Combating Political Corruption with Policy Bundles*

Frederico Finan[†] Maurizio Mazzocco[‡]
UC-Berkeley UCLA

April 2024

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 audit policies effectively reduce corruption, a multi-pronged approach that bundles an audit policy with other policies can achieve much higher welfare gains.

Keywords: Anti-corruption Policies, Corruption, Reelection Incentives, Political Selection, Dynamic Political Economy Model, Structural Estimation.

*We are grateful to participants at various seminars and conferences for helpful comments. Research reported in this paper was supported by the National Science Foundation, Award Number 922444.

†Department of Economics, email: ffinan@econ.berkeley.edu;

‡Department of Economics, email: 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 countries 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 evaluated during different periods and/or settings. For example, we have seen that audit policies can effectively reduce 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 from 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 also their future ones. It is difficult to capture these future margins of adjustment in the reduced form.

To address these gaps in the literature, we need to understand better why politicians engage in corruption over the course of their life cycle. Specifically, we need to understand the main incentives and constraints politicians face and 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 can then compare the effects of various anti-corruption policies within the same setting, 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 found to be corrupt.[1]

Our model is quite general, as it applies 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 number 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, is crucial for identifying and estimating the model's parameters. Third, Brazil allowed in 1997 mayors to hold office for two consecutive terms. This variation is essential for identifying the effect of reelection incentives on corruption. Finally, besides the data on corruption, we also have detailed information about all candidates who 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. For example, we can match the difference in stealing between mayors in their first term and those in their second and final term. This comparison is important because it captures a combination of two electoral forces. First, it reflects a dynamic decision by mayors to forgo stealing in their first term in order to get reelected to a second term. We find that this electoral incentive accounts for most of the difference in stealing between first and second-term mayors. Second, the comparison also captures selection

[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.

effects created by elections. We find that second-term mayors are, on average, more positively selected than first-term mayors in their observed and unobserved characteristics. These selection effects help explain why second-term mayors can, on average, steal more than first-term mayors and still provide higher public consumption.

Given these insights, we use our estimated model to quantify the effects on stealing and welfare of four commonly proposed anti-corruption policies. These include increasing the probability of a federal audit, extending the number of terms mayors can serve, 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 individual 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 34 percent. For comparison, adding an additional term to the number of consecutive terms a mayor can hold office is about a third as effective, with a 12% reduction. The least effective policy to reduce corruption is to double mayors' wages. This policy reduces corruption by less than 10%.

While combating corruption is a meaningful objective, what we ultimately care about is how the policy affects voters' welfare. Based on this metric, the audit policy is not the most effective, largely due to its costs. Instead, Brazil's Clean Record Act is the most effective, as voters are willing to pay 1% of their annual income for this policy. In comparison, voters are willing to pay less than 0.4% for the audit policy. Our model also allows us to compute the audit probability that maximizes voter's welfare. We find an optimal audit probability of 16%, which interestingly is not too far from Brazil's current policy. Thus, even at the optimal audit probability, voters will still prefer the Clean Records Act policy.

The fact that we can compare different policies 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. However, 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 who 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 affects earlier terms through the no-run restriction. By combining the two policies, the average willingness to pay increases to 1.2% of annual income, and corruption reduces by more than 60%. 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 adding the audit policy set at the optimal audit probability for the bundle. When we do so, the optimal audit probability for the combined policy reduces to 8%, and this policy proves to be our most effective. Citizens are willing to pay 1.3% of their lifetime income for this policy, and it reduces corruption by more than 80%.

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's anti-corruption program. They find that disclosing information about corruption in a municipality reduces corruption levels, but only in the short run. In subsequent terms, municipal corruption levels increased, especially among those who refrained from corruption before the first audit. Avis, Ferraz, and Finan (2018) analyze Brazil's anti-corruption program and 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 suggested 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

[2]For excellent surveys on corruption, see Olken and Pande (2012); Rose-Ackerman and Palifka (2016).
[3]See Ashworth (2012) for an excellent review of the literature.

Mexico that allowed mayors, who had been limited to a single term, to run for reelection 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 among bureaucrats. For example, Di 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 strengths and limitations of the policies. In contrast, the model we estimate captures many of the mechanisms that affect the decision of politicians to engage in corruption. This enables us to assess empirically their relative importance and the advantages and disadvantages the policies can present. The understanding of these mechanisms is critical for the design of policy as a redress for corruption. Second, our approach enables us to 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 also relates 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. This section describes the data and presents 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) Is corruption lower among mayors who have reelection incentives? 3) Is corruption associated with lower levels of public consumption? 4) Does being found to be corrupt affect the decision to run for reelection? 5) Are corrupt politicians less likely to be reelected? 6) Is being found to be corrupt associated with lower future wages? The correlations we present below are not necessarily specific to Brazil; other studies have documented similar findings in different settings.

Public Funds and Corruption. In 2003, Brazil's Comptroller General of the Union (Controladoria Geral da União – CGU), a functionally autonomous branch of the Federal Government, started a national program called *Programa de Fiscalização por Sorteios Públicos* to audit municipalities for their use of federal funds. The program selects approximately 60 municipalities by public lottery in a given round. All municipalities with a population of up to 500,000 inhabitants are eligible for selection.[4] As of February 2015, the program had conducted 2,241 audits across 40 lotteries, involving over R$22 billion dollars worth of federal funds.[5]

The CGU audits municipalities by issuing a random selection of inspection orders based on all the federal transfers the municipality received during the previous three to four years. Each order stipulates an audit task for a specific government project (e.g., school construction, purchase of medicine, etc.) within a specific sector. Given these audit tasks, the CGU will send 10 to 15 auditors for one to two weeks to examine municipal accounts, inspect public works, and verify the delivery of public services. After the inspections are completed, a detailed report describing any

[4] This eligibility criterion has changed slightly over time.

[5] See (Avis, Ferraz, and Finan 2018) for detailed description of the program.

irregularities is submitted to the central office. These reports are posted on the internet and sent to the municipal legislative branch, the Federal Courts of Accounts (TCU), the Federal Police, and the Federal Prosecutors' Office (MPF) for potential legal prosecution. In fact, (Avis, Ferraz, and Finan 2018) estimate that the audits increased the likelihood of incurring a formal legal action by 20%.

Throughout the audit process, the CGU takes several steps to minimize the threat of auditor corruption and maximize the audit's efficacy. For starters, the entire process is highly transparent, which minimizes the incentives for wrongdoing. The CGU also hires auditors based on a competitive public examination and pays them highly competitive salaries. Moreover, inspections are done by teams to increase the likelihood of detection and to reduce further the opportunity for corruption among individual auditors.[6]

Our data on corruption comes 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.[7] 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, municipalities received on average R$2,038,274 (approximately US$886,206) of federal transfers per year from the federal government to provide local public goods, including public education, health, and sanitation.[8] Mayors stole 6.3 percent of these funds or R$122,907.92. To put this amount into context, Brazil's GDP per capita in 2000 was around R$5,913.

There is also a considerable amount of heterogeneity both in the amount municipalities receive and in what mayors steal. 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.

[6] Consistent with these organization features, Ferraz and Finan (2008) find no evidence that auditors manipulate the audit reports.

[7] 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.

[8] In 2001, the exchange rate was 2.3 Reais to the US dollar.

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.[9] 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.

Public Consumption. Mayors affect the welfare of their citizens mostly through the provision of public goods, such as education and health. They do this directly by funding their production and indirectly by setting policies that can affect the economy more generally. To capture the various ways 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.[10]

Even though per-capita GDP may be an imperfect measure of public consumption, it offers two advantages over alternative indicators. First, politicians provide many public goods (e.g., schools, parks, roads, or water sanitation), and it is difficult to find data to measure all of them. But all of them are likely to be captured in per-capita GDP. In the Appendix Table 8, we show that GDP per capita strongly correlates with various public goods indicators, including the share

[9]Although our specification is different, this result is similar to those found in Ferraz and Finan (2011).

[10]These data are constructed by Brazil's National Statistical Institute (IBGE) and are available yearly since 2000. We downloaded them at the site: www.ipeadata.gov.br. See the online appendix for a description of the databases and their corresponding variables used in the analysis.

of households with electricity, water, and sewage or even the UN's Human Development Index applied to Brazilian municipalities. Thus, GDP per capita is likely to provide a more complete picture of a mayor's effect on total public consumption than by relying on a single measure of public goods, such as the number of hospitals, or a summary measure that includes only the public goods observed in the data. Second, citizens may value the fact that public goods are often market-enhancing. For instance, the average resident may care to have roads not only for their convenience but also because of their productive capabilities. In this sense, we can interpret GDP per capita as a proxy for the dollar value of the bundle of public goods being provided.

Of course, 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 use as our private sector index, 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.

In column 2 of Table 2, we investigate the relationship between corruption and public consumption. We regress the log of per-capita public consumption on the log of federal funds, the log amount of funds diverted, a dummy for being in the second term, the private sector index, 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.[11] We summarize these characteristics for the mayors in our 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 per-capita public consumption during the term. We also control for the private sector index, mayor's age, log 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 caught stealing have a reelection probability of 15 percentage points lower than those 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, increasing per-capita public consumption positively correlates 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 extensive 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.

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

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.[12] Altogether, we can match 68 percent of politicians to at least one post-office wage.[13]

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 being identified as a corrupt politician affects future earnings: the coefficients on both the audit and corruption dummies are small and statistically insignificant.

To estimate the model, we also need to measure how much mayors earn while in office. In principle, mayors can set their own salaries, and while no readily accessible dataset 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 mayor's 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 a 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 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 goods the incumbent produces depends on the share of funds invested in its production. Fourth, it is important to model the decision to run

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

[13] There are two principal reasons why we could not match the remaining 32 percent. First, if a mayor became self-employed and did not hire any employees over the 8-year period, they would not appear in the RAIS. Second, if the mayor decided to 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.

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 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 producing public goods and, hence, the actual amount of public goods produced. Lastly, ex-mayors found to be corrupt in the past are not associated with lower wages.

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 in their ability to produce public goods.

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 the public consumption Q produced by the local government. 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 parameter $\eta \in [0, 1]$ measures the degree of rivalry, with $\eta = 0$ indicating no rivalry and $\eta = 1$, full rivalry. We represent these preferences with the utility function $u^i\left(c^i_t, \bar{Q}_t\right)$.

If the current mayor decides to run for reelection, they must pay a utility cost κ. And if elected, they derive a utility ρ from being in power. Thus, we can characterize these individuals' 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

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

where θ represents the relative taste for public consumption.[14]

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}. Second, mayors are heterogenous in their ability a_i to produce the public good. It is not observed by voters and it is drawn from a log-normal distribution with mean μ_a and standard deviation σ_a. Because public goods 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 inputs, z_t^{pr}. In the estimation, we assume that the production function for public consumption has the following form:

$$\frac{Q_t}{d_t} = \left(\frac{z_t^{pu}}{d_t}\right)^{\alpha_1} \left(z_t^{pr}\right)^{\alpha_2} \ a_i.$$

By specifying the production function in per-capita terms for both public consumption and inputs, we can 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.[15]

[14]The parameter θ can be interpreted as a purely egoistic parameter: mayors derive utility from experiencing the public good they produce. But it can also be interpreted as an altruistic parameter: it captures to what extent mayors care about the welfare created by public consumption in the municipality.

[15]This assumption is not binding when we estimate the model.

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 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. They also 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. We denote this decision with δ_{run}.

The central government audits municipalities at random with probability p_t^{au}. 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. To account for the possibility that not all mayors accused of corruption are convicted, we assume that a mayor who has been caught stealing is convicted with probability p_t^c. We let $\delta_c = 1$ denote a mayor who has been convicted, zero otherwise. If convicted, 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, on a case-by-case basis, fines that also include 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. Note that mayors consider all these possibilities when deciding how much to steal, how much to consume, and whether to run for reelection.

Most judicial systems are plagued by prosecutorial delays in the conviction of guilty politicians. In our model, we can capture these delays by drawing fines from a distribution whose mean is lower than what is estimated in the data. Moreover, fine data are often misreported. To account for these two issues, we introduce in the estimation of the model an additional parameter δ_τ that allows for the mean of the fine distribution to be different from the mean estimated in the data. Specifically, in our model $\mu_\tau = \bar{\mu}_\tau + \delta_\tau$, where $\bar{\mu}_\tau$ is the mean observed in the data.

Each individual in a municipality owns $\bar{h}$ units of labor, supplied inelastically in return for a wage w that depends on whether the person is currently a mayor or an ex-mayor. Mayors receive a deterministic wage $\bar{w}$ that depends 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. We also assume that the mayor's unobserved ability a_i, is a politician-specific skill that does not affect market wages.[16] Given these assumptions, we model the wage process using a parsimonious specification that depends on education, a second order polynomial in age, and indicators for whether the person resides in a medium or large municipality. We introduce the municipality dummies to allow for municipal heterogeneity in wages. 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, δ_{lm} is equal to one if the municipality's population is above 50,000, and $\epsilon_t \sim N(0, \sigma_{pm})$ is an iid shock.

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 the incumbent or a challenger by comparing their own expected lifetime utility conditional on the incumbent being elected with the corresponding expected lifetime utility if the challenger wins the election.[17] The expectation conditional on the incumbent winning reelection is taken over the distribution of variables that determine the amount of per-capita public consumption the politician would produce in the next term. To account for learning about the

[16]While this might appear to be a strong assumption, it has received some support in the literature. For example, Diermeier, Keane, and Merlo (2005) found in their seminal study of post-congressional wages of U.S. Congressmen that a politician's unobserved skill-type has no effect on their post congressional wages either in the private or public sector.

[17]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.

incumbent's ability, we also condition on the amount of public consumption provided by the incumbent in the current term, whether the municipality was audited and, if audited, the amount stolen. The same approach is used for challengers except that there is no learning. The expected voter's lifetime utility conditional on reelection of the incumbent also includes an election shock, $\varepsilon_{i,t}$, that affects voters at the time of the election. It is assumed to be normally distributed with mean zero and variance one. The exact form of the voters' expected conditional lifetime utilities will be presented after we introduce the recursive formulation of their decisions.

