first-term mayors (17%) more than second-term mayors (6%). This is expected because the CRA policy works via electoral incentives, which are stronger for first-term mayors than second-term mayors. We can see this clearly when we compare the effects of the policy on mayors who run versus those who do not. In the base case, first term mayors who run steal 8.25 percent of the funds. Under the CRA policy, this number declines substantively to 5.25. In contrast, among the first-term mayors who forgo reelection, the CRA policy reduces corruption only slightly from 12.95 to 12.65. The policy's effect on second-term mayors is limited to a small 0.73 percentage point decline in corruption, which is driven only by selection effects.

Increasing the audit probability for mayors with past charges (audit-if-caught policy).
The CRA only applies to individuals who have been convicted of wrongdoing. But it is common in Brazil and other developing countries, for corrupt politicians to remain in office while their cases proceed through the courts, which can sometimes last for decades. One way to combat these delays would be to implement a policy whereby mayors who have been charged with wrongdoing face a higher audit probability. In doing so, we are targeting mayors that have shown a higher propensity to engage in corruption, which is an unobservable trait. To simulate such a policy, we increase the audit probability to one for mayors who had been caught stealing in the past but won reelection.

In Figure 1, we see that this policy reduces corruption, but only by 0.31 percentage points. This is because the policy only affects mayors who have been caught stealing and reelected – a low-probability event – given an audit probability of 5 percent and the reelection chances of corrupt mayors. The policy does reduce corruption for both first and second-term mayors, but for different reasons. First-term mayors steal less because if they do, the value of being in power in the second term declines due to the increase in audit risk. The reduction in corruption among second-term mayors comes from the reelected politicians who were caught stealing in the first term and face an audit with certainty during their second term.

3-term limit policy. If reelection incentives discourage mayors from stealing and elections enable voters to select better mayors, then increasing the number of terms a mayor can stay in power may help to reduce corruption. In Figure 1, we simulate the effects of allowing mayors to be reelected for a third term. We find that the possibility of a third term reduces corruption by 23%. This reduction comes mostly from second-terms mayors who now have reelection incentives and reduce their stealing by 5.66 percentage points relative to the base case. As expected, corruption in the final term is much higher than in the other two terms. But interestingly, the amount of corruption in the final term of the 3-term limit is much lower than in the final term of a 2-term limit, a difference that is due to selection. Compared to second-term mayors, those that get reelected to a third term are more positively selected in terms of having a college degree (33.7 vs 35.1) and their unobserved ability (0.84 vs 0.88).

Doubling wages. The last individual policy we consider is to double the mayor's salary. The rationale for such a policy is simple. If we increase the value of holding office, this will increase the incentives to being reelected, which in turn should discourage mayors from stealing. But in Figure 1, we see the exact opposite result: corruption actually *increases* from 10.36% to 10.64%. The explanation makes sense once we compare the stealing levels of mayors who run for reelection versus those that do not. When we condition on running for reelection, first-term mayors under the wage policy steal less than first-term mayors in the base case (8.19% vs 8.34%), which is in line with our ex-ante expectations. The policy's increase in overall corruption comes from the mayors who choose not to run. For these mayors, higher wages weakens the deterrent effect of financial punishments and increases stealing. This result indicates that wage policies are less effective when we account for the endogeneity of running for reelection.

5.5 Policy Bundles

We have considered several policies that successfully reduce corruption. But they are not without their limitations. For example, the policy to increase the audit probability was our most effective policy at reducing corruption, but it is also expensive. The CRA policy reduced corruption among mayors with electoral incentives, but only had a minimal effect on term-limited mayors. The same limitation applies to our 3-term limit policy. In Table 10, we summarize the limitations of each policy and suggest how we might combine these policies to increase the effectiveness of the anti-corruption reforms. Importantly, we can use our model to simulate the effects of these combined policies on corruption. We present these results in Figure 2.

The 3-term limit policy has two shortcomings. It has modest effects on first-term mayors and no direct effect on incumbents in their last term. To solve the first limitation, we can simply combine it with a policy that affects mayors who are not term limited before the policy, such as the CRA policy. When we combine these two policies, we reduce corruption to a level that is below the more expensive audit-probability policy. Stealing declines to 5.77%, with most of the decrease occurring in the first and second terms. First-term mayors reduce their fraction stolen from 8.39% to 6.55% and second-term mayors from 7.19% to 4.42%, relative to the basic 3-term policy. This policy combination also has a small effect on the last term through its impact on selection.

To address the second limitation of the 3-term limit policy, we can combine it with an increase in the audit probability for mayors in their last term. When we increase the audit probability for last-term mayors from the base value of 5% to 16.8%, corruption in the third term declines from 10.8% in the basic 3-term policy to 6.28%. But stealing in the second term increases from 7.19% in the basic 3-term policy to 7.96%, as politicians substitute corruption toward the term with the lower audit probability. Because of this response, the overall effect on corruption of this reform is about the same as the basic 3-term policy with 8.05% of public funds diverted. This result indicates that increasing the audit probability for just one term is not effective.

If we combine all three policies, we achieve an overall effect that is slightly larger than the impact of the 3-term-CRA policy, with a drop in the fraction stolen to 5.7%. The decline is noticeable in all terms, as the fraction stolen decreases to 6.2% in the first term, 4.83% in the second, and 6.11% in the third. With this policy, mayors can still substitute corruption to terms with lower audit probability, which reduces the policy efficacy. But, the CRA components of the policy limits how much politicians are willing to substitute, generating the observed decline in corruption.

The main drawback of the audit-probability policy is its costs. Thus, the policy needs to be as effective as possible. One option is to pair it with a policy that has lower costs and can increase its efficacy, such as auditing mayors that have been caught in the past. This combination reduces corruption by an additional 0.74 percentage points (5.33% versus 6.07%).

The CRA policy affects only mayors with electoral incentives. To overcome this limitation, we can increase the audit probability for mayors in their last term to 16.8%. This change substantially improves the CRA's effects on corruption, further reducing stealing by about 1.89 percentage points. Much of the reduction occurs for second-term mayors whose stealing is cut almost in half.

The findings of this section indicate that the performance of popular anti-corruption policies can be greatly improved by pairing them with other interventions that have the potential to counter their main weaknesses. When we use this strategy, we find that the most effective intervention with no direct cost combines the 3-term and the CRA policies.

5.6   Willingness to Pay for the Policies

So far we have analyzed the effects of the anti-corruption policies on the fraction of public funds stolen. But we might also be interested in how these policies affect individual welfare or, equivalently, an individual's willingness to pay for the policies. These effects are necessary for determining whether it is beneficial to implement a policy. Importantly, our model allows us to calculate an individual's willingness to pay for each of our policies.

We measure the willingness to pay by computing the reduction in initial wealth that makes a person living in a municipality indifferent between having an anti-corruption policy in place for the rest of their life versus not having it. All the calculations are for 2005 – the first year after our sample period – at median age (33), education (high school completion), and wealth ($R\$47,387$). We report the willingness to pay as a share of initial wealth. We also calculate the willingness to pay as a fraction of yearly income, which is easier to interpret, using a life expectancy for Brazil of 75 years.

We report our findings in Figure 3. The wage policy is associated with a negative willingness to pay, implying that it is welfare reducing. All the other policies have a positive willingness to pay. With the exception of policies that include an increase in the audit probabilities, all other interventions have negligible direct costs. If a policy maker had to choose only one among the individual non-audit policies, they would adopt the 3-term policy. Residents are willing to give up as much as 0.60% of their annual income to implement it. If government can implement more than one of these non-audit policies at the same time, a policy that combines the 3-term limit with the CRA policy provides the highest welfare improvement. Individuals are willing to pay more than 1.3% of their annual income for their joint implementation.