Voters' learning about the incumbent's ability introduces the second main deterrent to corruption: electoral incentives. This aspect of the model is 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. Voters' decisions generate selection on ability since 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 to maximize lifetime utility

$$E\left[\sum_{t=1}^{T}\beta^{t}\left(u\left(c_{t}^{i},\bar{Q}_{t}\right)+\rho 1_{\{\delta_{el,t}^{i}=1\}}-\kappa 1_{\{\delta_{run,t}^{i}=1\}}\right)\right],$$

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

$$c_{t}^{i}+b_{t}^{i}=w_{t}^{i}\bar{h}+1_{\{\delta_{el,t}^{i}=1\}}s_{t}^{i}+R_{t}b_{t-1}^{i}-1_{\{\delta_{el,t-1}^{i}=1,\delta_{au,t-1}^{i}=1,\delta_{c,t-1}^{i}=1\}}g\left(s_{t-1}^{i}\right)\qquad\text{for each } t \text{ and } \omega,$$

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

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:

$$z_t^{pu} + s_t^i = f_t^{pu} \qquad \text{for each } t \text{ and } \omega.$$[18]

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

$$\frac{Q_t}{d_t} = \left(\frac{z_t^{pu}}{d_t}\right)^{\alpha_1} (z_t^{pr})^{\alpha_2}\, a_i \qquad \text{for each } t \text{ and } \omega.$$

At the end of the term, the current mayor also decides whether to run for reelection, and all individuals choose whether to vote for the incumbent.

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 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 contain measurement errors, as only a subsample of a municipality's expenditure is audited. We account for them by introducing a measurement error ν_t^i such that the stealing variable we and voters observe as the outcome of an audit is: $s_t^{i,o} = s_t^i \nu_t^i$, with ν_t^i log-normal distributed with mean 0 and standard deviation σ_ν. The existence of measurement errors imply that voter cannot infer the exact mayor's ability even in case of an audit. They can only update the politician's ability distribution using observed public consumption and stealing.

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, based on the audit outcomes, collects fines from the previous-term mayors that were caught stealing and convicted. The current mayor, who has a given 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 between candidates that cannot credibly commit to their

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

campaign platform. The winning candidate will govern in term t. If an incumbent chooses not to run (i.e. an open election) or loses the election, a new mayor is elected with characteristics drawn from the distribution $f_m(Z)$, where Z includes the mayor's age, education, savings, and ability. 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: Current and Former Mayors

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 ran 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:

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

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

where the wage is drawn from the wage distribution of ex-mayors $p_w(w|Z)$, $\kappa 1_{\{\delta_{run,t}^i=1\}}>0$ if the ex-mayor ran at t but lost the election, and the dummy $1_{\{\delta_{el,t-1}^i=1,\delta_{au,t-1}^i=1,\delta_{c,t-1}^i=1\}}$ indicates that ex-mayors can be fined at t only if they were in power in the previous term, audited, and convicted.

If the current mayor is term-limited, the value function corresponds to the one computed for a past mayor. The decision problem of a current mayor that is not term-limited is more complicated because it includes the choice to run for reelection at the end of the term. If this mayor decides to run, they win 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:

$$V_{WM}^{i}\left(S_{t}^{M},t\right)=\max_{c_{t}^{i},b_{t}^{i},s_{t}^{i}}\ u^{i}\left(c_{t}^{i},Q_{t}\right)+\rho-\kappa+\beta E_{t}\left[V_{M}^{i}\left(S_{t+1}^{M},t+1\right)\right]$$

$$s.t.\ \ c_{t}^{i}+b_{t}^{i}=w_{t}^{i}\bar{h}+s_{t}^{i}+R_{t}b_{t-1}^{i}-1_{\{\delta_{el,t-1}^{i}=1,\delta_{au,t-1}^{i}=1,\delta_{c,t-1}^{i}=1\}}g\left(s_{t-1}^{i}\right)$$

$$z_{t}^{pu}+s_{t}^{i}=f_{t}^{pu}$$

$$\frac{Q_{t}}{d_{t}}=\left(\frac{z_{t}^{pu}}{d_{t}}\right)^{\alpha_{1}}\left(z_{t}^{pr}\right)^{\alpha_{2}}a_{i},$$

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 is not term-limited and chooses to run for reelection, $V_{RM}^{i}\left(S_{t}^{M},t\right)$, as:

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

To estimate the model, we also assume that the decision to run is affected by a shock $\epsilon_{R}\sim N(\mu_{R},\sigma_{R})$. If the mayor is not term-limited and decides to forgo reelection, the corresponding value function is equal to that of an ex-mayor, $V_{PM}^{i}\left(S_{t}^{PM},t\right)$. Thus, a mayor that is not term-limited will choose to run for reelection if

$$V_{RM}^{i}\left(S_{t}^{M},t\right)+\epsilon_{R}\geq V_{PM}^{i}\left(S_{t}^{PM},t\right).$$

Therefore, the value function of this type of current mayor is:

$$V_{M}^{i}\left(S_{t}^{M},t\right)=\max\ \left\{V_{RM}^{i}\left(S_{t}^{M},t\right)+\epsilon_{R},V_{PM}^{i}\left(S_{t}^{PM},t\right)\right\}.$$

Given that both the utility cost of running for election κ and the shock to the running decision enter additively in the politician's utility, we cannot separately identify κ from the mean of the shock μ_{R}. We therefore normalize κ to zero and estimate μ_{R}. The normalization is without loss of generality as we do not need to know κ separately from μ_{R} in our policy evaluations.

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, the amount of public goods produced and the amount stolen in the previous period, whether they were audited and convicted, the mayor's savings, the probability that the municipality will be audited, the conviction probability, 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, convicted, and amount stolen in the past, their savings, and the amount of adjusted per-capita public goods produced by the current mayors.

3.6 Recursive Formulation: Voters

In the data, the probability that a citizen of a municipality runs for election is below 1 percent. We will therefore abstract from the decision to become a candidate for mayor. Under this assumption, a voter's value function corresponds to the value function of a past mayor that did not run for reelection in the current period and whose wages are drawn from the population distribution. We can therefore determine a voter's expected lifetime utility using $V_{PM}^i\left(S_t^{PM}, t\right)$.

Specifically, at the time of the election, the voter's expected lifetime utility from $t+1$ onward, conditional on the incumbent being reelected, is equal to:

$$V_{vot}^i\left(Inc\right) = E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\middle|\, \bar{Q}_t, 1_{\{\delta_{au,t}^i=1\}}, s_t^i 1_{\{\delta_{au,t}^i=1\}}\right] + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ is an iid electoral shock, and we condition on public consumption $\bar{Q}_t$, whether the incumbent was audited $1_{\{\delta_{au,t}^i=1\}}$, and whether and how much the incumbent stole $s_t^i 1_{\{\delta_{au,t}^i=1\}}$ to account for the fact that voters' learn about the incumbent's ability based on observed variables. Voters use those variables to update the incumbent's distribution of ability and, hence, of public consumption for the next term.

The corresponding voter's value function conditional on the challenger winning the election takes the following form:[19]

$$V_{vot}^{i}\left(Ch\right) = E_{t}^{Ch}\left[V_{PM}^{i}\left(S_{t+1}^{PM}, t+1\right)\right].$$

Thus, a citizen of a municipality votes for the incumbent if:

$$V_{vot}^{i}\left(Inc\right) > V_{vot}^{i}\left(Ch\right).$$

To make the estimation of the model feasible, we assume that there is only one representative voter deciding the election. The probability that the incumbent wins is therefore

$$\Phi\left(\bar{V}_{vot}^{i}\left(Inc\right) - V_{vot}^{i}\left(Ch\right)\right),$$

where $\bar{V}_{vot}^{i}\left(Inc\right)$ is the voter's value function if the incumbent is elected without the shock and Φ is the standardized normal cdf.

To compute the voter's expected lifetime utility if the incumbent is elected, we need the distribution of ability conditional on $\bar{Q}_t$, $1_{\{\delta_{au,t}^{i}=1\}}$, and $s_t^i 1_{\{\delta_{au,t}^{i}=1\}}$. Given the complexity of our model, this distribution is not known. To overcome this issue, we approximate the voter's expected utility using the following specification:[20]

$$V_{vot}^{i}\left(Inc\right) \approx E_{t}^{Inc}\left[V_{PM}^{i}\left(S_{t+1}^{PM}, t+1\right)\right] + \delta_1 + \delta_2 \bar{Q}_t + \delta_3 1_{\{\delta_{au,t}^{i}=1, s_t^i=0\}} - \delta_4 s_t^i 1_{\{\delta_{au,t}^{i}=1\}} + \varepsilon_{i,t}.$$

We can then write the probability that the incumbent wins reelection as follows:

$$\Phi\left(E_{t}^{Inc}\left[V_{PM}^{i}\left(S_{t+1}^{PM}, t+1\right)\right] - E_{t}^{Ch}\left[V_{PM}^{i}\left(S_{t+1}^{PM}, t+1\right)\right] + \delta_1 + \delta_2 \bar{Q}_t + \delta_3 1_{\{\delta_{au,t}^{i}=1, s_t^i=0\}} - \delta_4 s_t^i 1_{\{\delta_{au,t}^{i}=1\}}\right).$$

In the estimation of the model, we use a dummy equal to 1 if the incumbent was audited and had stolen in place of $s_t^i 1_{\{\delta_{au,t}^{i}=1\}}$, because the effects on voters' decisions of the amount stolen if

[19] The voter's value function conditional on the challenger winning is computed by taking the expectation over all possible challengers. We do this to reduce the computation burden because otherwise we would have to keep track of all the challenger's state variables.

[20] The approximation is exact if the conditional distribution is normal. If a variable y is normally distributed conditional on a vector x, then $E\left[y\,|x\right] = \mu_y + \Sigma_{yx}\Sigma_x^{-1}\left(x - \mu_x\right)$, where μ_y and μ_x are the means of y and x, Σ_{yx} is the vector of covariances between y and x, and Σ_x is the matrix of variances of x.

audited is not statistically significant.

4 Identification and Estimation

In this section, we present arguments for the identification and estimation of the model's parameters. We estimate all of the model's parameters with the exception of the curvature of the utility function δ, the discount factor β, and the probability of conviction conditional on being audited and caught stealing. For the curvature of the utility function and the discount factor, we follow the literature and set $\delta = 2$ (e.g. Attanasio and Weber (1995)) and $\beta = 0.98$ (e.g. Attanasio, Low, and Sánchez-Marcos (2008)). For the probability of being convicted conditional on being audit and caught stealing, we rely on the estimates presented in Avis, Ferraz, and Finan (2018) and set it to 0.756.[21]

Production Function and Ability Parameters: α_1, α_2, μ_a, σ_a.

Identification. We can identify the production function parameters and the parameters of the ability distribution, if we observe the variables in the production function for first-term mayors. Let $\bar{a}_i = \log a_i - \mu_a$ denote demeaned log ability. Then, the log of the production function of first-term mayors has the following form:

$$\log \frac{Q_t}{d_t} = \alpha_1 \log \frac{z_t^{pu}}{d_t} + \alpha_2 z_t^{pr} + \mu_a + \bar{a}_i, \tag{1}$$

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 $Z_t = [z_t^{pr}, z_t]$, where z_t^{pr} is the private sector inputs. Taking the expectation

[21] To see how we computed this number, note that the probability of being convicted conditional on being audited and caught stealing is $Pr(Conv|Audit = 1\&Stealing = 1) = Pr(Conv = 1\&Audit = 1\&Stealing = 1)/Pr(Audited = 1\&Stealing = 1)$. The joint probability $Pr(Conv = 1\&Audit = 1\&Stealing = 1)$ is equal to the fraction of mayors that incurred a legal action as a result of the audit. Using the data and estimates presented in (Avis, Ferraz, and Finan 2018), this number is 0.072. In our data, the probability of being audited and caught stealing $= Pr(Audit = 1)Pr(Stealing = 1|Audit = 1) = 0.13 * 0.73$. Thus, $Pr(Conv|Audit = 1\&Stealing = 1) = 0.072/(0.13 * 0.73) = 0.756$.

conditional on Z_t of equation (1), we obtain:

$$E\left[\log \frac{Q_t}{d_t}\middle| Z_t\right] = \alpha_1 E\left[\log \frac{z_t^{pu}}{d_t}\middle| Z_t\right] + \alpha_2 z_t^{pr} + \mu_a + E\left[\bar{a}_i\middle| Z_t\right].$$

Since Z_t is mean independent of $\bar{a}_i$, we have that $E\left[\bar{a}_i\middle| Z_t\right] = E\left[\bar{a}_i\right] = 0$, and the previous equation becomes:

$$E\left[\log \frac{Q_t}{d_t}\middle| Z_t\right] = \alpha_1 E\left[\log \frac{z^{pu}}{d_t}\middle| Z_t\right] + \alpha_2 z^{pr} + \mu_a.$$

Thus, the parameters α_1, α_2, and μ_a are identified if the variables $\log \frac{Q_t}{d_t}$, $\log \frac{z_t^{pu}}{d_t}$, z_t^{pr}, and z_t are observed. Lastly,

$$\log \frac{Q_t}{d_t} - E\left[\log \frac{Q_t}{d_t}\middle| Z_t\right] = \bar{a}_i.$$

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

Estimation. We estimate the parameters α_1, α_2, 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} 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. Without the randomization, the estimates of the production function parameters would only apply to the selected sample.

Because a mayor's ability may affect how much federal transfers a municipality receives, we instrument the endogenous variable $\log \frac{z_t^{pu}}{d_t}$ using population thresholds. These population thresholds 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, and ability. We assume that the mayor's ability, whose distribution we have shown is identified, is independent of the remaining characteristics. We can then identify non-parametrically the joint distribution of age, education, and savings, since we observe those variables.

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

Remaining Parameters

There are ten remaining parameters: the relative taste for public consumption parameter, θ; the parameter accounting for prosecutorial delays and measurement errors in the fine data δ_τ; the electoral parameters $\delta_1 - \delta_4$; the rivalry parameter η; the utility from being in power parameter, ρ; the mean and variance of the shock to run for reelection, μ_R and σ_R; and the standard deviation of the measurement error σ_ν.

Given the complexity of the model, it is difficult to provide a proof for their identification. But we can prove that there exists one moment for each parameter that is monotonic in the parameter of interest for any value of the other model parameters. This is a useful result because it is a necessary condition for identification and it provides reliable moments that can be used as targets for the estimation of the ten parameters.

Proposition 1 *The expected value of stealing for second-term mayors is a strictly decreasing function of θ.*

Proof. In the online appendix. ■

To provide the intuition behind this 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. Given the parameters determining the first two factors, Proposition 1 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.[22]

Proposition 2 *Stealing of first-term mayors is a strictly decreasing function of δ_τ.*

Proof. In the online appendix. ■

[22] 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.

For the intuition behind this proposition observe that stealing in the first term can be thought of as buying an asset that provides a negative return equal to the fine if audited and convicted. The negative return is log-normally distributed with mean $\mu_\tau + \delta_\tau$ and standard deviation σ_τ. An increase in δ_τ has therefore the effect of raising the negative return on this asset and, thus, reducing its demand.

Proposition 3 *The probability that a mayor wins reelection is strictly increasing in* δ_1*.*

Proof. In the online appendix. ■

The parameter δ_1 account for general learning about the incumbent ability to provide public consumption. Consequently, larger values of δ_1 produce higher probabilities of reelection.

Proposition 4 *The probability that a mayor wins reelection if it produces more than* $\bar{Q}_{y-th}$*, the y-th percentile of the* $\bar{Q}$ *distribution, is monotonically increasing in* δ_2*.*

Proof. In the online appendix. ■

The electoral parameter δ_2 determines the effect that observed per-capital public consumption has on learning about incumbents' ability and thus, their probability of being reelected. Consequently, the probability that an incumbent wins reelection conditional on providing more than the $y-th$ percentile of the distribution of per-capital public consumption is monotonically increasing in δ_2.

Proposition 5 *The probability that a mayor wins reelection if audited and found to not have stolen is monotonically increasing in* δ_3*.*

Proof. In the online appendix. ■

The electoral parameter δ_3 accounts for learning about the incumbent's ability from a clean audit. This parameter should therefore have a positive effect on their probability of being reelected if a mayor was audited and didn't steal.

Proposition 6 *The probability that an audited mayor who stole wins reelection is strictly decreasing in* δ_4*.*

Proof. In the online appendix. ■

The parameter δ_4 plays a similar role to δ_3 but for stealing. It determines the effect of stealing on learning about the incumbent's ability and on its reelection probability. Thus, the election probability of an incumbent who stole is monotonic in this parameter.

Proposition 7 *Average stealing among mayors is an increasing function of the rivalry parameter η. Moreover, the rate of increase in average stealing with η is larger for municipalities with larger population size.*

Proof. In the online appendix. ■

The intuition behind this proposition is straightforward. Higher η implies a higher degree of rivalry for the public good. The benefits of investing public funds in the production of the public good therefore decline with η. Moreover, the decline is faster for municipalities with larger populations.

Proposition 8 *Given all the other model parameters, ρ, μ_R, and σ_R are identified by the following moments: the probability of running for mayors that were audited and did not divert resources; the same probability for mayors that were audited and caught stealing; the probability of running unconditional of the audit outcome.*

Proof. In the online appendix. ■

For the intuition behind this proposition, note that incumbents running for reelection draw shocks from a distribution with mean μ_R and standard deviation σ_R 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 running and winning than audited mayor who were found to be corrupt. Therefore, a larger ρ increases the difference in the likelihood of running between these two groups of incumbents, independent of μ_R and σ_R. Given ρ, the parameters μ_R and σ_R can then be identified using the probability of running of audited mayors not caught stealing and the probability of running unconditional on the audit outcome, as those two parameters change the location and scale of the two probabilities.

The following proposition establishes that the variance of the fraction stolen is a strictly monotonic function of σ_ν, for any value of the other parameters.

Proposition 9 *The variance of the fraction of public funds diverted is strictly increasing in* σ_ν.

Proof. In the online appendix. ■

Estimation. We estimate the parameters θ, δ_τ, $\delta_1 - \delta_4$, η, ρ, μ_R, σ_R, 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 begin by presenting our model's fit and estimation results. We then use our estimated model to understand why second-term mayors steal more than first-term mayors. We conclude this section by analyzing the effects of various anti-corruption policies, including policy bundles.

5.1 Model Fit

Fit of Moments Used in the SMM Estimation. We use ten moments for our SMM estimation: (i) average stealing among first-term mayors; (ii) average stealing among second-term mayors; (iii) reelection rates conditional on running; (iv) reelection rates conditional on running and on public consumption being above the 75^{th} percentile; (v) reelection rates conditional on running and having been caught stealing; (vi) reelection rates conditional on running and having a clean audit; (vii) probability of running conditional on having a clean audit; (viii) probability of running conditional on having been caught stealing; (ix) difference in average stealing between small and large municipalities; (x) the unconditional probability of running.

In panel A of Table 3, we compare these data moments with our simulated moments. Our model matches these moments extremely well. Importantly, we can exactly match the actual fraction stolen by both first (5.6%) and second-term mayors (7.3%), which we use to identify the effect of delays and measurement error in the fine data and the relative taste for public consumption, respectively. To identify the voting parameters $(\delta_1, \delta_2, \delta_3, \delta_4,)$, we rely on the reelection rates of different subgroups. It is reassuring that we match all of those moments almost perfectly. The largest difference between the data and simulated moments is in the probability of running for reelection conditional on a clean audit, which we use to identify the utility from being in power and the two moments of the distribution of the running shock. But even in this case, the difference is only 0.2 percentage points.

Fit of Moments Not Used in the Estimation. In panel B of Table 3, we assess how well our model matches moments not used in the estimation. We compare eight moments and, in general, do quite well. For example, we can match the share of incumbents who forgo reelection (29.4% versus 30.2%). 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, 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. Even though 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 -0.38 versus a mean in the data of -0.42.

In Section 2, we documented the right skewness of the stealing distribution. The skewness stems from 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 the municipality receives, 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. The last part of Table 3 shows that our model, despite its parsimony, can generate this skewness.

5.2 Parameter Estimates

Preference Parameters. We present estimates of our model parameters in Table 4. We estimate a relative taste for public consumption equal to 0.075. This implies that individuals value the utility from private consumption about 12.7 times more than the utility from public consumption at the mean of their public and private consumption. Our rivalry parameter ($\eta = 0.97$) suggests a relatively high degree of rivalry in public consumption. Mayors also enjoy high levels of utility from being in power. We estimate the utility from being in power to be 1.67, which corresponds to 73% of the average utility a mayor experiences from private and public consumption.

Electoral Parameters. Our estimate of the informational value of incumbency (δ_1) is 0.341. To interpret its magnitude, recall that the probability of getting reelected is 57.3%. If we re-simulated the model with this parameter set to zero, this probability decreases to 50.4%, a 12.7% decline. Voters also punish mayors that have been found to be corrupt. If we re-simulate the model with the coefficient on having been caught stealing ($\delta_4 = -0.270$) set to zero, the probability of reelection conditional on stealing increases by 9.8%, and stealing among first-term mayors increases by 5.1%. Similarly, voters reward mayors who were audited and not found to be corrupt. Here, the effects are modest because a change in δ_3 impacts both the mayor's decision to run and reelection probabilities. But they have the expected sign. If we set δ_3 to zero, the probability of running conditional on not stealing declines from 77.2% to 76.1%, and the reelection probability goes down from 57.3% to 56.7%.

The difference between the incumbent and the average challenger in expected public goods production also affects electoral decisions. If we set this parameter (δ_2) to zero, the probability of reelection declines by 0.7 percentage points (from 57.3% to 56.6%). But this small effect masks larger effects in the probability of running: the probability of running conditional on stealing goes up by 2 percentage points, and the probability of running conditional on not stealing goes down by 2.1 percentage points.

Production Function. Panel C of Table 4 reports the estimated coefficients for the production function. We estimate that a 10% increase in public inputs increases public consumption by 1.9%, whereas a 10% increase in private inputs increases public consumption by 4.3%. An important source of unobserved heterogeneity in our model is the mayor's ability to produce public goods. We estimate that the mean of log-ability is -0.431 in the first term and -0.424 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 4.5 percent. The difference in coefficients between first and second-term mayors implies an increase in per-capita public consumption of 1 percent. The standard deviation of log-ability equals 0.439.

Fine Parameters. For the mean and standard deviation of the fine distribution, we estimate coefficients of 0.094 and 0.284, respectively. To account for possible prosecutorial delays and measurement errors in the fine data, we allow for a reduction in the mean of the fine distribution, which we estimate at -0.890. Overall, these estimates suggest that corrupt mayors, on average, have to repay the original amount and pay a fine equal to 45% of the amount stolen.

Wage Process Parameters. In Panel E of Table 4, 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.

5.3 Model Simulations

Elections can play two important roles in promoting voters' welfare. First, because voters care about public goods, politicians can refrain from stealing and provide more public goods to improve their reelection chances. This is often referred to as an *electoral 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 are still able to provide more public goods. Our model can provide a third channel. Politicians in our model can also save, and if second-term mayors have accumulated more wealth than first-term mayors, then this too can explain part of the difference in stealing between first and second-term mayors. A key feature of our model is the ability to identify the effects of all three mechanisms separately.

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

Second-term mayors steal about 19 percent more than first-term mayors.[23] Public consumption is also higher among second-term mayors, which indicates positive selection. We can see this more clearly in the subsequent rows. Second-term mayors are, for example, wealthier, younger, and have higher ability. Interestingly, they are also less likely to have a college degree, consistent with what we see in the data.

We can decompose the differences in stealing and public consumption between first and second-term mayors due to reelection incentives, selection effects, and wealth effects. It is worth noting that in a static model, we would typically model wealth as a fixed attribute of the politician, similar to education level or ability. In this case, including wealth as part of the selection effect would make sense. But in our dynamic model, wealth is endogenous as politicians choose how much to save over their life cycle. Thus, it is important to distinguish selection effects separately from wealth effects.

To isolate the effect of each channel, we simulate the model under different environments.

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

To measure the effects of *reelection incentives*, we compare the decisions of first-term mayors with and without the possibility of reelection. For this comparison, mayors are drawn from the same distribution of observable and unobservable characteristics and, thus, are identical. The only difference lies in the possibility of reelection. For *selection effects*, we compare the decisions of first-term mayors to those of first-term mayors with the same distribution of observable and unobservable characteristics as second-term mayors. Thus, these two groups of first-term mayors have the same incentives but differ in their observable and unobservable traits. Finally, we calculate *wealth effects* by comparing the decisions of first-term mayors to those of first-term mayors who have been given the same distribution of initial savings as second-term mayors. Thus, we will compare two groups with the same incentives, the same distribution of observable and unobservable traits, but different savings levels. In Table 6, we present the effect of each channel as a percentage of the total of the three effects.

The removal of reelection incentives would increase stealing by 44.1%. This effect is approximately six times the size of the effect of selection (59.8% vs 9.0%) and twice the size of the wealth effect (59.8% vs 31.2%). In contrast, the selection effect is only a third of the size of the wealth effect.

Regarding per-capita public consumption, we see that most of the difference between first and second-term mayors come from selection effects, which account for more than 84.9% of the total effects. This result makes sense given that ability has a direct effect on the production of public consumption and the positive selection of second-term mayors relative to first-term mayors. Reelection incentives account for an additional 9.4% of the total effects, whereas wealth effects contribute the remaining 6%. Unlike the selection effects, the effects of these other channels are indirect.

In sum, reelection incentives play a significant role in a mayor's decision to engage in corruption. The selection effects generated by elections are also important, particularly for public consumption and, as we will document, welfare. The results of this decomposition are useful for the design of anti-corruption policies. Given our findings, it stands to reason that policies enhancing reelection incentives might effectively combat corruption. Moreover, if we can extend a politician's time horizon, we may not only heighten reelection incentives but also allow voters more opportunities to screen for better politicians. We will explore such policies in the next section.

5.4 Policy Evaluation

Our model offers several advantages for evaluating policies. First, it allows us to compare several anti-corruption policies, all within the same setting. Second, it allows us to bundle policies. This is important because, as we will show, each individual policy has its limitations, and by combining certain policies, we can mitigate these weaknesses to limit corruption further. Third, while combatting corruption is a meaningful objective, we ultimately care about how policy affects individual welfare or, equivalently, an individual's willingness to pay for a particular policy. Importantly, our model allows us to calculate an individual's willingness to pay for each of our policies.

In this section, we use our estimated model to evaluate four individual anti-corruption policies: 1) an increase in the audit probability; 2) the clean record act policy; 3) extending term limits to a third term; and 4) doubling mayors' wages. We compare these policies to our base case, which sets the audit probability to 5% in all periods. In the subsequent section, we demonstrate how to combine some of these policies to reduce corruption further.

Brazil's Audit Program (Audit). In 2003, Brazil's Federal Government introduced a program to reduce corruption in local governments. As we discussed in Section 2, the program audits 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 plot on the right y-axis the average fraction stolen for different audit probabilities. As expected, stealing declines as the probability of an audit increases. For example, as we go from a 6% audit probability to a 20%, average stealing decreases by almost half, from 9.3% to 5.7%. For the case of Brazil's audit program, our simulations suggest that the program has reduced corruption by about 3.9 percentage points, which represents a 34.8% reduction from our base case.

Despite their effectiveness, audits can be quite expensive. According to Zamboni and Litschig (2018), the direct cost of a single audit in 2004 was estimated to be around US$50,000.[24] From

[24]Note that this estimate only includes transportation costs, the auditors' salaries, and per diem. It excludes any

2003-2008, the program had audited 1,401 municipalities at an estimated total direct cost of \$70.05 million.

Given these costs, a natural question is whether this audit probability is optimal. To answer this question, we need to measure a citizen's willingness to pay for the audit policy at different audit probabilities after accounting for the corresponding cost. To measure the willingness to pay of a citizen, we compute 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 yearly income using a life expectancy for Brazil of 75 years.

In Figure 1, we plot on the left y-axis a citizen's willingness to pay for an audit policy at different audit probabilities. As we can see from the figure, the current government policy is not too far off from the optimal one. Based on a citizen's willingness to pay, it would be optimal for the program to reduce its audit probability from 16.8% to 16%. At this point, the willingness to pay is maximal at 0.40, suggesting that the representative citizen would be willing to pay 0.40% of their annual income over their life cycle to implement such a policy. For comparison, the willingness to pay for the current audit program is 0.39.