The individual policy with the highest average willingness to pay is the audit-probability policy. Individuals are willing to pay as much as 1.23% of annual income (24.6% of their initial wealth) to enjoy the corresponding increase in public goods for the remainder of their lives. Thus, people would approve an increase in taxes as high as 1.23% to finance this policy. We argued earlier that the main limitation of this policy is its costs. To evaluate whether it is welfare improving, we need to answer the following question: Would a 1.23% increase in income taxes be sufficient to pay for the audit probability policy? The answer is yes. As we mentioned above, the estimated cost of an audit is approximately $R\$450,000$. The policy increases the probability of an audit from 5% to

16.8%. Given Brazil's $5,570$ municipalities, the policy increases the number of audits from 279 to 936 a year. The total cost of this policy is thus $657 \times 450,000 = R\$295,650,000$, which is less than 0.01% of Brazil's tax revenue in 2005 (according to IMF data). Thus, an increase in income taxes around 1%, which household are willing to pay, would generate well above the amount required to implement the policy.

Even though the audit-probability policy is welfare improving, we can still do better with a combined policy. A policy that combines the 3 term-limit policy with the CRA policy, and that increases the audit probability only for mayors in the last term will outperform the simple audit-probability policy. Individuals are willing to pay 1.36% of annual income for this multi-pronged approach. Individuals are also willing to pay 1.46% of their annual income for a policy that combines the audit probability policy with an audit-if-caught policy. But because of the additional audit costs associated with this latter policy, the former would still be preferred.

6   Conclusions

In this paper, we develop and estimate a dynamic model of decisions for local politicians who can engage in corruption. Using data from Brazil, including objective measures of local corruption, we estimate the model to quantify the importance of the incentives and constraints politicians face when making decisions over a finite horizon about what to consume, save, steal, and seek re-election. The model offers important insights into what determines corruption and how we can design policy to combat it.

We show that policies that strengthen the power of re-election incentives, such as extending term limits or banning corrupt politicians from running for office, can substantially reduce corruption among politicians who are eligible for reelection. But for politicians with shorter time horizons, such as those who have been term-limited, these policies are much less effective. In contrast, an audit policy can reduce corruption among both groups of politicians because it both promotes electoral accountability and brings about legal punishments. But audits are also costly and as a result, are not necessarily the best option. Combining the policies that enhance re-election incentives (i.e. the 3-term and the Clean Record Act) with an increase in the audit probability only for mayors in their last term could achieve a greater reduction in corruption at lower cost.

Our findings suggest that residents in Brazil are willing to pay more than 1.36% of their annual income for such a multi-pronged approach, which is more than what they would be willing to pay for Brazil's current audit policy and more than twice of what they would be willing to pay for any individual policy that does not require audits.

Our estimates and policy analysis clearly apply for the case of Brazil. But our framework is quite general and can be used to understand local corruption in any setting in which local politicians control large public budgets and are elected representatives. Our approach also highlights the importance of being able to compare across different policies and combinations of policies all within a common setting.

References

Aruoba, S. Boragan, Allan Drazen, and Razvan Vlaicu. 2015, May. "A Structural Model of Electoral Accountability." Nber working papers 21151, National Bureau of Economic Research, Inc.

Ashworth, Scott. 2012. "Electoral Accountability: Recent Theoretical and Empirical Work." *Annual Review of Political Science* 15 (1): 183–201.

Attanasio, Orazio, Hamish Low, and Virginia Sánchez-Marcos. 2008. "Explaining Changes in Female Labor Supply in a Life-Cycle Model." *The American Economic Review* 98 (4): 1517–1552.

Attanasio, Orazio P., Costas Meghir, and Ana Santiago. 2012. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *The Review of Economic Studies* 79 (1): 37–66.

Attanasio, Orazio P., and Guglielmo Weber. 1995. "Is Consumption Growth Consistent with Intertemporal Optimization? Evidence from the Consumer Expenditure Survey." *Journal of Political Economy* 103 (6): 1121–1157.

Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126 (5): 1912–1964.

Bai, Jie, Seema Jayachandran, Edmund J. Malesky, and Benjamin A. Olken. 2017. "Firm Growth and Corruption: Empirical Evidence from Vietnam." *The Economic Journal* n/a, no. n/a.

Basford, K. E., and G. J. McLachlan. 1985. "Likelihood Estimation with Normal Mixture Models." *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 34 (3): 282–289.

Becker, Gary S., and George J. Stigler. 1974. "Law Enforcement, Malfeasance, and Compensation of Enforcers." *The Journal of Legal Studies* 3 (1): 1–18.

Bobonis, Gustavo J., Luis R. Camara Fuertes, and Rainer Schwabe. 2016. "Monitoring Corruptible Politicians." *American Economic Review* 106 (8): 2371–2405 (August).

Brender, Adi, and Allan Drazen. 2005. "Political budget cycles in new versus established democracies." *Journal of Monetary Economics* 52 (7): 1271–1295 (October).

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini. 2013. "The Political Resource Curse." *American Economic Review* 103 (5): 1759–96.

Chen, Ting, and James Kai-sing Kung. 2018. "Busting the "Princelings": The Campaign Against Corruption in China's Primary Land Market*." *The Quarterly Journal of Economics* 134 (1): 185–226 (10).

Chong, Alberto, Ana L. De La O, Dean Karlan, and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice, and Party Identification." *The Journal of Politics* 77 (1): 55–71.

Colonnelli, E, and M Prem. 2019, September. "Corruption and firms." Documentos de trabajo 017430, Universidad del Rosario.

Corbi, Raphael, Elias Papaioannou, and Paolo Surico. 2018. "Regional Transfer Multipliers." *The Review of Economic Studies* 86 (5): 1901–1934 (11).

Costas-Perez, Elena, Albert Sole-Olle, and Pilar Sorribas-Navarro. 2012. "Corruption scandals, voter information, and accountability." *European Journal of Political Economy* 28 (4): 469–484.

Coviello, Decio, and Stefano Gagliarducci. 2017. "Tenure in Office and Public Procurement." *American Economic Journal: Economic Policy* 9 (3): 59–105 (August).

Diermeier, Daniel, Michael Keane, and Antonio Merlo. 2005. "A Political Economy Model of Congressional Careers." *American Economic Review* 95 (1): 347–373 (March).

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effect of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123 (2): 703–745.

———. 2011. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." *American Economic Review* 101 (4): 1274–1311 (June).

Finan, Frederico, and Maurizio Mazzocco. 2020. "Electoral Incentives and the Allocation of Public Funds." *NBER Working Paper, National Bureau of Economic Research*, no. w21859.

Fisman, Ray, and Miriam A. Golden. 2017. *Corruption: What Everyone Needs to Know.* OUP Catalogue no. 9780190463977. Oxford University Press.

Gamboa-Cavazos, Mario, and Frank Schneider. 2007. "Bankruptcy as a Legal Process." *SSRN Electronic Journal*, 06.

Gourieroux, Christian, and Alain Monfort. 1996. *Simulation-Based Econometric Methods.* New York: Oxford University Press.

Kaboski, Joseph P., and Robert M. Townsend. 2011. "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative." *Econometrica* 79 (5): 1357–1406.

Kang, Karam. 2015. "Policy Influence and Private Returns from Lobbying in the Energy Sector." *The Review of Economic Studies* 83 (1): 269–305 (07).

Kang, Karam, and Bernardo Silveira. 2020. "Understanding Disparities in Punishment: Regulator Preferences and Expertise." *Working Paper*.

Keane, Michael P., and Kenneth I. Wolpin. 1994. "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence." *The Review of Economics and Statistics* 76 (4): 648–672 (November).

Kiefer, J., and J. Wolfowitz. 1956. "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters." *The Annals of Mathematical Statistics* 27 (4): 887–906.