Besides increasing the probability of an audit, we can also induce deterrence effects by increasing the fines associated with getting caught or the probability of conviction. We consider such policies in Appendix Figures 4 and 5. To simulate an increase in fines, we increase the mean of the lognormal distribution for fines (x-axis) and compute average stealing (y-axis). As expected, larger expected fines decrease stealing among mayors. In principle, one could increase the size of a fine sufficiently high so that mayors refrain from corruption. In our model, this occurs when mayors expect to pay 16 times the amount they have stolen. Why the government does not impose such punishments is an interesting question but one that we feel is beyond the scope of this paper. In Appendix Figure 5, we simulate the effects of decreasing and increasing the probability of being convicted. Stealing monotonically decreases with a higher probability of conviction. Interestingly, corruption does not go to zero even if one increases the probability of conviction from 0.76 to one. Such a policy, if feasible, would only decrease stealing by 5.6%.

overhead costs. If we, instead, divided the program's annual budget (\$25 million) by the number of yearly audits, we would estimate an average cost of \$150,000 per audit.

Clean Record Act (CRA). 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 law raises the implicit cost of corruption and thus incentivizes political candidates to refrain from stealing. In our model, we can simulate the effects of this policy by prohibiting incumbents who have been convicted of stealing from participating in future elections.[25]

In Figure 2, we graph the results of the CRA policy, along with the other individual policies. In plot (a), we present average stealing across all mayors. In plot (b), we report average stealing by term. In plot (c), we present the voters' willingness to pay for each policy. In the case of the audit policy, we set the audit probability to the optimal one after accounting for its cost.

We find that the CRA policy reduces corruption by 30.9%, relative to the base case. The policy affects first-term mayors (54.6%) more than second-term mayors. In fact, the policies even led to a slight increase among second-term mayors relative to the base case. This is expected because the CRA policy works via electoral incentives, which exist only for first-term mayors. Despite its minimal effect on second-term mayors, citizens are willing to pay 1.00% of their lifetime annual income for this policy. As shown in Figure 2c, this is much higher than the willingness to pay for the audit program or any other individual policy we consider.

3-term limit policy. In the previous section, we documented the importance of electoral incentives and selection effects in explaining the levels of stealing and public consumption we observe in the data. Because both of these are a byproduct of elections, increasing the number of terms a mayor can stay in power may help to reduce corruption. To evaluate this policy, we simulate the effects of allowing mayors to be reelected for a third term. Under this scenario, corruption reduces by 12.11% relative to the base case, and citizens are willing to pay 0.47% of their lifetime income for such a policy (see Figure 2). The largest reduction in stealing comes mostly from second-term mayors who now have reelection incentives and reduce their stealing by 25.24% relative to second-term mayors in the base case. Interestingly, corruption in the third (final) term is much lower than in the other two terms. This again highlights the selection effect that elections induce. Relative to first and second-term mayors, third-term mayors are much more positively selected in terms of their ability and initial wealth – traits that are both associated with less stealing. For

[25] We are implicitly assuming that if caught, the mayor will be convicted with probability p^c before the next electoral term, i.e., within four years.

example, the average ability among third-term mayors is 10.1 and 4.8 percent higher than the average ability of first and second-term mayors, respectively.

Given the policy's strong incentives and positive selection effects, we might wonder why average stealing over the three terms reduces to only 8.5%. This is due to political turnover. Even though third-term mayors are highly selected, only a small fraction of municipalities are expected to elect mayors to a third term.

Doubling wages. Our final individual policy is to double the mayor's salary. The rationale for this policy is simple. Increasing the value of holding office will increase the incentives for being elected, which should discourage mayors from stealing. In Figure 2a, we report the policy effects when we account for the cost of the policy. We see that a doubling of wages decreases corruption by 9.8%. Despite its implied costs, citizens are willing to pay 0.47% of their lifetime annual income to adopt this policy.

Robustness

A key assumption underlying our simulations is that the model's parameters are invariant to these policy reforms. Because we do not model the decision to become a politician, one might be concerned that our policies might affect the ability distribution of the candidate pool. In particular, we might think that the candidate pool might become more positively selected as we increase the probability of an audit or extend term limits.

We test our assumption by exploiting the randomness of the audits in our data and by evaluating whether being audited affects the type of candidates running in the subsequent election. Specifically, we regress characteristics of the candidate pool (e.g., the share of candidates with a college degree) on an indicator for whether the municipality had been audited. In Panel A of Appendix Table 9, we present these results only for the 2004 elections, whereas in Panel B, we present pool data for three elections, 2004, 2008, and 2012. For the results reported in Panel B, the regression also accounted for municipality and election year fixed effects. We do not find any evidence that having been audited affects either the size or composition of the candidate pool, as measured by their gender, education, or age.

We also re-simulate all of our policies under the assumption that they increase the average ability of the candidate pool. Specifically, we increase the mean of the ability distribution by 25%. For completeness, we also consider the opposite effect (i.e., 25% decrease) in case the policies reduce the average ability of the candidate pool. We present these results in Appendix Table 10. In general, the effects of the policies become stronger as the candidate pool becomes more positively selected. Nevertheless, the effects are similar to our original simulations, and our general conclusions remain unchanged.

5.5 Policy Bundles

Thus far, 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 effectively reduces corruption but is expensive. The same can be said about the wage policy. The CRA policy reduces corruption among mayors with electoral incentives but does not affect term-limited mayors. The 3-term policy also effectively reduces stealing, but its incentive effects are limited with so few mayors surviving to a third term. In Table 7, 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 policy bundles on corruption.

Among our individual policies, the Clean Records Act produced the highest willingness to pay. But as we have noted, it is less effective on term-limited mayors. Thus, we can enhance this policy by combining it either with the 3-term policy, an audit policy, or both. We present these results in Figure 3. Combining the CRA with a 3-term policy reduces corruption by 60.6%, compared to 34.8% (for the CRA) or 12.1% (for the 3-term) when the policies were applied separately. The fact that the effects of the combined policy are larger than the sum of the two individual policies is evidence of important complementaries. We can see this when we compare corruption levels across terms (Figure 3b). By itself, the CRA did not affect corruption among second-term mayors. This bundled policy reduces corruption among second-term mayors by 65.8%. Importantly, citizens are willing to pay 1.21% of their lifetime annual income for this policy bundle, which is 22 basis points higher than their willingness to pay for just the CRA.

We find smaller improvements for a policy bundle that combines the 3-term with an 11% probability of being audited. We set the audit probability to 11% because it is the audit probability that maximized the willingness to pay for this combined policy. As shown in Figure 3a, combining the 3-term with an 11% audit probability reduces corruption by 31.0%. While this is an impressive reduction, it is not as large as the 34.8% reduction one could achieve by implementing just the (optimal) audit policy. Nevertheless, citizens would prefer this policy bundle to the individual optimal audit policy by almost a factor of two (WTP for bundle = 0.67% versus WTP for audit = 0.39%). By combining the 3-term with an audit policy, we not only strengthen the incentives of the 3-term policy but also reduce the costs of the audit policy by lowering its audit probability.

We also consider combining the audit policy with the CRA policy. The audit policy should enhance the incentive effects of the CRA policy, and the CRA policy helps lower the optimal audit probability. As shown in Figure 3, this policy can be quite effective at reducing corruption (47.5% relative to baseline), with most of the reduction coming from first-term mayors. The willingness to pay for the CRA-audit policy bundle is almost identical to the 3-term+CRA policy bundle.

Citizens' highest willingness to pay combines all three policies: CRA, 3-term, and 8% audit probability. For this policy bundle, citizens are willing to pay almost 1.29% of their lifetime annual income (see Figure 3c). The policy also reduces average stealing by an impressive 81.3% relative to the base case. It is also equally effective among first-term mayors (81.3% reduction relative to the base case) and second-term mayors (76.2% reduction relative to the base case).

The last policy we consider is to combine the audit policy with a doubling of mayors' wages. In Figure 2a, we saw that doubling mayors' wages would reduce corruption by 9.8% under the base case probability of an audit of 5%. When combining these two policies, the optimal audit probability is 9%, and stealing falls by 22.1% relative to the base case (see Figure 3a). If we compare this reduction to the effect of each policy separately – 9% audit policy = 12.3% reduction in corruption and wage policy = 9.8% reduction in corruption – we see that the effects of the combined policy are almost additive. Also, compared to the other policy bundles, this policy resulted in our lowest willingness to pay.

Mayoral Effectiveness vs Dynamic Corruption Incentives. In Section 5.3, we quantified the effects of electoral incentives and selection on the differences in stealing and per-capita consumption between first and second-term mayors. In comparing policies, it is also useful to distinguish the welfare impacts that come through changes in a mayor's ability versus changes in other aspects of the model. In Appendix Figure 6, we report the willingness to pay for each policy under the assumption that second-term mayors are drawn from the same ability distribution as first-term mayors, thus eliminating selection effects on ability. As we can see from the figure, welfare drops significantly. For example, voters' willingness to pay for the 3-term policy – a policy that induces substantial positive selection – drops from 0.47% to 0.16%, a 65% decline. Overall, these results suggest that selection on ability is an important channel for generating the effects of these policies on voter's welfare.

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 deciding what to consume, save, steal, and whether to seek reelection. 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 reelection 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 reelection incentives (i.e., the 3-term and the Clean Record Act) with an increase in the audit probability could reduce corruption at a lower cost. Our findings suggest that residents in Brazil are willing to pay 1.3% 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 30% of what they would be willing to pay for any individual policy.

Our estimates and policy analysis clearly apply to the case of Brazil. However, our framework is quite general and can be used to understand local corruption in any setting where 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.

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).

Di Tella, Rafael, 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.

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).

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.

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.

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).

Zamboni, Yves, and Stephan Litschig. 2018. "Audit risk and rent extraction: Evidence from a randomized evaluation in Brazil." *Journal of Development Economics* 134:133–149.

Tables & 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 the private sector inputs, 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. Robust standard errors are reported in brackets, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Comparing Data Moments with Model Moments

	Model	Data	Corresponding Parameter
<u>Panel A</u>: Moments Used in the Estimation			
Average Stealing in the Second Term	0.073	0.073	θ
Average Stealing in the First Term	0.056	0.056	δ_τ
Pr[Reelection]	0.573	0.573	δ_1
Pr[Reelection|Public Cons. > 75^{th}perc.]	0.612	0.612	δ_2
Pr[Reelection|Audited and Stealing = 0]	0.652	0.650	δ_3
Pr[Reelection|Audited and Stealing > 0]	0.516	0.515	δ_4
Pr[Running|Audited and Stealing = 0]	0.772	0.779	ρ
Pr[Running|Audited and Stealing > 0]	0.686	0.684	μ_R
Difference in Stealing, Small Vs Large Municipalities	0.241	0.240	η
Pr[Running]	0.718	0.718	σ_R
Variance of Fraction Stolen	0.103	0.103	σ_ν
<u>Panel B</u>: Moments Not Used in the Estimation			
Fraction Incumbents Not Running for Reelection	0.294	0.302	
Mean of Log Ability, Second term	-0.38	-0.42	
Fraction of Mayors not Stealing	69.0	73.3	
25th Percentile Fraction of Funds Stolen	0.00	0.00	
Median Fraction of Funds Stolen	0.040	0.021	
75th Percentile Fraction of Funds Stolen	0.092	0.076	
90th Percentile Fraction of Funds Stolen	0.166	0.196	
Per-capita consumption	3.863	3.809	

Notes: This table presents the moments used to estimate the model's parameter, and the moments used to evaluate the model's goodness of fit. Column 1 reports simulated moments based on 500 simulations for each municipality. Column 2 reports the data moments. All reelection rates are conditional on running. Per-capita consumption is measured with the rivalry parameter.

Table 4: Parameter Estimates

	Parameter	Estimate	Standard Error
Panel A: Preference Parameters			
Relative Value of Public Consumption	θ	0.075	0.005
Rivalry parameter	η	0.966	0.009
Utility of being in power	ρ	1.666	0.009
Panel B: Electoral Parameters			
Informational Value of Incumbency	δ_1	0.341	0.025
Difference in public consumption (incumbent vs challenger)	δ_2	0.583	0.062
Dummy if Audited and not Caught Stealing	δ_3	0.387	0.085
Dummy if Audited and Caught Stealing	δ_4	-0.270	0.009
Mean of Running Shock	μ_R	-0.289	0.011
Std Dev. of Running Shock	σ_R	1.378	0.013
Panel C: Production Function Parameters			
Log Public Inputs	α_1	0.194	0.044
Log Private Inputs	α_2	0.434	0.088
Mean Log Ability Distribution	μ_a	-0.431	0.407
Mean Log Ability Distribution Second term	μ'_a	-0.424	0.411
Standard Deviation Log Ability Distribution	σ_a	0.439	0.027
Std. Dev. Measurement Error in Stealing	σ_ν	0.609	0.154
Panel D: Fine Parameters			
Mean of Fine Data	μ_τ	0.094	0.020
Standard Deviation of Fine Data	σ_τ	0.284	0.045
Measurement Error in Fine Data	δ_τ	-0.890	0.030
Panel E: Wage Process Parameters			
Constant	γ_0	-3.086	0.509
College Education	γ_1	4.097	0.179
Age	γ_2	1.035	0.183
Age^2	γ_3	-0.098	0.018
Medium Municipality	γ_4	1.144	0.211
Large Municipality	γ_5	3.215	0.339
Standard Deviation	σ_{pm}	6.335	0.202

Notes: Panels A and B and the last line of Panel D report the SMM estimates of the model parameters. The standard errors are computed using the asymptotic distribution of the SMM estimator. When deriving it, we account for the variation introduced by the parameters estimated outside the model by using a two-step estimator. Panel C 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 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. Panel D 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 5: Political Outcomes & Selection – Baseline Model

	Full Sample (1)	First-term mayors (2)	Second-term mayors (3)
Fraction Stolen	0.097	0.091	0.108
Per-Capita Public Cons.	3.837	3.746	4.001
Age	49.482	50.177	48.237
Initial Wealth	442.353	336.139	632.520
College Education	0.363	0.389	0.315
Ability	0.717	0.696	0.753

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 6: Decomposing Corruption: Effects of Incentives, Selection, and Savings

	$\%\Delta\{2^{nd} - 1^{st}\}$ mayors (1)	Reelection Incentives (2)	Selection Effect (3)	Wealth Effect (4)
Fraction Stolen	19.12%	59.79%	9.03%	31.18%
Public Consumption	6.79%	9.40%	84.96%	5.64%

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

Table 7: Policy Limitations and Potential Solutions

Individual Policies	Pros	Cons	Potential Solution
Audits	Most effective single policy in reducing corruption	Expensive	Combine with alternative policies to reduce optimal audit probability
CRA	Most effective single policy in terms of WTP	Limited effect on term-limited mayors	Combine with alternative policies that incentives term-limited mayors (e.g., 3-term limits or audits)
3-term limit	Reduces corruption, induces positive selection	Effects limited due to political turnover	Combine with alternative policies that creates stronger incentives (e.g., CRA or audits)
Double Wages	Reduces corruption	Expensive	Combine with audit policy to increase threat of punishment

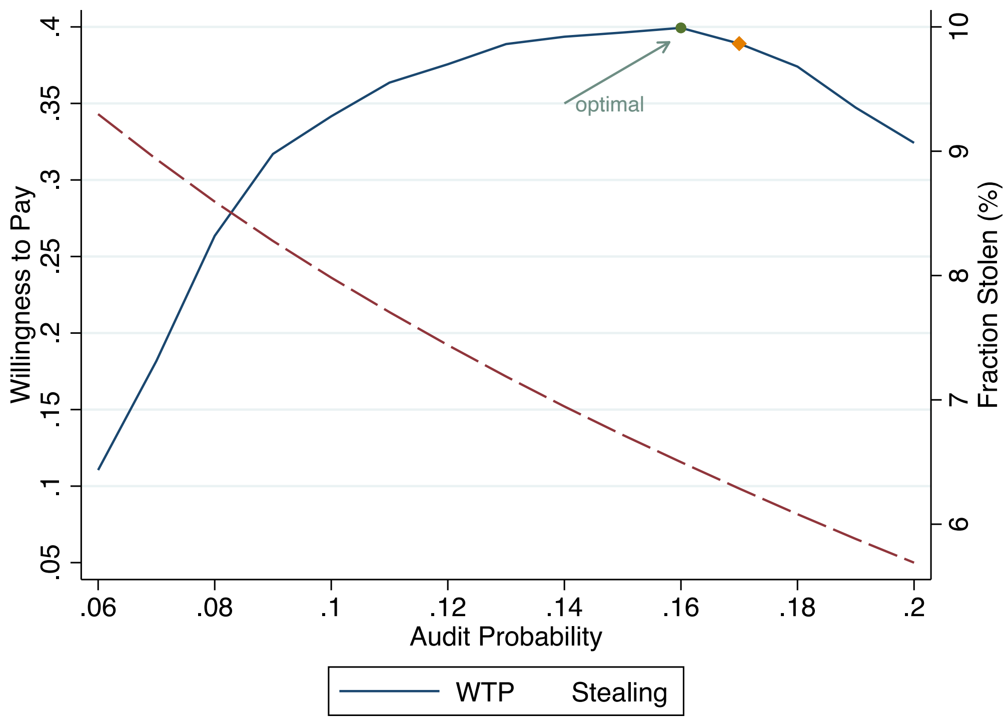
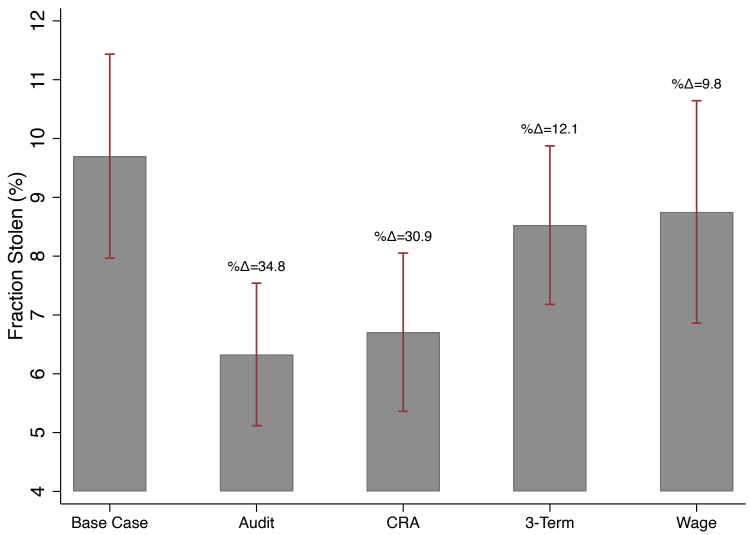
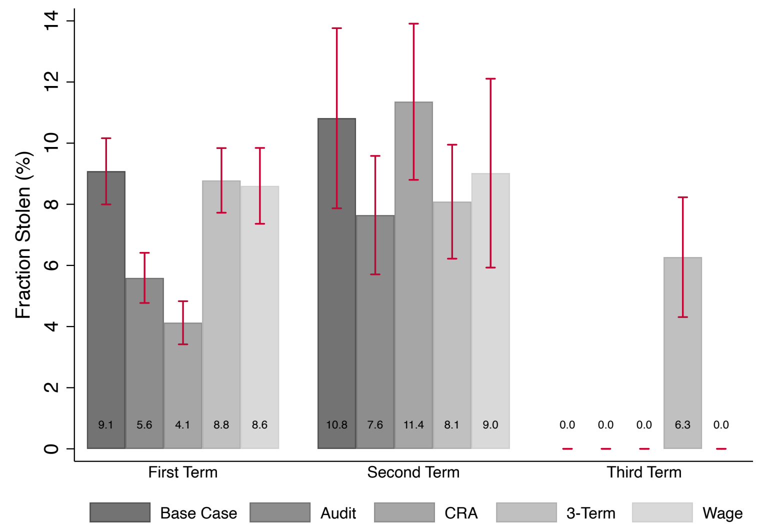


Figure 1: Optimal Audit Policy

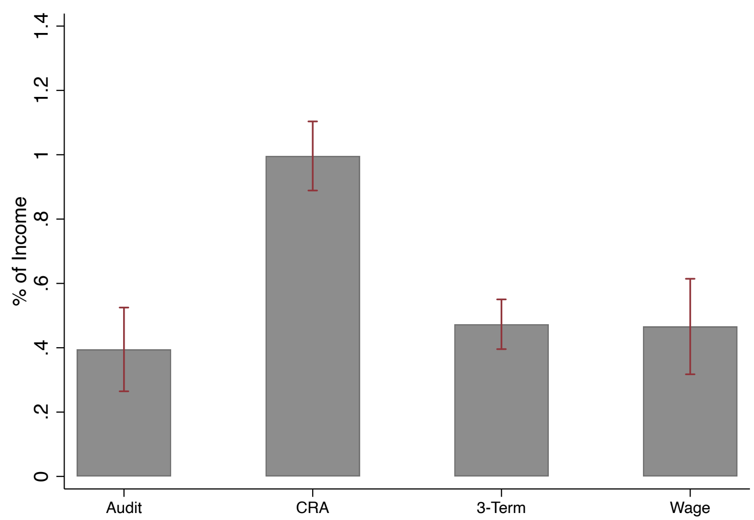
Notes: This figure presents results based on 500 simulations for each municipality.



(a) Average Stealing



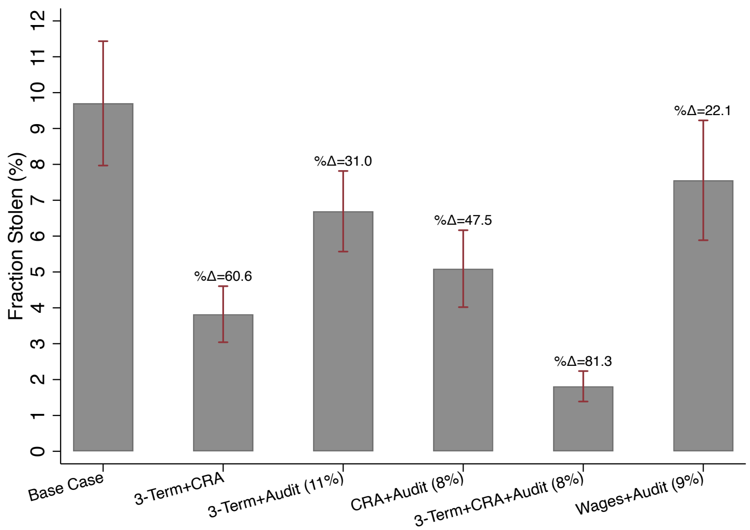
(b) Average Stealing by Term



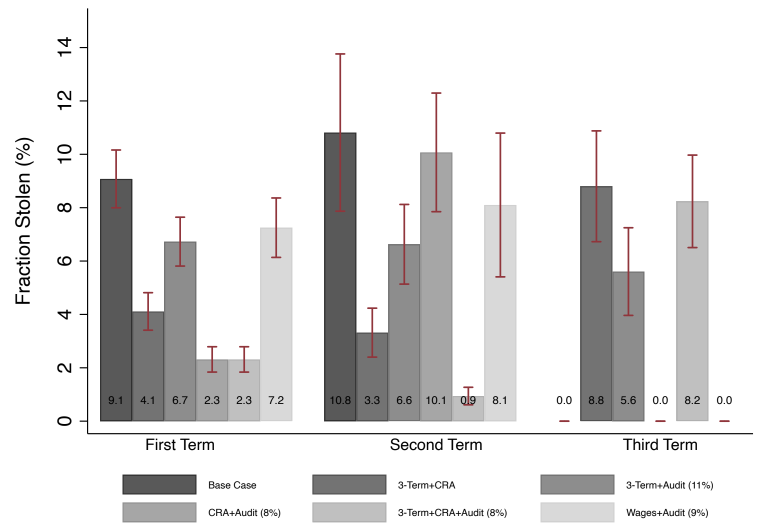
(c) Willingness to Pay

Figure 2: The Effects of Anti-Corruption Policies on Corruption

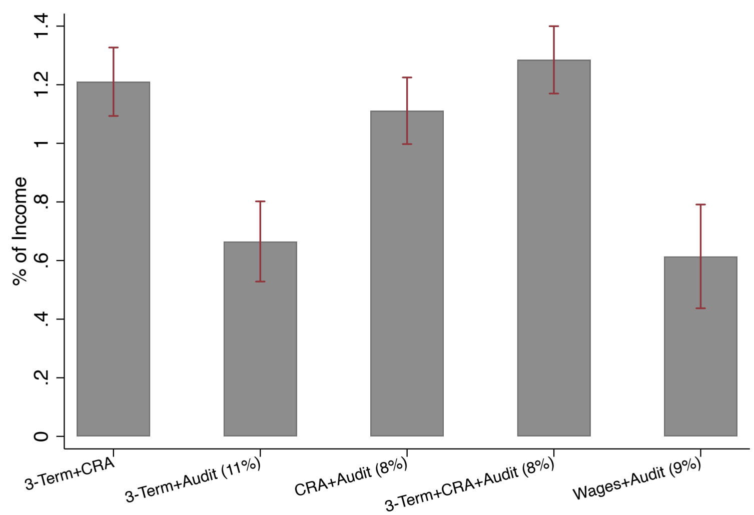
Notes: Each reported statistic is computed based on 100 simulations for each municipality. The 95% confidence intervals are computed by drawing the model parameters from their asymptotic distribution 256 times, calculating the relevance statistics for each draw of the parameters, and computing the confidence intervals using the 256 observations.



(a) Average Stealing



(b) Average Stealing by Term



(c) Willingness to Pay

Figure 3: The Effects of Anti-Corruption Policy Bundles on Corruption

Notes: Each reported statistic is computed based on 100 simulations for each municipality. The 95% confidence intervals are computed by drawing the model parameters from their asymptotic distribution 256 times, calculating the relevance statistics for each draw of the parameters, and computing the confidence intervals using the 256 observations.

A Proofs

A.1 Proof of Proposition 1

We will show that the expected value of stealing for second-term mayors is strictly decreasing in the taste for the public good parameter θ. 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.

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:

$$\frac{Q}{d} = \left(\frac{z^{pu}}{d}\right)^{\alpha_1} (z^{pr})^{\alpha_2} \;\; a_i = \bar{\lambda}(s),$$

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:

$$\begin{gathered} \max_{c,b,s} u(c) + v\left(\bar{Q}\right) + \beta E\left[u(\tilde{c})\right] + \beta E\left[v\left(\tilde{\bar{Q}}\right)\right] - \theta\eta\log(d) \\ s.t. \quad c = y - b + s, \qquad \tilde{c} = \tilde{y} + Rb - \tilde{R}s, \\ Q = \lambda(s), \qquad \tilde{R} = \delta e^{\mu+\sigma z}, \qquad z \sim N(0,1), \qquad \delta \in [0,1], \end{gathered} \tag{2}$$

where $v\left(\bar{Q}\right) = \theta\bar{Q}$, y includes labor earning and a possible fine for the amount stolen in the previous term, $\tilde{\bar{Q}}$ is adjusted 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:

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

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

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

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:

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

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

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

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

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

We have

$$
\begin{aligned}
& R^2 E\left[u''(\tilde{c})\right] E\left[\left(R-\tilde{R}\right)^2 u''(\tilde{c})\right] - R^2 E\left[\left(R-\tilde{R}\right) u''(\tilde{c})\right]^2 \\
&= R^2 E\left[u''(\tilde{c})\right]\left\{E\left[R^2 u''(\tilde{c})\right] + E\left[\tilde{R}^2 u''(\tilde{c})\right] - 2RE\left[\tilde{R} u''(\tilde{c})\right]\right\} \\
&- \left\{R^4 E\left[u''(\tilde{c})\right]^2 + R^2 E\left[\tilde{R} u''(\tilde{c})\right]^2 - 2R^3 E\left[u''(\tilde{c})\right] E\left[\tilde{R} u''(\tilde{c})\right]\right\} \\
&= R^2 E\left[u''(\tilde{c})\right] E\left[\tilde{R}^2 u''(\tilde{c})\right] - R^2 E\left[\tilde{R} u''(\tilde{c})\right]^2 = R^2\left\{E\left[u''(\tilde{c})\right] E\left[\tilde{R}^2 u''(\tilde{c})\right] - E\left[\tilde{R} u''(\tilde{c})\right]^2\right\} \\
&> R^2\left\{E\left[\sqrt{(-u''(\tilde{c}))}\sqrt{\tilde{R}^2(-u''(\tilde{c}))}\right]^2 - E\left[\tilde{R} u''(\tilde{c})\right]^2\right\} = R^2\left\{E\left[\tilde{R} u''(\tilde{c})\right]^2 - E\left[\tilde{R} u''(\tilde{c})\right]^2\right\} = 0.
\end{aligned}
$$

where the inequality follows from the Cauchy-Schwarz inequality: $E[XY]^2 < E[X^2]\,E[Y^2]$, applied to $X = \sqrt{(-u''(\tilde{c}))}$ and $Y = \sqrt{\tilde{R}^2(-u''(\tilde{c}))}$. 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:

$$
\begin{aligned}
u'(c^*) - \beta R E\left[u'\left(\tilde{c^*}\right)\right] &= 0 \\
v'\left(\bar{Q}^*\right)\lambda'(s^*) + \beta E\left[u'\left(\tilde{c^*}\right)\left(R-\tilde{R}\right)\right] &<= 0,
\end{aligned}
\tag{3}
$$

where the inequality accounts for the possibility that $s^* = 0$ (c^* is always greater than 0, as $\lim_{c\to 0} u'(c) = \infty$). Consider an increase in θ. This change only affects $v'\left(\bar{Q}\right)$. Specifically, $\frac{\partial v'(\bar{Q})}{\partial\theta} = 1 > 0$. Moreover, $\lambda'(s^*) < 0$. Hence, at the new θ, but old solution for s and c, the left hand side of (3) decreases and the inequality changes to

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

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 (3) with respect to s, which take the following form:

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

where the inequality follows from $v\left(\bar{Q}\right)$ 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 (3) 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 (3) 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.