Lim, Claire S. H. 2013. "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges." *American Economic Review* 103 (4): 1360–97.

Lopez-Videla, Bruno. 2020. "Political Time Horizons and Government Investments: Evidence from Mexico." *Working paper*.

Mauro, Paolo. 1995. "Corruption and Growth." *Quarterly Journal of Economics*, no. 110:681–712.

Meghir, Costas, Ahmed Mushfiq Mobarak, Corina D. Mommaerts, and Melanie Morten. 2019, July. "Migration and Informal Insurance: Evidence from a Randomized Controlled Trial and a Structural Model." Nber working papers 26082, National Bureau of Economic Research, Inc.

Niehaus, Paul, and Sandip Sukhtankar. 2013. "Corruption Dynamics: The Golden Goose Effect." *American Economic Journal: Economic Policy* 5 (4): 230–69 (November).

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.

Olken, Benjamin A., and Rohini Pande. 2012. "Corruption in Developing Countries." *Annual Review of Economics* 4 (1): 479–509.

Redner, Richard. 1981. "Note on the Consistency of the Maximum Likelihood Estimate for Nonidentifiable Distributions." *The Annals of Statistics* 9 (1): 225–228.

Rose-Ackerman, Susan, and Bonnie J. Palifka. 2016. *Corruption and Government: Causes, Consequences, and Reform.* 2. Cambridge University Press.

Sieg, Holger, and Chamna Yoon. 2017. "Estimating Dynamic Games of Electoral Competition to Evaluate Term Limits in US Gubernatorial Elections." *American Economic Review* 107 (7): 1824–57 (July).

Stromberg, David. 2008. "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida." *American Economic Review* 98 (3): 769–807.

Tella, Rafael Di, and Ernesto Schargrodsky. 2003. "The Role of Wages and Auditing during a Crackdown on Corruption in the City of Buenos Aires." *The Journal of Law and Economics* 46 (1): 269–292.

Todd, Petra E., and Kenneth I. Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility." *American Economic Review* 96 (5): 1384–1417 (December).

Tables and Figures

Table 1: Summary Statistics

	Observations	Mean	Standard Deviation
Public Consumption (R$1,000)	5,514	3,043.1	3,410.67
Index of Private Inputs	5,461	0.95	0.343
Federal transfers (R$1,000)	5,328	2,038.3	3,397.7
Uncond. share of resources found to be corrupt	491	0.063	0.103
Fraction of corrupt mayors	491	0.733	0.443
Population in 2001	5,514	21,797.51	39,278.11
Small municipality	5,514	0.491	0.500
Medium municipality	5,514	0.414	0.492
Large municipality	5,514	0.094	0.293
Literacy rates in 2000	5,461	80.00	11.617
Second-term mayor	5,514	0.400	0.490
Re-election rates among those that ran, 2004	2,375	0.573	0.499
Re-election rates 2004, unconditional	3,307	0.411	0.492
Mayor has college education	5,498	0.379	0.485
Mayor's age	5,514	47.928	7.962
Relative Campaign Contributions	1,837	2.139	2.549
Self-reported wealth 2008	4,610	574,658.1	1438,309

Notes: This table presents summary statistics for the main variables used in the analysis. See the online appendix for a description of each variable.

Table 2: Reduced-form Evidence

	Fraction Stolen	Public Consumption per capita (logs)	Ran 2004	Reelected 2004	Log Wages of Ex-Mayors
	(1)	(2)	(3)	(4)	(5)
Second-term mayor	0.016*	0.144			
	(0.009)	(0.283)			
Amount Stolen (logs)		-0.028*			
		(0.016)			
Federal Transfers	-0.003	0.433***			
	(0.009)	(0.089)			
Audited			0.068	0.089	0.593
			(0.047)	(0.056)	(0.763)
Fraction Stolen × Audited			-0.123**	-0.150**	-0.022
			(0.056)	(0.067)	(0.829)
Per Capita Public Consumption (logs)			0.135***	0.204***	
			(0.036)	(0.048)	
College Education	-0.017*				4.260***
	(0.010)				(0.188)
Mayor's Age	-0.002		-0.033***	-0.039***	1.266***
	(0.002)		(0.004)	(0.005)	(0.187)
Mayor's Age^2					-0.107***
					(0.019)
Number of observations	474	349	3254	2333	3389
R^2	0.02	0.63	0.03	0.10	0.14

Notes: In column 1, the dependent variable is the fraction of resources audited that were classified as corruption. The regression also controls for log population and GDP per capita in 2001. In column 2, the dependent variable is public consumption per capita averaged over 2001-2004, expressed in logs. The regression also controls for the log of population, literacy rate, and GDP at the beginning of the term. We also allow the effects of literacy rate and GDP to vary according to whether the mayor is in his second term. In column 3, the dependent variable is an indicator for whether the incumbent ran for reelection in 2004. In column 4, the dependent variable is an indicator for whether the incumbent was reelected in 2004, conditional on running. The regression in column 3 and 4 also control for log of private per-capita GDP, mayor's age, log of population, and literacy rate. In column 5, the dependent variable is the log wages of ex-mayors. The regression also controls for population. See the online appendix for a description of each variable.

Table 3: Electoral Rule and Fine Distribution

Parameters	Estimates	Standard Errors	$\frac{dy}{dx}$
Audit	0.328	0.239	0.095
Audit× 1{$Stealing > 0$}	−0.554	0.274	-0.161
Log Public Consumption	0.553	0.169	0.161
Mayor's Age	−0.128	0.018	-0.037
Log Population	−0.469	0.100	-0.136
Large Municipality	0.611	0.162	0.177
Relative Campaign Contributions	0.325	0.043	0.094
Log Private GDP	−0.336	0.113	-0.097
Literacy Rate	−0.001	0.0005	-0.0000345
Low Type Constant	−1.763	1.733	-0.001
Increase in Constant for High Type	4.857	1.571	65.8
High Type Probability Cond. on Running	0.899	0.034	
Mean of Fine Distribution	0.094	0.020	
Variance of Fine Distribution	0.284	0.045	
Degree of Rivalry η	0.848	0.158	

Notes: This table reports the results of estimating Equation (2) by maximum likelihood. The dependent variable is an indicator for whether incumbent mayor was reelected conditional on running. Column 1 reports the point estimate. Column 2 reports standard errors. Column 3 reports the marginal effects evaluated at the variable mean. The estimation sample consists of 2,333 observations.

Table 4: Production Function

Parameters	Estimates	Standard Errors
Log Public Inputs	0.153	0.065
Log Private Inputs	0.429	0.094
College Education	0.043	0.034
Constant, Normalized	0	–
Mean Log Ability Distribution	−0.294	0.673
Mean Ability Log Distribution Second term	−0.283	0.690
Standard Deviation Ability Distribution	0.399	0.021

Notes: This table reports the GMM estimates of the production function. The dependent variable is the log of per capita public consumption. In addition to the variables report here, we also control for a fifth-order polynomial in population, literacy rate, public consumption in 2001, an indicator for whether the mayor is the second term, the interaction between literacy rate and being second-term mayor, and the interaction between public consumption in 2001 and being a second term mayor. The excluded instruments include the FPM indicators as discussed in Section 4. The estimation sample consists of 474 observations.

Table 5: Log Wages of Ex-Mayors

Parameters	Estimates	Standard Errors
Constant	-3.086	0.509
College Education	4.097	0.179
Age	1.035	0.183
Age^2	-0.098	0.018
Medium Municipality	1.144	0.211
Large Municipality	3.215	0.339
Standard Deviation	6.335	0.202

Notes: This table presents the wage regression used for ex-mayors. See the online appendix for a description of each variable. The sample consists of 3,389 observations.

Table 6: Preferences, High Type Probability, and Measurement Errors

Parameters	Estimates	Standard Errors
Relative Taste for Per-Capita Public Consumption	0.161	0.00011
Utility Cost of Running	0.027	0.00021
Utility from Being in Power	0.591	0.00049
High Type Probability	0.642	0.00121
Standard Deviation of the Measurement Errors	0.063	0.00015

Notes: This table reports the SMM estimates of the preference parameters, the probability of a high electoral appeal mayor, and standard deviation of the measurement errors. The standard errors are computed using the asymptotic distribution of the SMM estimator.

Table 7: Moments

	Model	Data
<u>Panel A</u>: Moments Used in the Estimation		
Average Fraction of Funds Stolen, First Term	5.59%	5.59%
Average Fraction of Funds Stolen, Second Term	7.31%	7.31%
Fraction Audited Incumbents Who Did Not Steal and Ran	75.60%	77.92%
Fraction Audited Incumbents Who Stole and Ran	69.39%	68.39%
Fraction of Incumbents Running with High Appeal	90.54%	89.90%
Fraction of Audited Mayors Caught Stealing	73.14	73.32
<u>Panel B</u>: Moments Not Used in the Estimation		
Fraction Incumbents Not Running for Reelection	29.37%	28.18%
Fraction Incumbents Winning Reelection	58.89%	57.26%
Fraction Audited Incumbents Who Stole Reelected	51.73%	51.11%
Mean of Log Ability, Second term	−0.245	−0.283
25th Percentile Fraction of Funds Stolen	0.0%	0.0%
Median Fraction of Funds Stolen	4.60%	2.10%
75th Percentile Fraction of Funds Stolen	9.94%	7.62%
90th Percentile Fraction of Funds Stolen	15.20%	19.56%

Notes: This table presents the moments used to estimate the model's parameter and moments used to evaluate the goodness of fit of the model. Column 1 reports simulated moments based on 500 simulations for each municipality. Column 2 reports the data moments.

Table 8: Simulations for Baseline Model (5% Audit Probability in All Terms)

	Full Sample (1)	First-term mayors (2)	Second-term mayors (3)
Fraction Stolen	0.103	0.088	0.129
Per-Capita Public Cons.	11.658	11.395	12.105
Age	49.482	50.421	47.894
Initial Wealth	445.092	331.4451	637.348
College Education	0.314	0.300	0.337
Ability	0.807	0.784	0.846
Electoral Appeal	0.775	0.642	1.000

Notes: This table presents simulated moments using the baseline model with a 5% audit probability in all terms, based on 500 simulations for each municipality.

Table 9: Decomposing Selection from Incentives and Decisions

	2nd mayors - 1st mayors (1)	Reelection Incentives (2)	Selection (3)	Optimal Decisions (4)
Fraction Stolen	0.040	42.8%	-32.3%	89.5%
Public Consumption	0.71	-5.0%	74.7%	30.3%

Notes: This table presents simulated moments based on 500 simulations for each municipality.

Table 10: Limitations of Policies and Possible Solutions

Policy	Issue 1	Issue 2	Solution 1	Solution 2
Audit-Probability	Expensive		Combine It with No-run or Audit-if-Caught Policy	
CRA	Limited Effects on Last Term Mayors		Increase Number of Terms	Increase Audit Probability in Last Term
Audited-If-Caught	Small Effects with Low Audit Probability		Combine with Audit Probability Policy	– –
3-Term	Limited Effects in First Term	Limited Effects in Last Term	Combine It with No-Run or Audit-if-Caught Policy	Increase Audit Probability in Last Term
Double Wages	Expensive	Ineffective If Not Running Is an Option	Not to Be Used If Not Running Is an Option	– –

Figure 1: The Effects of Anti-Corruption Policies on Corruption

Notes: This figure presents results based on 500 simulations for each municipality.

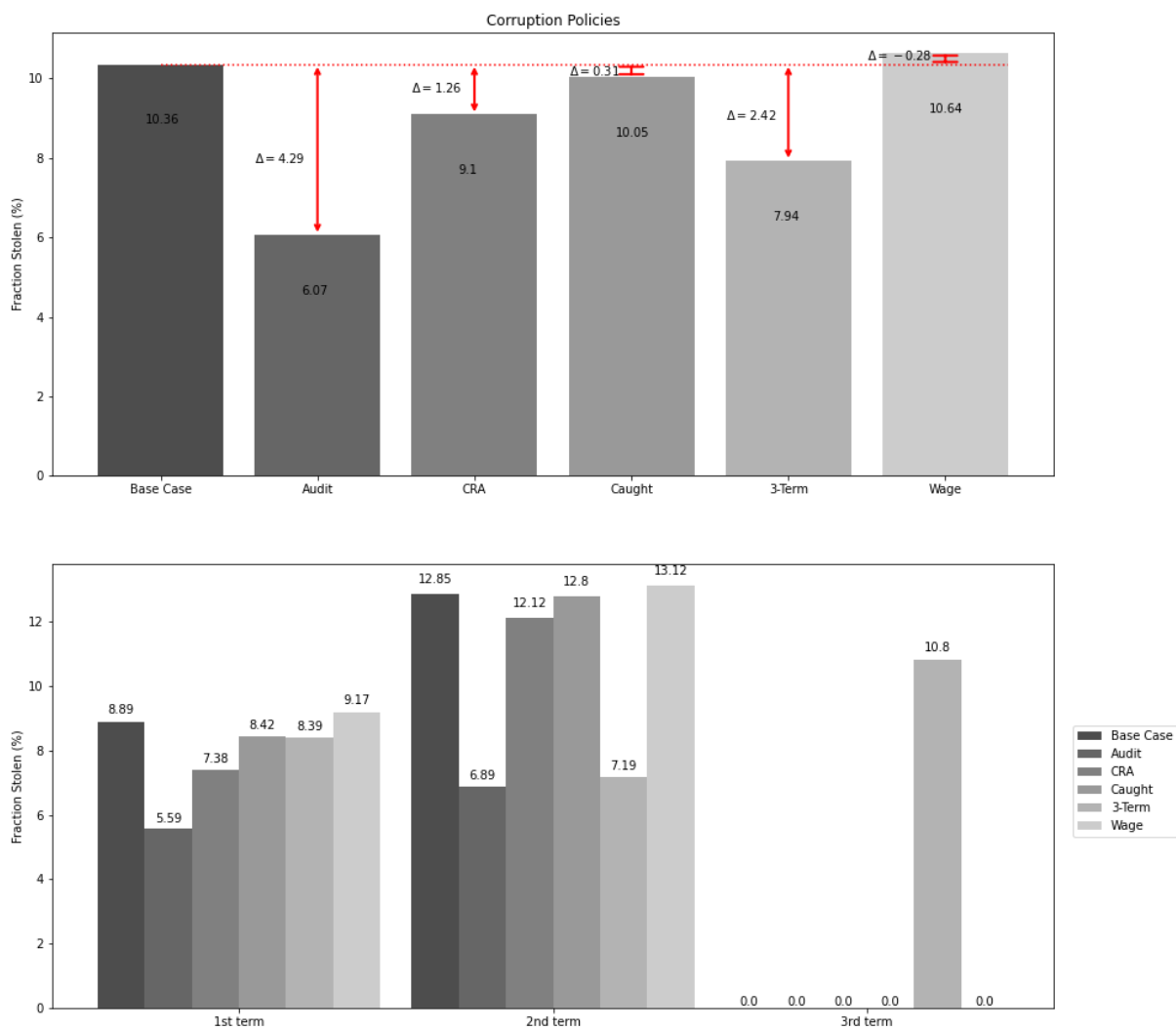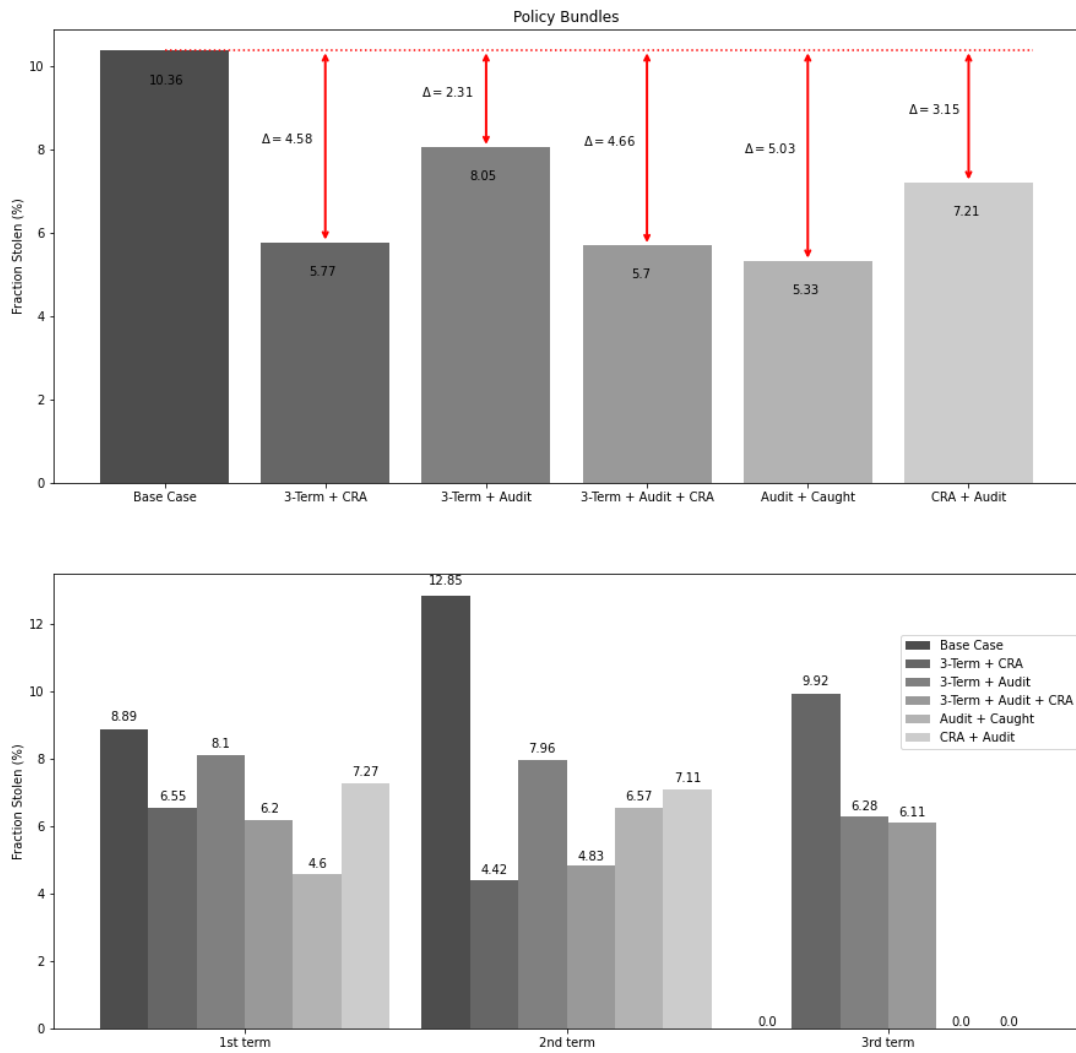Figure 2: The Effects of Policy Bundles on Corruption