A.2 Proof of Proposition 2

In Proposition 1, we have shown that stealing in the first term has a negative return equal to $\tilde{R}=\delta e^{\mu+\delta_{\tau}+\sigma z}$ distributed log-normally with mean $\mu+\delta_{\tau}$ and standard deviation σ (see problem (2). Thus an increase in δ_{τ} raises the negative return on stealing and reduced the quantity demanded of this variable. To show this, without loss of generality, we focus on a three-period problem.

A first-term mayor chooses private consumption, savings, and stealing as the solution of the following problem:

$$\max_{c_1,b_1,s_1,\tilde{c}_2,\tilde{b}_2,\tilde{s}_2,\tilde{c}_3} u(c_1)+v(\bar{Q}_1)+p(s_1)\beta E\left[u(\tilde{c}_2)+v\left(\tilde{\bar{Q}}_2\right)|Inc\right]+(1-p(s_1))\beta E\left[u(\tilde{c}_2)+v\left(\tilde{\bar{Q}}_2\right)|Ch\right]+$$
$$+\beta^2 E\left[u(\tilde{c}_3)+v\left(\tilde{\bar{Q}}_3\right)|Inc\right]-\theta\eta\left(1+\beta+\beta^2\right)\log(d)$$
$$s.t. \quad c_1=y_1-b_1+s_1, \quad \tilde{c}_2=\tilde{y}_2-\tilde{b}_2+R_2b_1+\tilde{s}_2 1_{\{\delta^i_{el,2}=1\}}-\tilde{R}_2 s_1, \quad \tilde{c}_3=\tilde{y}_3+R_3\tilde{b}_2-\tilde{R}_3\tilde{s}_2 1_{\{\delta^i_{el,2}=1\}},$$
$$Q_1=\lambda(s_1), \quad \tilde{Q}_t=\lambda(\tilde{s}_t), \quad \tilde{R}_t=\delta e^{\mu+\delta_\tau+\sigma z_t}, \quad z\sim N(0,1), \quad t=2,3, \quad \delta\in[0,1],$$

where $v\left(\bar{Q}\right)=\theta\bar{Q}$, y includes labor earning and a possible fine for the amount stolen in the previous term, $\tilde{\bar{Q}}$ is the adjusted 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:

$$\max_{c_1,s_1,\tilde{b}_2,\tilde{s}_2} u(y_1-b_1+s_1)+v(\bar{Q}_1)+p(s_1)\beta E\left[u\left(\tilde{y}_2-\tilde{b}_2+R_2b_1+\tilde{s}_2-\tilde{R}_2s_1\right)+v\left(\tilde{\bar{Q}}_2\right)|Inc\right]+$$
$$(1-p(s_1))\beta E\left[u\left(\tilde{y}_2-\tilde{b}_2+R_2b_1-\tilde{R}_2s_1\right)+v\left(\tilde{\bar{Q}}_2\right)|Ch\right]+$$
$$+\beta^2 E\left[u\left(\tilde{y}_3+R_3\tilde{b}_2-\tilde{R}_3\tilde{s}_2 1_{\{\delta^i_{el,2}=1\}}\right)+v\left(\tilde{\bar{Q}}_3\right)\right]-\theta\eta\left(1+\beta+\beta^2\right)\log(d).$$

s_1, b_1, $\tilde{s}_2$, and $\tilde{b}_2$ must satisfy the following Euler equations:

$$\begin{aligned}
u'(c_1)+v'\left(\bar{Q}_1\right)\lambda'(s_1)-\beta E\left[u'(\tilde{c}_2)\tilde{R}_2\right]+p'(s_1)\left(E\left[u(\tilde{c}_2)|Inc\right]-E\left[u(\tilde{c}_2)|Ch\right]\right) &<= 0\\
u'(c_1)-\beta RE\left[u'(\tilde{c}_2)\right] &= 0\\
u'(\tilde{c}_2)+v'\left(\bar{Q}_2\right)\lambda'(s_2)-\beta E\left[u'(\tilde{c}_3)\tilde{R}_3\right] &<= 0\\
u'(\tilde{c}_2)-\beta RE\left[u'(\tilde{c}_3)\right] &= 0
\end{aligned}$$

where the inequalities account for the possibility that zero stealing is optimal.

Consider first the case of a first-term mayor who chooses not to run for reelection. In this case, the relevant Euler equation are:

$$\begin{aligned} u'(c_1) + v'\left(\bar{Q}_1\right)\lambda'(s_1) - \beta E\left[u'(\tilde{c}_2)\tilde{R}_2\right] &<= 0 \\ u'(c_1) - \beta R E\left[u'(\tilde{c}_2)\right] &= 0. \end{aligned}$$

Consider an increase in δ_τ and hence in $\tilde{R}_2$ for each realization of the fine shock. Using $E[yx] = Cov(yx) + E[y]E[x]$, we have,

$$E\left[u'(\tilde{c}_2)\tilde{R}_2\right] = Cov\left[u'(\tilde{c}_2), \tilde{R}_2\right] + E\left[u'(\tilde{c}_2)\right] E\left[\tilde{R}_2\right].$$

As, $\tilde{c}_2$ declines with δ_τ and u' is a decreasing function, $u'(\tilde{c}_2)$ increases with $\tilde{c}_2$. Consequently, $Cov\left[u'(\tilde{c}_2), \tilde{R}_2\right]$ increases with $\tilde{c}_2$. A a first order Taylor expansion of $u'(\tilde{c}_2)$ around $\bar{c}_2 = E[\tilde{c}_2] = E[\tilde{y}_2] + R_2\tilde{b}_2 - E[\tilde{R}_2]\tilde{s}_2 1_{\{\delta^i_{el,2}=1\}}$ makes this clear. Let $u'(\tilde{c}_2) = u'\left(\tilde{x} - z\tilde{R}_2\right)$ and $u'(\bar{c}_2) = u'\left(\bar{x} - z\bar{R}_2\right)$. Then

$$u'(\tilde{c}_2) \approx u'(\bar{c}_2) + u''(\bar{c}_2)\left(\tilde{x} - z\tilde{R}_2 - (\bar{x} - z\bar{R}_2\right).$$

Hence,

$$Cov\left[u'(\tilde{c}_2), \tilde{R}_2\right] = -u''(\bar{c}_2)\, zV\left(\tilde{R}_2\right).$$

An increase in δ_τ reduces $\bar{c}_2$, reduces $u''(\bar{c}_2)$ and, hence, increases $Cov\left[u'(\tilde{c}_2), \tilde{R}_2\right]$.

An increase in δ_τ also clearly increases $E\left[\tilde{R}_2\right]$. Hence, $E\left[u'(\tilde{c}_2)\tilde{R}_2\right]$ increases and, at the old optimum,

$$u'(c_1) - \beta R E\left[u'(\tilde{c}_2)\right] > 0.$$

As a result, at the new optimum c_1 must decline.

The reduction in c_1 can be achieved by increasing savings b_1, by reducing stealing s_1, or by change b_1 and s_1 in a way that reduces $s_1 - b_1$. Suppose that the optimal choice is to increase savings and not to reduce stealing. Without loss of generality, we consider the borderline case in which b_1 increases and s_1 does not change.

Let c_1^*, b_1^*, s_1^*, b_2^*, and s_2^* be the initial optimum. Also, at the new optimum, $b_1^{**} = b_1^* + \delta_1$, $s_1^{**} = s_1^*$, $b_2^{**} = b_2^* + \delta_2$, and $s_2^{**} = s_2^* + \xi_2$. Since, $\delta_\tau > 1$ (the amount stolen must be returned), we can write $\tilde{R}_2' = \tilde{R}_2 + \Delta$. Then,

$$c_1^{**} = y_1 - b_1^* - \delta_1 + s_1^* \quad \text{and} \quad \tilde{c}_2^{**} = \tilde{y}_2 + R_2\left(b_1^* + \delta_1\right) - \left(\tilde{R}_2 + \Delta\right) s_1^* - \left(b_2^* + \delta_2\right) + s_2^* + \xi_2.$$

Consider the alternative optimal solution $b_1^{**} = b_1^* + \delta_1 - \epsilon$, $s_1^{**} = s_1^* - \epsilon$, with $\epsilon > 0$, and b_2^{**} and s_2^{**} don't change. In this case, c_1^{**} stays the same and

$$\tilde{c}_2^{**'} = \tilde{y}_2 + R_2\left(b_1^* + \delta_1\right) - \left(\tilde{R}_2 + \Delta\right) s_1^* - \left(b_2^* + \delta_2\right) + s_2^* + \xi_2 + \left(\tilde{R}_2 + \Delta - R_2\right)\epsilon.$$

To evaluate the expected utility generated by $\tilde{c}_2^{**'}$, we use a first order Taylor expansion of $u\left(\tilde{c}_2\right)$ around the expected value of the initial optimum $\bar{c}_2^{**}$, to obtain

$$\begin{aligned}
&E\left[u\left(\tilde{c}_2^{**'}\right)\right] \approx u\left(\bar{c}_2^{**}\right) + u'\left(\bar{c}_2^{**}\right)\left(E\left[\tilde{y}_2\right] + R_2\left(b_1^* + \delta_1\right) - \left(E\left[\tilde{R}_2 + \Delta\right]\right) s_1^*\right. \\
&\left.- \left(b_2^* + \delta_2\right) + s_2^* + \xi_2 + \left(E\left[\tilde{R}_2 + \Delta\right] - R_2\right)\epsilon - \bar{c}_2^{**}\right) = \\
&u\left(\bar{c}_2^{**}\right) + u'\left(\bar{c}_2^{**}\right)\left(E\left[\tilde{R}_2 + \Delta\right] - R_2\right)\epsilon.
\end{aligned}$$

The alternative policy has also an effect on the utility generated by public consumption $v\left(Q_1\right)$, as optimal stealing declines. To evaluate its effect, we take a first order Taylor expansion of $v\left(Q\left(s_1^{**'}\right)\right)$ around the initial optimal stealing s_1^{**}:

$$v\left(\bar{Q}\left(s_1^{**'}\right)\right) \approx v\left(\bar{Q}\left(s_1^{**}\right)\right) + v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)\left(s_1^{**'} - s_1^{**}\right) = v\left(\bar{Q}\left(s_1^{**}\right)\right) - v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)\epsilon.$$

The total change in utility generated by the alternative choice is therefore

$$G^{**} = u'\left(\bar{c}_2^{**}\right)\left(E\left[\tilde{R}_2 + \Delta\right] - R_2\right)\epsilon - v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)\epsilon.$$

For the solution b_1^{**}, s_1^{**}, and b_2^{**} after the increase in δ_τ to be optimal, it must be $G^{**} \leq 0$, or the alternative solution would increase total utility. Equivalently, it must be

$$E\left[\tilde{R}_2 + \Delta\right] - R_2 \leq \frac{v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)}{u'\left(\bar{c}_2^{**}\right)}.$$

It must therefore also be that

$$E\left[\tilde{R}_2\right] - R_2 < \frac{v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)}{u'\left(\bar{c}_2^{**}\right)}. \tag{4}$$

Let's go back to the initial optimal decision c_1^*, b_1^*, s_1^*, b_2^*, and s_2^*. Consider the alternative with identical c_1^*, b_2^*, , and s_2^*, but $b_1^{*'} = b_1^* + \epsilon$, $s_1^{*'} = s_1^* + \epsilon$, with $\epsilon > 0$. Then,

$$\tilde{c}_2^{*'} = \tilde{y}_2 + R_2 b_1^* - \tilde{R}_2 s_1^* - b_2^* + s_2^* - \left(\tilde{R}_2 - R_2\right)\epsilon.$$

The corresponding expected utility is given by

$$\begin{aligned} E\left[u\left(\tilde{c}_2^{*'}\right)\right] \approx\ & u\left(\bar{c}_2^*\right) + u'\left(\bar{c}_2^*\right)\left(E\left[\tilde{y}_2\right] + R_2 b_1^* - E\left[\tilde{R}_2\right] s_1^* - b_2^* + s_2^* - \left(E\left[\tilde{R}_2\right] - R_2\right)\epsilon - \bar{c}_2^{*'}\right) = \\ & u\left(\bar{c}_2^*\right) - u'\left(\bar{c}_2^*\right)\left(E\left[\tilde{R}_2\right] - R_2\right)\epsilon. \end{aligned}$$

The change in stealing generates the following change in the utility for public consumption:

$$v\left(\bar{Q}\left(s_1^{*'}\right)\right) \approx v\left(\bar{Q}\left(s_1^*\right)\right) + v'\left(\bar{Q}\left(s_1^*\right)\right)\lambda'\left(s_1^*\right)\left(s_1^{*'} - s_1^*\right) = v\left(\bar{Q}\left(s_1^*\right)\right) + v'\left(\bar{Q}\left(s_1^*\right)\right)\lambda'\left(s_1^*\right)\epsilon.$$

The total change in utility generated by the alternative choice is therefore

$$G^* = -u'\left(\bar{c}_2^*\right)\left(E\left[\tilde{R}_2\right] - R_2\right)\epsilon + v'\left(\bar{Q}\left(s_1^*\right)\right)\lambda'\left(s_1^*\right)\epsilon.$$

For the solution b_1^*, s_1^*, b_2^*, and s_2^* before the increase in δ_τ to be optimal, it must be $G^* \leq 0$, or the alternative solution would increase total utility. Hence, we must have

$$E\left[\tilde{R}_2\right] - R_2 > \frac{v'\left(\bar{Q}\left(s_1^*\right)\right)\lambda'\left(s_1^*\right)}{u'\left(\bar{c}_2^*\right)} = \frac{v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)}{u'\left(\bar{c}_2^*\right)}.$$

where the equality follows from the initial statement $s_1^{**} = s_1^*$, i.e. stealing does not change after the increase in δ_τ.

The last thing to prove is that $u'\left(\bar{c}_2^*\right) < u'\left(\bar{c}_2^{**}\right)$. From the previous result that, with a higher

δ_τ, c_1 must increase, we have

$$u'(\bar{c}_2^*) \approx \beta R E\left[u'(\tilde{c}_2^*)\right] = u'(c_1^*) < u'(c_1^{**}) = \beta R E\left[u'(\tilde{c}_2^{**})\right] \approx u'(\bar{c}_2^{**}).$$

Hence,

$$E\left[\tilde{R}_2\right] - R_2 > \frac{v'\left(\bar{Q}\left(s_1^*\right)\right)\lambda'\left(s_1^*\right)}{u'\left(\bar{c}_2^*\right)} = \frac{v'\left(\bar{Q}\left(s_1^{**}\right)\right)\lambda'\left(s_1^{**}\right)}{u'\left(\bar{c}_2^*\right)} > \frac{v'\left(\bar{Q}\left(s_1^*\right)\right)\lambda'\left(s_1^*\right)}{u'\left(\bar{c}_2^{**}\right)}.$$

This contradict the previous inequality (4) required for the new solution with $s_1^{**} = s_1^*$ to be optimal. We can therefore conclude that stealing must decline in the new optimal choice with higher δ_τ.

A.3 Proof of Proposition 3

The probability that a mayor wins reelection is given by

$$\Phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\middle|\, \bar{\bar{Q}}, \bar{s}\right] - E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\right] + \delta_1 + \delta_2\bar{Q}_t + \delta_3 1_{\{\delta_{el,t}^i=1, s_t^i=0\}} - \delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\right).$$

Its derivative with respect to δ_1 is therefore

$$\phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\middle|\, \bar{\bar{Q}}, \bar{s}\right] - E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\right] + \delta_1 + \delta_2\bar{Q}_t + \delta_3 1_{\{\delta_{el,t}^i=1, s_t^i=0\}} - \delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\right) >$$

where ϕ is the standardized normal density function. This probability is therefore strictly increasing in δ_1.

A.4 Proof of Proposition 4

The probability that a mayor wins reelection conditional on producing more than $\bar{Q}_{y-th}$ is given by

$$\begin{aligned}\Phi\Big(&E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\middle|\, \bar{\bar{Q}}, \bar{s}, \bar{Q} > \bar{Q}_{y-th}\right] - E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM}, t+1\right)\right] + \\ &+ \delta_1 + \delta_2\bar{Q}_t + \delta_3 1_{\{\delta_{el,t}^i=1, s_t^i=0\}} - \delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\,\Big|\, \bar{Q} > \bar{Q}_{y-th}\Big).\end{aligned}$$

Its derivative with respect to δ_2 is therefore

$$\phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\middle|\,\bar{\bar{Q}},\bar{s},\bar{Q}>\bar{Q}_{y-th}\right]-E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\right]+\right.$$
$$\left.+\,\delta_1+\delta_2\bar{Q}_t+\delta_3 1_{\{\delta_{el,t}^i=1,s_t^i=0\}}-\delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\middle|\,\bar{Q}>\bar{Q}_{y-th}\right)\bar{Q}_t>0,$$

which concludes the proof.

A.5 Proof of Proposition 5

The reelection probability of a mayor who was audited and found to not have stolen stole is given by

$$\Phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\middle|\,\bar{\bar{Q}},\bar{s},s_t^i>0,1_{\{\delta_{el,t-1}^i=1\}}=1\right]-E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\right]+\right.$$
$$\left.+\,\delta_1+\delta_2\bar{Q}_t+\delta_3 1_{\{\delta_{el,t}^i=1,s_t^i=0\}}-\delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\middle|\,s_t^i>0,1_{\{\delta_{el,t}^i=1\}=1}\right).$$

Its derivative with respect to δ_3 is therefore

$$-\phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\middle|\,\bar{\bar{Q}},\bar{s},s_t^i>0,1_{\{\delta_{el,t-1}^i=1\}}=1\right]-E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\right]+\right.$$
$$\left.+\,\delta_1+\delta_2\bar{Q}_t+\delta_3 1_{\{\delta_{el,t}^i=1,s_t^i=0\}}-\delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\middle|\,s_t^i>0,1_{\{\delta_{el,t}^i=1\}=1}\right)1_{\{\delta_{el,t}^i=1\}}>0.$$

This concludes the proof.

A.6 Proof of Proposition 6

The reelection probability of a mayor who was audited and stole is given by

$$\Phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\middle|\,\bar{\bar{Q}},\bar{s},s_t^i>0,1_{\{\delta_{el,t-1}^i=1\}}=1\right]-E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\right]+\right.$$
$$\left.+\,\delta_1+\delta_2\bar{Q}_t+\delta_3 1_{\{\delta_{el,t}^i=1,s_t^i=0\}}-\delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\middle|\,s_t^i>0,1_{\{\delta_{el,t}^i=1\}=1}\right).$$

Its derivative with respect to δ_4 is therefore

$$-\phi\left(E_t^{Inc}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\middle|\,\bar{\bar{Q}},\bar{s},s_t^i>0,1_{\{\delta_{el,t-1}^i=1\}}=1\right]-E_t^{Ch}\left[V_{PM}^i\left(S_{t+1}^{PM},t+1\right)\right]+\right.$$
$$\left.+\,\delta_1+\delta_2\bar{Q}_t+\delta_3 1_{\{\delta_{el,t}^i=1,s_t^i=0\}}-\delta_4 s_t^i 1_{\{\delta_{el,t}^i=1\}}\middle|\,s_t^i>0,1_{\{\delta_{el,t}^i=1\}=1}\right)s_t^i 1_{\{\delta_{el,t}^i=1\}}<0.$$

This concludes the proof.

A.7 Proof of Proposition 7

We focus on second term mayors because the notation is less cumbersome. But the same argument can be applied to first term mayors.

In proof of Proposition 2, we have shown that, for second-term mayors stealing is determined by the following Euler equation:

$$u'\left(\tilde{c}_2\right)\leq\beta E\left[u'\left(\tilde{c}_3\right)\tilde{R}_3\right]-v'\left(\bar{Q}_2\right)\lambda'\left(s_2\right),$$

where the inequality accounts for zero stealing. It the mayor steals a positive amount of funds the Euler equation becomes a strict inequality:

$$u'\left(\tilde{c}_2\right)=\beta E\left[u'\left(\tilde{c}_3\right)\tilde{R}_3\right]-v'\left(\bar{Q}_2\right)\lambda'\left(s_2\right),$$

The rivalry parameter η affects only the term $v'\left(\bar{Q}_2\right)$ by reducing it and, hence, by reducing $-v'\left(\bar{Q}_2\right)\lambda'\left(s_2\right)$, as $\lambda'\left(s_2\right)<0$. Specifically, given that in our model $\frac{\partial v\left(\bar{Q}_2\right)}{\partial Q}=\frac{\theta}{d^\eta}$, we have

$$\frac{\partial}{\partial\eta}\left(\frac{\partial v\left(\bar{Q}_2\right)}{\partial Q}\right)=-\frac{\theta\ln d}{d^\eta}<0.$$

This implies that, if η increases without changing the optimal choice, the previous Euler equation is replaced by the following inequality:

$$u'\left(\tilde{c}_2\right)>\beta E\left[u'\left(\tilde{c}_3\right)\tilde{R}_3\right]-v'\left(\bar{Q}_2\right)\lambda'\left(s_2\right).$$

The get back the equality, stealing must increase so that the left hand side declines by an increase in $\tilde{c}_2$ and the right hand side, which measures the cost of stealing, increases. Consequently, if for a fraction of mayors it is optimal to steal a positive amount, which is the case in our data, average stealing among mayors is an increasing function of η. This proves the first part of the Proposition.

To prove the second part, note that the reduction in $-\frac{\partial v(\bar{Q}_2)}{\partial Q}\lambda'(s_2)$ is increasing in population size d:

$$\frac{\partial}{\partial d}\left(\frac{\partial}{\partial \eta}\left(\frac{\partial v\left(\bar{Q}_2\right)}{\partial Q}\right)\right) = \frac{\theta}{d^\eta}\left(\frac{d\left(\ln d\right)^2 - 1}{d}\right) > 0.$$

Consequently, stealing has to increase more with η if population size is larger for the Euler equation to be satisfied. Thus, if for a fraction of mayors it is optimal to steal a positive amount, average stealing among mayors increases faster with η if population size is larger. This concludes the proof.

A.8 Proof of Proposition 8

Remember that S is the vector of state variables that affect the decisions of an incumbent, that $E\left[V_{RM}\left|S\right.\right]$ and $E\left[V_{NRM}\left|S\right.\right]$ 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

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

where $\epsilon_R \sim N\left(\mu_R, \sigma_R\right)$ is a shock to the decision to run. 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:

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

Since the utility from being in power ρ and the cost of running κ 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:

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

and

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

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

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

In the model, $E\left[\bar{V}_{LM}\left|S\right.\right] = E\left[V_{NRM}\left|S\right.\right]$ as the only difference between an incumbent who runs and loses and an incumbent who chooses to run is the cost κ. Consequently, the inequality simplifies to

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

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

$$\begin{aligned}
P\left[R\left|S\right.\right] &= P\left[P\left(S\right)\left(E\left[\bar{V}_{WM}\left|S\right.\right] - E\left[\bar{V}_{LM}\left|S\right.\right]\right) - \kappa + P\left(S\right)\rho \geq -\epsilon_R\right] = \\
&= P\left[\frac{P\left(S\right)\left(E\left[\bar{V}_{WM}\left|S\right.\right] - E\left[\bar{V}_{LM}\left|S\right.\right]\right) + P\left(S\right)\rho - \kappa + \mu_R}{\sigma_R} \geq \frac{-\epsilon_R + \mu_R}{\sigma_R}\right] = \\
&= \Phi\left(\frac{P\left(S\right)\left(E\left[\bar{V}_{WM}\left|S\right.\right] - E\left[\bar{V}_{LM}\left|S\right.\right]\right) + P\left(S\right)\rho - \kappa + \mu_R}{\sigma_R}\right),
\end{aligned}$$

where Φ is the cumulative density function of a standard normal. The previous steps make clear that we can only identify the sum of the cost of running κ and mean of the running shock μ_R. We therefore normalize κ to be equal zero. This is without loss of generality as in our policy simulation we do not need to know κ separately from μ_R. Then, we have,

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

By taking the expectation over S, we then have

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

In the previous equation, the quantities $K_1(S) = E_S[\Phi^{-1}(P[R|S])]$, $K_2(S) = E_S[P(S)|S]$, and $K_3(S) = E\left[P(S)\left(E\left[\bar{V}_{WM}|S\right] - E\left[\bar{V}_{LM}|S\right]\right)\right]$ do not depend on ρ, μ_R, and σ_R. Thus, for any given set of the other model parameters, they are known if the probabilities of running and winning are known. The previous equation can therefore be written as a linear function of the three parameters of interest:

$$K_1(S)\,\sigma_R - K_2(S)\,\rho + \mu_R = K_3(S)\,.$$

Consider now three groups of mayors: the group composed of audited mayors that did not steal, with state variables S^N; the group formed by audited mayors who chose to steal, with S^S, and the group that includes all the mayors unconditional of the audit outcome, with S^U. These three groups of mayors have different sets of state variables and, hence, different quantities $K_1(S)$, $K_2(S)$, and $K_3(S)$. We therefore have three linear equations in the three parameters we are interested in:

$$\begin{aligned} K_1\left(S^N\right)\sigma_R - K_2\left(S^N\right)\rho + \mu_R &= K_3\left(S^N\right) \\ K_1\left(S^S\right)\sigma_R - K_2\left(S^S\right)\rho + \mu_R &= K_3\left(S^S\right) \\ K_1\left(S^U\right)\sigma_R - K_2\left(S^U\right)\rho + \mu_R &= K_3\left(S^U\right). \end{aligned}$$

The parameters ρ, μ_R, and σ_R are therefore identified if the probabilities of running and winning are known separately for S^N, S^S, and S^U.

A.9 Proof of Proposition 9

As the measurement error on stealing is independent of the actual fraction stolen, the variance of the observed fraction of funds diverted is given by

$$V\left(\frac{s_t\nu_t}{f_t^{pu}}\right) = V\left(\frac{s_t}{f_t^{pu}}\right)V(\nu_t) + V\left(\frac{s_t}{f_t^{pu}}\right)E(\nu_t)^2 + V(\nu_t)\,E\left(\frac{s_t}{f_t^{pu}}\right)^2 = \left[V\left(\frac{s_t}{f_t^{pu}}\right) + E\left(\frac{s_t}{f_t^{pu}}\right)^2\right]V(\nu_t)\,,$$

where the second equality follows from $E(\nu_t) = 0$. Since $\left[V\left(\frac{s_t}{f_t^{pu}}\right) + E\left(\frac{s_t}{f_t^{pu}}\right)^2\right]$ does not depend on σ_ν, for any value of the other parameters, $V\left(\frac{s_t\nu_t}{f_t^{pu}}\right)$ is increasing in $\sigma_\nu = \sqrt{V(\nu_t)}$.

B 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.