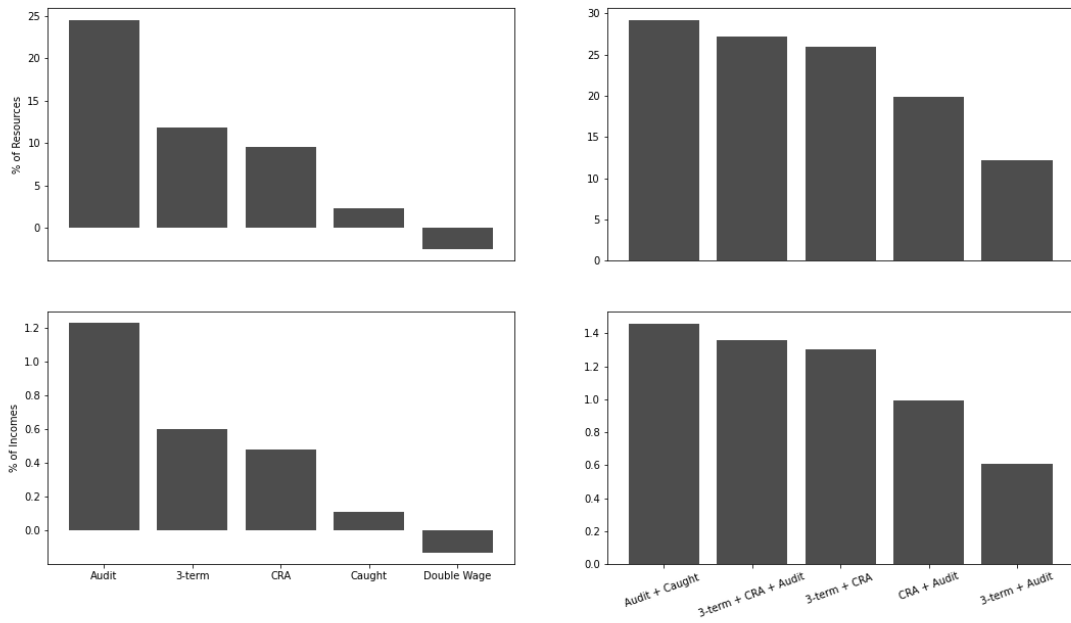Notes: This figure presents results based on 500 simulations for each municipality.

Figure 3: Willingness to Pay for Corruption Policies

Notes: This figure presents results based on 500 simulations for each municipality.

A Proof of Proposition 1

Let $P(\phi_H)$ and $P(\phi_H | d_R)$ be, respectively, the unconditional and conditional probability of high electoral appeal. Then,

$$P(\phi_H | d_R) = \frac{P(d_R | \phi_H) P(\phi_H)}{P(d_R)} =$$

$$\frac{P(d_R | \phi_H) P(\phi_H)}{P(d_R | \phi_H) P(\phi_H) + P(d_R | \phi_L)(1 - P(\phi_H))} = \frac{1}{1 + \dfrac{P(d_R | \phi_L)(1 - P(\phi_H))}{P(d_R | \phi_H) P(\phi_H)}}$$

Consider $\frac{P(d_R | \phi_L)(1 - P(\phi_H))}{P(d_R | \phi_H) P(\phi_H)}$. The terms $P(d_R | \phi_L)$ and $P(d_R | \phi_L)$ are not affected by changes in $P(\phi_H)$, since in the model the mayor's decision to run for reelection does not depend on the fraction or number of opponents with high electoral appeal. Hence,

$$\frac{\partial}{\partial P(\phi_H)}\left(\frac{P(d_R | \phi_L)(1 - P(\phi_H))}{P(d_R | \phi_H) P(\phi_H)}\right) = -\frac{P(d_R | \phi_L)}{P(d_R | \phi_H)}\frac{1}{P(\phi_H)^2} < 0$$

and

$$\frac{\partial P(\phi_H | d_R)}{\partial P(\phi_H)} = \frac{\partial}{\partial P(\phi_H)}\left(\frac{1}{1 + \dfrac{P(d_R | \phi_L)(1 - P(\phi_H))}{P(d_R | \phi_H) P(\phi_H)}}\right) > 0.$$

Thus, $P(\phi_H)$ is identified if $P(\phi_H | d_R)$ is known.

B Proof of Proposition 2

So far we have shown identification of all parameters except θ, the variance of the measurement error, the cost of running parameter κ, and the utility from being in power parameter ρ. We will now show identification of θ by considering the decision problem of second-term mayors. Since the parameters κ and ρ affect only the decisions of mayors who can run for reelection and, in the data, second-term mayors cannot be reelected, we can ignore them when studying their problem. We can also ignore the variance of the measurement errors, as they have no effect on a mayor's decisions. Given this, we will show identification of θ by proving that the expected value of stealing for second-term mayors is strictly decreasing in that parameter.

The proof is provided in two steps. We first show that the problem of a second-term mayor has a solution and the solution is unique. We then show that, at the unique solution, the amount stolen by a second term-mayor is monotonically decreasing in θ, hence the result.

For second-term mayors, re-election concerns are irrelevant and stealing can only take place in their current term. Without loss of generality, we can therefore restrict our attention to a two-period model in which the first period corresponds to the mayor's second term and the second period to the rest of her or his life. To simplify the notation, let δ be the dummy for being audited at the end of the term ($\delta^i_{au,t}$ in the paper), and note that, under our assumptions, the fine schedule takes the following form: $g(s) = e^{\mu + \sigma z} s$, with $z \sim N(0,1)$. Also note that we can rewrite the production function for public consumption as follows:

$$\frac{Q}{d} = \left(\frac{z^{pu}}{d} \right)^{\alpha_1} (z^{pr})^{\alpha_2} \exp \left\{ \alpha_3 + \alpha_4 e_i + \sum_{j=1}^{N_d} \alpha_{4+j} (d)^j \right\} \quad a_i = \bar{\lambda}(s),$$

where $\bar{\lambda}(s)$ highlights that the only decision variable for the mayor is the fraction of resources diverted. Let $\lambda(s) = \bar{\lambda}(s) d$. Then, the second-term mayor chooses private consumption, savings, and stealing as the solution of the following problem:

$$\max_{c,b,s} u(c) + v(Q) + \beta E[u(\tilde{c})] + \beta E\left[v\left(\tilde{Q}\right)\right] - \theta \eta \log(d)$$

$$s.t. \quad c = y - b + s, \qquad \tilde{c} = \tilde{y} + Rb - \tilde{R}s,$$

$$Q = \lambda(s), \qquad \tilde{R} = \delta e^{\mu + \sigma z}, \qquad z \sim N(0,1), \qquad \delta \in [0,1],$$

where $v(Q) = \theta \log(Q)$, y includes labor earning and a possible fine for the amount stolen in the previous term, $\tilde{Q}$ is the public consumption provided by future mayors, and $\tilde{y}$ denotes future labor earnings, which are assumed to be uncorrelated with current stealing following the empirical evidence provided in Section 2. By replacing the constraints in the objective function, the problem can be rewritten as follows:

$$\max_{c,s} u(c) + l(s) + \beta E\left[u\left(\tilde{y} + Ry - Rc + \left(R - \tilde{R}\right)s\right)\right] + \beta E\left[v\left(\tilde{Q}\right)\right] - \theta \eta \log(d).$$

where $l(s) = \theta \log(\lambda(s))$ and $E\left[v\left(\tilde{Q}\right)\right]$ does not depend on the decision variables c and s. Given

53

our functional-form assumptions, λ is strictly concave in s, which implies that l is strictly concave in s.

A solution to the problem exists and is unique if the objective function is strictly concave. Consider the functions

$$h_1\left(c, s\right) = u\left(c\right) + l\left(s\right), \qquad h_2\left(c, s\right) = E\left[u\left(\tilde{y} + Ry - Rc + \left(R - \tilde{R}\right)s\right)\right].$$

The objective function is strictly concave if h_1 and h_2 are concave and at least one is strictly concave in c and s, as the weighted sum of concave functions with positive weights is concave. The Hessian matrix of h_1 takes the following form:

$$H_1 = \begin{bmatrix} u''\left(c\right) & 0 \\ 0 & l''\left(s\right) \end{bmatrix}.$$

Given the strict concavity of $u\left(c\right)$ and $l\left(s\right)$, we have $u''\left(c\right) < 0$, $l''\left(s\right) < 0$, and $u''\left(c\right) l''\left(s\right) > 0$. The function h_1 is therefore strictly concave. The Hessian matrix of h_2 takes the following form:

$$H_2 = \begin{bmatrix} R^2 E\left[u''\left(\tilde{c}\right)\right] & -RE\left[\left(R - \tilde{R}\right) u''\left(\tilde{c}\right)\right] \\ -RE\left[\left(R - \tilde{R}\right) u''\left(\tilde{c}\right)\right] & E\left[\left(R - \tilde{R}\right)^2 u''\left(\tilde{c}\right)\right] \end{bmatrix}.$$

Given the strict concavity of $u\left(c\right)$ we have $R^2 E\left[u''\left(\tilde{c}\right)\right] < 0$ and $E\left[\left(R - \tilde{R}\right)^2 u''\left(\tilde{c}\right)\right] < 0$. It is left to prove that the product of the diagonal terms minus the product of the off-diagonal terms is positive, i.e.

$$R^2 E\left[u''\left(\tilde{c}\right)\right] E\left[\left(R - \tilde{R}\right)^2 u''\left(\tilde{c}\right)\right] - R^2 E\left[\left(R - \tilde{R}\right) u''\left(\tilde{c}\right)\right]^2 > 0.$$

We have

$$R^2 E\left[u''\left(\tilde{c}\right)\right] E\left[\left(R-\tilde{R}\right)^2 u''\left(\tilde{c}\right)\right] - R^2 E\left[\left(R-\tilde{R}\right)u''\left(\tilde{c}\right)\right]^2$$

$$= R^2 E\left[u''\left(\tilde{c}\right)\right]\left\{E\left[R^2 u''\left(\tilde{c}\right)\right] + E\left[\tilde{R}^2 u''\left(\tilde{c}\right)\right] - 2RE\left[\tilde{R}u''\left(\tilde{c}\right)\right]\right\}$$

$$- \left\{R^4 E\left[u''\left(\tilde{c}\right)\right]^2 + R^2 E\left[\tilde{R}u''\left(\tilde{c}\right)\right]^2 - 2R^3 E\left[u''\left(\tilde{c}\right)\right] E\left[\tilde{R}u''\left(\tilde{c}\right)\right]\right\}$$

$$= R^2 E\left[u''\left(\tilde{c}\right)\right] E\left[\tilde{R}^2 u''\left(\tilde{c}\right)\right] - R^2 E\left[\tilde{R}u''\left(\tilde{c}\right)\right]^2 = R^2\left\{E\left[u''\left(\tilde{c}\right)\right] E\left[\tilde{R}^2 u''\left(\tilde{c}\right)\right] - E\left[\tilde{R}u''\left(\tilde{c}\right)\right]^2\right\}$$

$$> R^2\left\{E\left[\sqrt{\left(-u''\left(\tilde{c}\right)\right)}\sqrt{\tilde{R}^2\left(-u''\left(\tilde{c}\right)\right)}\right]^2 - E\left[\tilde{R}u''\left(\tilde{c}\right)\right]^2\right\} = R^2\left\{E\left[\tilde{R}u''\left(\tilde{c}\right)\right]^2 - E\left[\tilde{R}u''\left(\tilde{c}\right)\right]^2\right\} = 0.$$

where the inequality follows from the Cauchy-Schwarz inequality: $E\left[XY\right]^2 < E\left[X^2\right] E\left[Y^2\right]$, applied to $X = \sqrt{\left(-u''\left(\tilde{c}\right)\right)}$ and $Y = \sqrt{\tilde{R}^2\left(-u''\left(\tilde{c}\right)\right)}$. The function h_2 is therefore concave. Hence, the object function is concave and the problem has a unique solution.