C Appendix Tables

Table 8: GDP Per Capita and Public Goods

Variables	GDP per capita	HDI	Water	Sewage	Electricity	Schools	Computers	Health Clinics
GDP per capita	1.00							
HDI	0.47	1.00						
	(0.00)							
Water	0.40	0.86	1.00					
	(0.00)	(0.00)						
Sewage	0.29	0.61	0.66	1.00				
	(0.00)	(0.00)	(0.00)					
Electricity	0.31	0.73	0.76	0.58	1.00			
	(0.00)	(0.00)	(0.00)	(0.00)				
Schools	-0.02	-0.08	-0.14	-0.08	-0.14	1.00		
	(0.16)	(0.00)	(0.00)	(0.00)	(0.00)			
Computers	0.28	0.48	0.41	0.29	0.32	0.03	1.00	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)		
Health clinics	0.10	0.15	0.07	0.07	0.09	0.74	0.18	1.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	

Notes: This table presents a correlation matrix between GDP per capita (measured in 2001) and various indicators of public goods. HDI is the UN's Human Development Index computed by the IBGE for Brazil's municipalities (measured in 2000). Water, Sewage, Electricity indicate the share of households who have access to running water, sewage, electricity, respectively. These data were computed from the 2000 Census. Schools refers to the number of schools in the municipality (measured in 2001). Computers refers to share of schools with a computer lab (measured in 2001). Health Clinics refers the number of health establishments in the municipality (measured in 2005).

Table 9: The Effects of the Audits on the Candidate Pool

	Number of Candidates	Share Male	Share College	Average Age	Campaign Spending
	(1)	(2)	(3)	(4)	(5)
Panel A: 2004 elections					
Audited	-0.065	0.012	0.007	0.297	-19.962
	(0.044)	(0.008)	(0.014)	(0.270)	(24.023)
Number of observations	5476	5476	5476	5476	5476
R^2	0.27	0.03	0.15	0.05	0.02
Panel B: 2004-2012 elections					
Audited	-0.00123	-0.000403	-0.00379	0.0620	7.402
	(0.04)	(0.07)	(0.42)	(0.33)	(0.86)
Observations	16432	16432	16432	16432	16432
R^2	0.57	0.52	0.62	0.54	0.34

Notes: This table presents regression output from two separate regressions. In Panel A, we regress the dependent variable indicated by the columns on an indicator for whether the municipality had been audited prior to the 2004 elections. The regression also controls for log population 2000 and state fixed effects. The unit of observation is a municipality and robust standard errors are reported in parentheses. In panel B, we estimate a regression pooling the 2004, 2008, 2012 elections. The dependent variable is denoted in column and the main independent variable is an indicator for having been audited prior to that election. The regression also includes election year fixed effects and municipal fixed effects. The unit of observation is a municipality, election year. Robust standard errors clustered at the municipality level are reported in parentheses.

Table 10: Simulations for Baseline Model - Policy Robustness

		Average Stealing			
	WTP	Full Sample	1^{st} mayors	2^{nd} mayors	3^{rd} mayors
	(1)	(2)	(3)	(4)	(5)
25% Increase in ability					
Audit	0.428	0.057	0.051	0.067	0.000
CRA	1.014	0.060	0.038	0.101	0.000
3-term	0.489	0.078	0.081	0.072	0.054
Wages	0.524	0.078	0.079	0.077	0.000
3-term+Audit	0.702	0.061	0.062	0.059	0.049
CRA+Audit	1.136	0.045	0.021	0.090	0.000
3-term+CRA+Audit	1.315	0.016	0.021	0.008	0.074
Wage+Audit	0.676	0.067	0.066	0.069	0.000
25% Decrease in ability					
Audit	0.350	0.070	0.061	0.087	0.000
CRA	0.975	0.074	0.045	0.127	0.000
3-term	0.457	0.093	0.094	0.091	0.072
Wages	0.417	0.098	0.093	0.105	0.000
3-term+Audit	0.620	0.073	0.073	0.074	0.064
CRA+Audit	1.083	0.057	0.026	0.112	0.000
3-term+CRA+Audit	1.253	0.020	0.026	0.011	0.092
Wage+Audit	0.549	0.084	0.079	0.094	0.000

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.

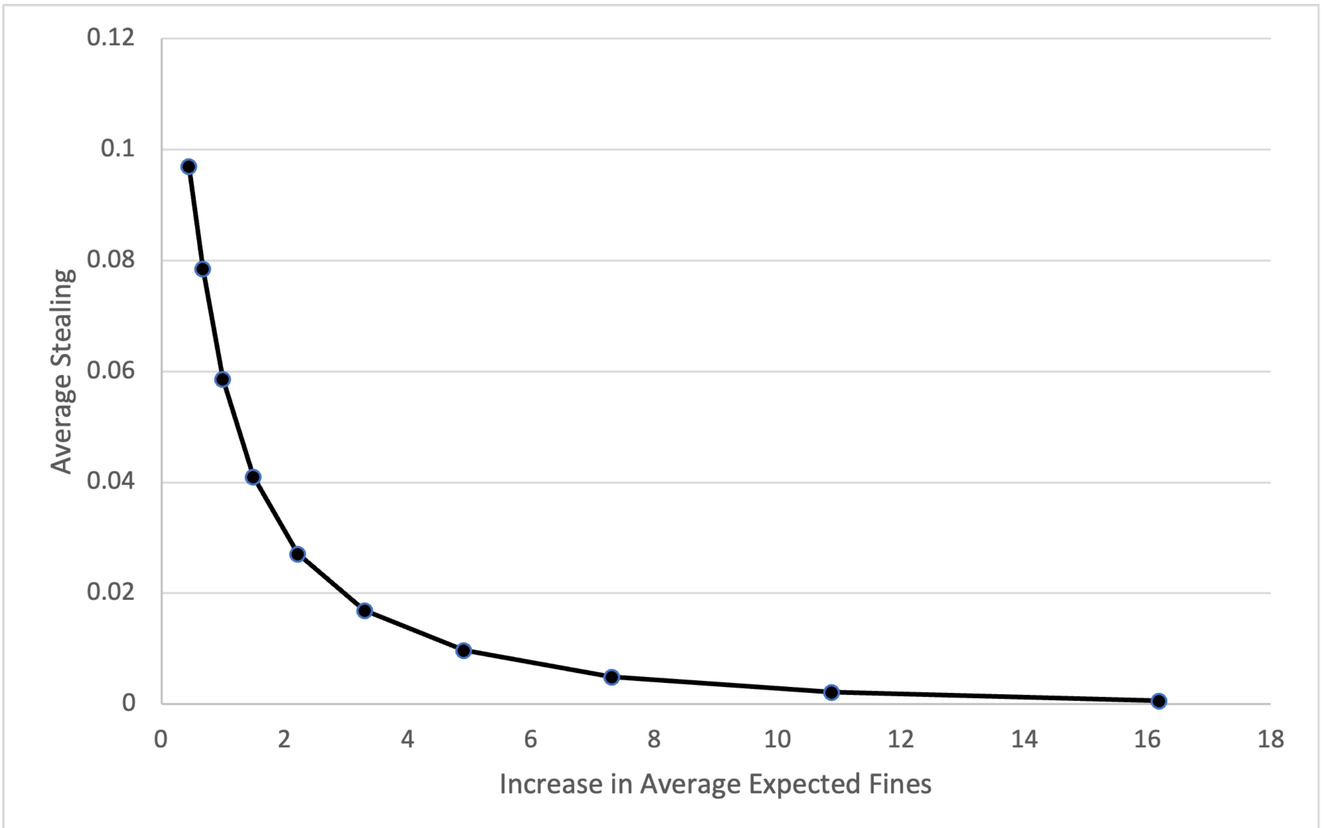


Figure 4: The Effects of Fines on Stealing

Notes: This figure presents results based on 100 simulations for each municipality.

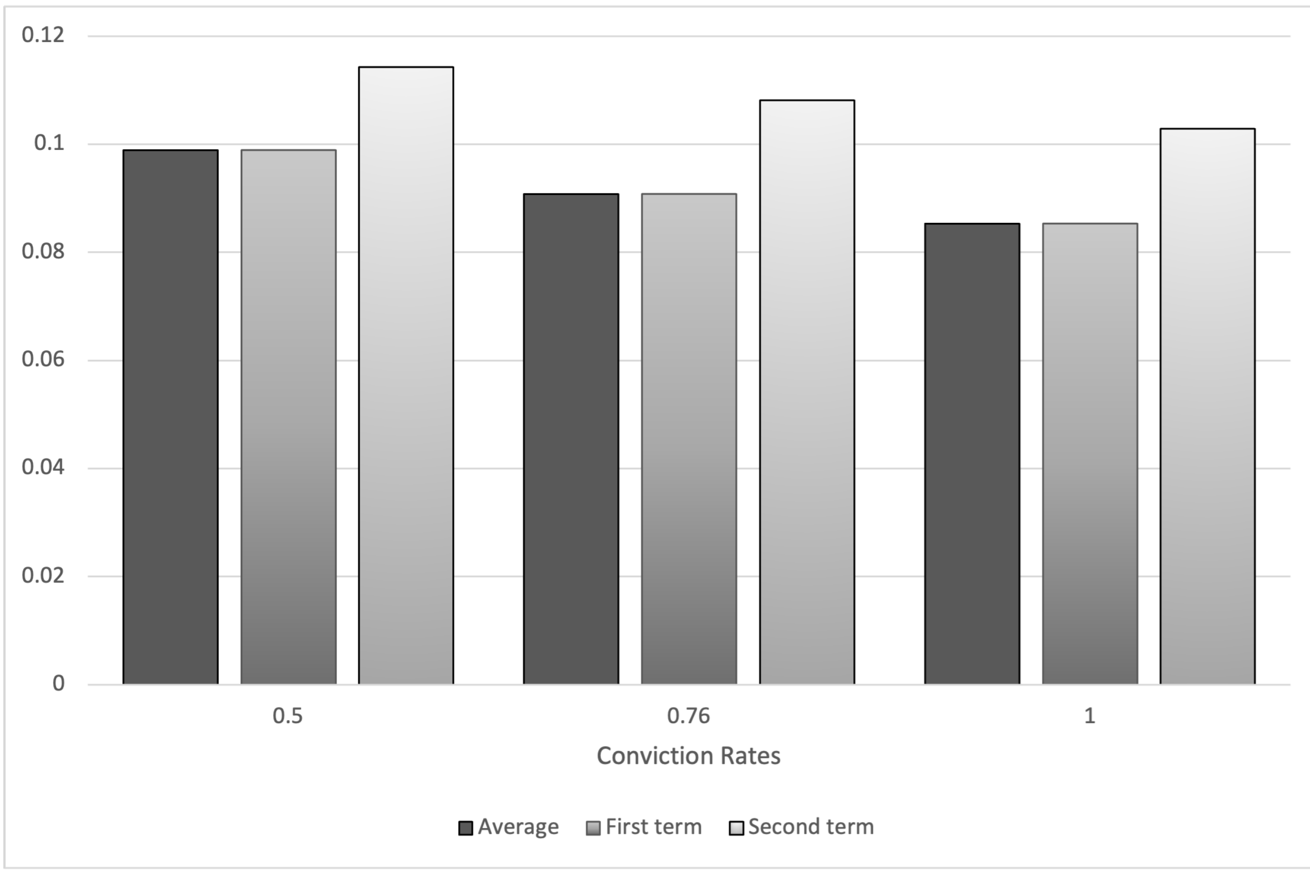


Figure 5: The Effects of Conviction Rates on Average Stealing

Notes: This figure presents results based on 100 simulations for each municipality.

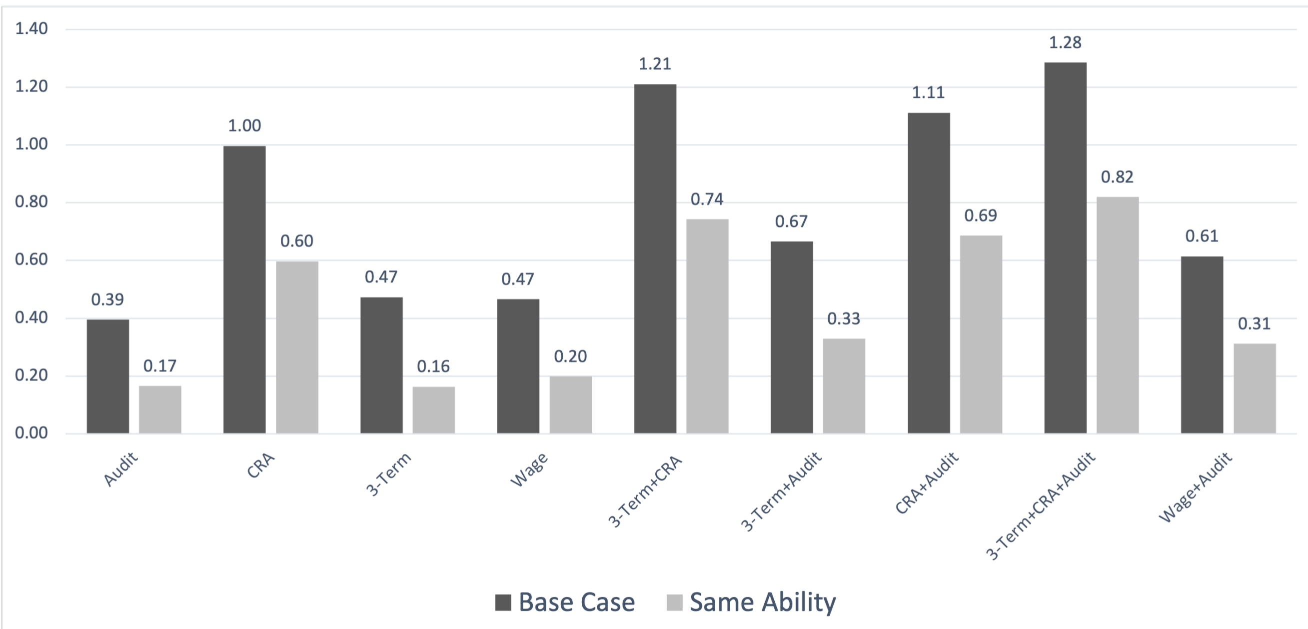


Figure 6: Willingness to Pay: The Effects of Ability

Notes: This figure presents results based on 100 simulations for each municipality.