We will now show that, at the unique solution, stealing is decreasing in θ. The unique solution for c and s must satisfy the following two Euler equations:

$$u'\left(c^*\right) - \beta R E\left[u'\left(\tilde{c}^*\right)\right] = 0$$
$$v'\left(Q^*\right)\lambda'\left(s^*\right) + \beta E\left[u'\left(\tilde{c}^*\right)\left(R-\tilde{R}\right)\right] <= 0, \tag{4}$$

where the inequality accounts for the possibility that $s^* = 0$ (c^* is always greater than 0, as $\lim_{c\to 0} u'\left(c\right) = \infty$). Consider an increase in θ. This change only affects $v'\left(Q\right)$. Specifically, $\frac{\partial v'(Q)}{\partial \theta} = \frac{\partial(\theta/Q)}{\partial \theta} = 1/Q > 0$. Moreover, $\lambda'\left(s^*\right) < 0$. Hence, at the new θ, but old solution for s and c, the left hand side of (4) decreases and the inequality changes to

$$\bar{v}'\left(Q^*\right)\lambda'\left(s^*\right) + \beta E\left[u'\left(\tilde{c}^*\right)\left(R-\tilde{R}\right)\right] < 0.$$

Whether these mayors will choose to increase or decrease s^* in response to the increase in θ depends on the sign of the derivative of the left hand side of (4) with respect to s, which take the

following form:

$$\frac{\partial}{\partial s} \left(v'\left(Q^*\right)\lambda'\left(s^*\right) + \beta E\left[u'\left(\tilde{c}^*\right)\left(R - \tilde{R}\right)\right]\right)$$

$$= v''\left(Q^*\right)\lambda'\left(s^*\right)^2 + v'\left(Q^*\right)\lambda''\left(s^*\right) + \beta E\left[u''\left(\tilde{c}^*\right)\left(R - \tilde{R}\right)^2\right] < 0,$$

where the inequality follows from $v\left(Q\right)$ being increasing and concave, and the concavity of $\lambda\left(s\right)$ and $u\left(c\right)$. The mayors can therefore be divided into three groups based on their response. If their optimal stealing at the old θ was zero, they continue to steal zero funds. If s^* was positive at the old θ and the left hand side of (4) is still negative when choosing stealing equal to zero, the mayors will optimally choose to divert zero funds. They will therefore reduce the amount stolen. Lastly, if s^* was positive at the old θ and there is a new stealing amount $0 < s^{**} < s^*$ at which (4) is satisfied as an equality, the mayors will reduce their optimal amount of stealing to s^{**}.

We can therefore conclude that $\frac{\partial s^*}{\partial \theta} \leq 0$, with strict inequality for the second-term mayors that are not at a corner before the change in θ. Hence, if we take the expectation over mayors of $\frac{\partial s^*}{\partial \theta}$, we have $E\left[\frac{\partial s^*}{\partial \theta}\right] < 0$, provided that some second-term mayors steal for any relevant value of θ (in the data 73% of mayors divert resources). This concludes the proof.

C   Proof of Proposition 3

The only parameters for which we still need to prove identification are κ, ρ, and the variance of the measurement error in stealing σ_m. Since σ_m does not affect the incumbent's decision to run for reelection, in the proof we will abstract from it.

Remember that S is the vector of state variables that affect the decisions of an incumbent, that $E\left[V_{RM} | S\right]$ and $E\left[V_{NRM} | S\right]$ denote the expected value of running and not running conditional on S, and that $P\left(S\right)$ is the probability that the incumbent wins the election. Then, given S, an incumbent choose to run for reelection if

$$E\left[V_{RM} | S\right] \geq E\left[V_{NRM} | S\right] + \epsilon_R,$$

where $\epsilon_R \sim N\left(0, \sigma_R\right)$ is a shock to the decision to run and σ_R is normalized to be equal to 1. As V_{WM} is the value function of an incumbent who runs and wins the election and V_{LM} is the value

function of an incumbent who runs and loses the election, the previous inequality can be rewritten in the following form:

$$P(S) E[V_{WM}|S] + (1 - P(S)) E[V_{LM}|S] \geq E[V_{NRM}|S] + \epsilon_R,$$

Since the cost of running κ and the utility from being in power ρ enter additively the incument's utility function only for the current term, we can rewrite the expected values of running and winning and running and loosing as follows:

$$E[V_{WM}|S] = E[\bar{V}_{WM}|S] - \kappa + \rho$$

and

$$E[V_{LM}|S] = E[\bar{V}_{LM}|S] - \kappa.$$

where $\bar{V}_{WM}$ and $\bar{V}_{LM}$ denote the value functions without κ and ρ The previous inequality takes therefore the following form:

$$P(S)\left(E[\bar{V}_{WM}|S] - \kappa + \rho\right) + (1 - P(S))\left(E[\bar{V}_{LM}|S] - \kappa\right) \geq E[V_{NRM}|S] + \epsilon_R,$$

In the model, $E[\bar{V}_{LM}|S] = E[V_{NRM}|S]$ as the only difference between an incumbent who runs and loses and an incument who chooses not to run is the cost κ. Consequently, the inequality simplifies to

$$P(S)\left(E[\bar{V}_{WM}|S] - E[\bar{V}_{LM}|S]\right) - \kappa + P(S)\rho \geq \epsilon_R.$$

Hence, for any given S, the probability that an incumbent runs is given by

$$P[R|S] = P\left[P(S)\left(E[\bar{V}_{WM}|S] - E[\bar{V}_{LM}|S]\right) - \kappa + P(S)\rho \geq \epsilon_R\right] =$$
$$= \Phi\left(P(S)\left(E[\bar{V}_{WM}|S] - E[\bar{V}_{LM}|S]\right) - \kappa + P(S)\rho\right),$$

where Φ is the cumulative density function of a standard normal. As a result,

$$\Phi^{-1}(P[R|S]) = P(S)\left(E[\bar{V}_{WM}|S] - E[\bar{V}_{LM}|s]\right) - \kappa + P(S)\rho.$$

By taking the expectation over S, we then have

$$E_S\left[\Phi^{-1}\left(P\left[R\,|S\right]\right) - P\left(S\right)\left(E\left[\bar{V}_{WM}\,|S\right] - E\left[\bar{V}_{LM}\,|S\right]\right)\right] = -\kappa + E_S\left[P\left(S\right)|S\right]\rho. \qquad (5)$$

Consider now two groups of mayors: the group composed of audited mayors that did not steal, S^N, and the group formed by audited mayors who chose to steal, S^S. By taking differences of the previous equation for the two groups, we obtain

$$E_S\left[\Phi^{-1}\left(P\left[R\,|S^N\right]\right) - P\left(S\right)\left(E\left[\bar{V}_{WM}\,|S^N\right] - E\left[\bar{V}_{LM}\,|S^N\right]\right)\right] -$$

$$E_S\left[\Phi^{-1}\left(P\left[R\,|S^S\right]\right) - P\left(S\right)\left(E\left[\bar{V}_{WM}\,|S^S\right] - E\left[\bar{V}_{LM}\,|S^S\right]\right)\right] = \left(E_S\left[P\left(S\right)|S^N\right] - E_S\left[P\left(S\right)|S^S\right]\right)\rho,$$

which implies that

$$\rho = \frac{E_S\left[\Phi^{-1}\left(P\left[R\,|S^N\right]\right) - P\left(S\right)\left(E\left[\bar{V}_{WM}\,|Z^N\right] - E\left[\bar{V}_{LM}\,|S^N\right]\right)\right]}{E_S\left[P\left(S\right)|S^N\right] - E_S\left[P\left(S\right)|S^S\right]}$$
$$- \frac{E_S\left[\Phi^{-1}\left(P\left[R\,|S^S\right]\right) - P\left(S\right)\left(E\left[\bar{V}_{WM}\,|S^S\right] - E\left[\bar{V}_{LM}\,|S^S\right]\right)\right]}{E_S\left[P\left(S\right)|S^N\right] - E_S\left[P\left(S\right)|S^S\right]}.$$

All the parameters entering the value functions on the right hand side have been shown to be identified. Hence, if the probability of running and winning for audited mayors that did not steal and for audited mayors that stole are observed, there is only one value of ρ that corresponds to the observed data. Lastly, by substituting for ρ in (5) for audited non-stealing mayors using the previous equation, we can find the unique value of κ that corresponds to the data, which concludes the proof.

D Proof of Proposition 4

Given the vector of state variables S, let $s\left(S\right)$ be the optimal choice of stealing as a function of S. Then, the probability of observing zero stealing is given by $P\left(s\left(S\right) + \nu \leq 0\right)$, where ν is the measurement error and the probability is computed over the distributions of S and ν.

Consider a particular realization of S and, hence, a particular realization of $s\left(S\right)$ and denote with $z \sim N\left(0,1\right)$ a draw from a standard normal. Then, since $\nu \sim N\left(0, \sigma_\nu\right)$, the probability of

observing zero stealing conditional on S can be written as follows:

$$P\left(s\left(S\right)+\nu\le 0\middle|\,S\right) = P\left(\nu\le -s\left(S\right)\middle|\,S\right) = P\left(\sigma_\nu z\le -s\left(S\right)\middle|\,S\right)$$
$$= P\left(z\le -\frac{s\left(S\right)}{\sigma_\nu}\middle|\,S\right) = \Phi\left(-\frac{s\left(S\right)}{\sigma_\nu}\middle|\,S\right),$$

where Φ is the standard normal cumulative distribution function. By taking the derivative of $\Phi\left(-\frac{s(S)}{\sigma_\nu}\middle|\,S\right)$ with respect to σ_ν, one obtains

$$\frac{\partial\Phi\left(-\frac{s(S)}{\sigma_\nu}\middle|\,S\right)}{\partial\sigma_\nu} = \phi\left(-\frac{s\left(S\right)}{\sigma_\nu}\middle|\,S\right)\frac{s\left(S\right)}{\sigma_\nu^2} > 0.$$

Hence, conditional on S, the probability of zero stealing is monotonically increasing in σ_ν. The unconditional probability can be derived by integrating the conditional probability over S:

$$P\left(s\left(S\right)+\nu\le 0\right) = \int_S \Phi\left(-\frac{s\left(S\right)}{\sigma_\nu}\middle|\,S\right)dF\left(S\right).$$

Since S and ν are independent, by taking the derivative of $P\left(s\left(S\right)+\nu\le 0\right)$ with respect to σ_ν, we have

$$\frac{\partial P\left(s\left(S\right)+\nu\le 0\right)}{\partial\sigma_\nu} = \int_S \phi\left(-\frac{s\left(S\right)}{\sigma_\nu}\middle|\,S\right)\frac{s\left(S\right)}{\sigma_\nu^2}dF\left(S\right) > 0.$$

The probability of zero stealing is therefore monotonically increasing in σ_ν. Since all the model parameters that determine $s\left(S\right)$ have been shown to be identified, this implies that there is a unique value of σ_ν that corresponds to a particular value of $P\left(s^o\left(S\right)\le 0\right) = P\left(s\left(S\right)+\nu\le 0\right)$.

E   Data Appendix

In this section, we describe all the variables used in the analysis, their source of origin, and how they were constructed.

Corruption Data

These data come from Ferraz and Finan (2011). They are constructed from the official audit reports of the municipalities that were drawn from the first 11 lotteries. See Ferraz and Finan

(2011) for a detailed discussion for how the corruption measures were defined and coded. The corruption measures correspond to the period of 2001-2003. From these data, we created the following main variable:

Fraction Stolen The share of resources audited classified as corruption.

Audit An indicator for whether the municipality was audited during the first 11 lotteries.

Election Data

These data were downloaded from Brazil's electoral commission (https://www.tse.jus.br/) and cover the mayor elections for 2000, 2004, and 2008. The data contain detailed information on every candidate that ran for office, including their electoral outcomes and various socio-demographic characteristics. For our estimation sample, we only consider mayors who were in office during the 2001-2004 term. From these data, we create the following main variables:

Ran for reelection An indicator for whether the mayor ran for office in the 2004 elections

Reelection An indicator for whether the mayor was reelected in the 2004 elections.

Second-term An indicator for whether the mayor was in his second term during 2001-2004.

Age The age of the mayor as of the year 2000. When estimating the model, we discretize this variable into 4 year intervals. The variable ranges from 1 to 10.

College An indicator for whether the mayor has a college education

Wealth For each candidate, we use their wealth data measured in 2008. We had missing wealth information for 17% of the sample. For these candidates, we assigned them the sample average.

Relative Campaign Contribution Total 2004 campaign contributions of the incumbent divided by the campaign contributions of the second place candidate.

Municipality Data

These data come from Instituto de Pesquisa Econômica Aplicada (IPEA), a government-led research organization. IPEA has created a data repository (www.ipeadata.gov.br) containing information on various socio-economic characteristics of Brazil's municipalities. IPEA collects and aggregates these data from several government agencies, including the Instituto Brasileiro de Geografia e Estatistica (IBGE) and the National Treasury (Tesouro Nacional). For these data, we create the following main variables:

Public Consumption The average of total GDP (in R$1000) for the municipality for the years 2001-2004.

Private Inputs We constructed this variable using factor analysis. It is the first principal component of three variables: the number of firms in the municipality in 1995, average wages in the private sector in 2000, and rate of employment in 2000.

Federal Transfers Total amount of federal funds transferred to the municipality.

Public Inputs Federal transfers multiplied by one minus fraction stolen.

Population Population of the municipality measured in 2001.

Large Municipality Indicator for whether the municipality has a population larger than $50,000$.

Medium Municipality Indicator for whether the municipality has a population between $10,001$ and $50,000$.

Small Municipality Indicator for whether the municipality has a population less than or equal to $10,000$.

Literacy Rate Literacy rate of the adult population in 2000, measured in percentages.

Fines Data

These data were originally assembled by Avis, Ferraz, and Finan (2018), who downloaded them in 2013 from the National Council for Justice (CNJ). These data include the names of all individuals

charged with misconduct in public office. For each individual, the data set contains the type of irregularity (e.g. violation of administrative principles or diversion of resources), the court where the conviction took place, the fine, and the date. These data are matched to the electoral data based on where the individual was a mayor and the period he/she served in office. Individuals on this list are banned from running for any public office for at least five years. Using these data, we create the following variable:

Fine as a multiple of stealing We divide the fine amount by the amount stolen.

Mayor's Salary

To collect these data, we randomly sampled 10% of municipalities stratified by three population thresholds. We then downloaded the mayor's wage from the mayors' office website. The average monthly earnings paid to mayors in municipalities with population less than 10,000 residents were equal to R\$3,233. They were equal to R\$4,268 for municipalities with population between 10,000 and 50,000 residents, and to R\$5,077 for larger municipalities. These salaries have all been deflated to real terms based on the year 2000.

Private Sector Wages

These data come Relação Anual de Informações Sociais (RAIS), which is an employer-employee data set collected on an annual basis and captures the entirety of Brazil's formal sector employment. Our data covers the period of 2002-2013. These data are matched to the electoral data based on each candidate's national identification number (CPF). We were able to match 68% of all candidates that ran for mayor. From these data, we measure a mayor's wage once they leave office conditional on not being elected for future office.

Wages of ex-mayors Monthly wage of ex-mayors averaged over the period of 2005 to 2013.