Combating Political Corruption with Policy Bundles*

Frederico Finan† UC-Berkeley

Maurizio Mazzocco‡ UCLA

September 2024

Abstract

In this paper, we develop a dynamic model of politicians who can engage in corruption. The model provides important insights into the determinants of corruption and how to design policies 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, raising 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 extremely grateful for the time and effort the referees, and especially the editor, dedicated to our paper. Their comments and suggestions improved the paper significantly. We also thank 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 effective policies to reduce 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 policies evaluated during different periods and/or settings. For example, audit policies can effectively reduce corruption in Brazil, China, and Indonesia, but how do they compare to extending term limits in Mexico? Second, each policy has its strengths and weaknesses. Perhaps we can achieve better results by combining policies to minimize each one'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 likely affect not only their current choices but also their future ones. It is difficult to capture these future adjustments in reduced-form analysis.

To address these gaps in the literature, we need to better understand 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 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. 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, 8.6 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, in 1997 Brazil allowed 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 candidates who ran for mayor, 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 a sizable part of the difference in stealing between first and second-term mayors. Second, the comparison also captures selection effects created by elections. We find that second-term mayors are, on average, more

[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.

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 policies 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 mayors' 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 36 percent. For comparison, doubling mayors' wages is about a third as effective, with a 13% reduction. The least effective policy at reducing corruption is to extend the number of consecutive terms a mayor can hold office from two to three terms. 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 0.6% 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 15%, which interestingly is not too far from Brazil's current policy. But even at the optimal audit probability, voters would 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 can have limited effects on mayors who are in their last term or have electoral incentives absent the reform. The Clean Record Act has the weakness of only affecting politicians who plan to run for reelection. Given these limitations, we can increase the efficacy of the two policies by just combining them.

A ban on corrupt politicians from running for office is more effective because it subjects them to electoral incentives for an 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 34%. 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 9%, and this policy proves to be our most effective. Citizens are willing to pay 1.3% of their income for this policy, and it reduces corruption by more than 56%.

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 Mexico that allowed mayors, who had been limited to a single term, to run for reelection for an

[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.

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. However, 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.

Municipal Politics and Corruption. Local elections in Brazil occur every four years, allowing citizens to vote for a municipal mayor and a local council. Following the election, mayors are subject to a two-term limit. Given their pivotal role, mayors wield significant influence as they oversee the allocation of millions of dollars in federal and state funds, crucial for providing basic public services like education, healthcare, and sanitation.

In Brazil, combating political corruption is a paramount concern, both nationally and locally. This concern has spurred two noteworthy policy reforms that shape our subsequent analysis. First, during our sample period, the federal government in 2003 implemented an audit program significantly increasing the likelihood of audits for federal transfers, which constitute the bulk of municipal funds. Before this program, federal oversight accompanying such transfers was minimal, with a mere five percent effective audit probability; after the introduction of the program, the effective audit probability stood at 16.8 percent. Second, another pivotal reform came in 2010, after our sample period, in the form of the Clean Records Act. This legislation mandates an eight-year ineligibility from holding office for anyone convicted of a crime, even if appeals are pending.

Public Funds and the Audit Program. 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 450,000 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 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. To start, 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 further reduce the opportunity for corruption among individual auditors.[6]

Our data on corruption comes from the audit reports generated by this anti-corruption program. Specifically, 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 au-

[4] This population eligibility criterion has changed slightly over time after our sample period.

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

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

dited.[7] These data 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.[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 27% of audited mayors were not found to be corrupt.

As originally documented by Ferraz and Finan (2011), term limits affect corruption levels. In Brazil, mayors can only serve two consecutive terms. We find that mayors who were in their first term steal on average 5.6 percent of the allocated funds, whereas second-term mayors divert 7.3 percent, a 30 percent increase. In column 1 of Table 2, we show that this difference is also robust to controlling for various mayor and municipal characteristics.[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, public work costs 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

[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.

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

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. 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 9, we show that GDP per capita strongly correlates with various public goods indicators, including the share 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 impact on total public consumption than 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 con-

[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.

sumption. We regress the log of per-capita public consumption on the log of federal funds, the log amount of funds diverted, an indicator 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 that is 15 percentage points lower than

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

those who were audited but not found to be corrupt. This finding is consistent with the results originally found in Ferraz and Finan (2008), as well as with those found in other settings: 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 an 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.

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 from 2005 to 2013. We then compute their average wage in the private sector over this period.[12] Overall, 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, an indicator for being audited, and an indicator 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 indicators 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

[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.

variable, this information is publicly available on most municipalities' 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 were equal to R$3,233 for municipalities with a population of fewer 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 and less than 450,000.

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 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 earning lower future wages.

3 Model

In this section, we develop a 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 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 goods they receive (e.g. schools, police force, parks, and roads). Local governments produce these public goods using public funds. 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.[14] Individuals have preferences over a private good c and a public good 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 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 for individual i with the utility function $u^i\left(c_t^i, \bar{Q}_t\right)$.

If the current mayor decides to run for reelection, they must pay a utility cost κ. If elected, they derive a utility ρ from being in power. We can characterize these individuals' preferences with the following utility function:

$$u^i\left(c_t^i, \bar{Q}_t\right) + \rho - \kappa. \tag{1}$$

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.[15]

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

[14] We do not model the decision to enter into politics and abstract from selection issues associated with this decision.

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

are heterogeneous in their ability a_i to produce the public good. A mayor's ability is not observed by voters and 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:

$$\frac{Q_t}{d_t} = \left(\frac{z_t^{pu}}{d_t}\right)^{\alpha_1} \left(z_t^{pr}\right)^{\alpha_2} \ a_i. \tag{2}$$

By specifying the production function in per-capita terms for public consumption, 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.[16]

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 goods. They also decide how much to consume c_t versus save b_t of their resources, including any money they have stolen. Finally, provided they are not in their last term, mayors choose whether to run for reelection based on their expected lifetime utility conditional on running versus not, and a running shock $\epsilon_t^R \sim N(\mu_R, \sigma_R)$. We denote this decision with the indicator $\mathbb{1}_t^{run} = 1$ if the incumbent decides to run for an additional term at time t.

The central government audits municipalities at random with probability p_t^{au}. If a municipality is audited, which we will denote by $\mathbb{1}_t^{au} = 1$, and public resources have been diverted, the mayor is caught with probability one and the amount stolen becomes public knowledge. To account for the possibility that not all mayors accused of corruption are convicted, a mayor who has been caught stealing is convicted with probability p_t^c. We let the indicator $\mathbb{1}_t^c = 1$ denote a mayor who has been convicted. 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

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

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 standard deviation σ_τ. The fine schedule represents one of the main potential deterrents of corruption. 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 capture these delays by drawing fines from a distribution whose mean can be lower than what is estimated in the data. Specifically, we introduce in 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. Consequently, 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 omit this consideration. We also assume that the mayor's unobserved ability a_i, is a politician-specific skill that does not affect market wages.[17] 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 indicators to allow for municipal heterogeneity in wages. Specifically, we let

$$\ln w_t^{fm} = \gamma_0 + \gamma_1 e_t + \gamma_2 age_t + \gamma_3 age_t^2 + \gamma_4 pop_m + \gamma_5 pop_l + \epsilon_t^w, \tag{3}$$

where pop_m is an indicator equal to one if the population in the municipality is between 10,000 and 50,000, pop_l is equal to one if the municipality's population is above 50,000, and $\epsilon_t^w \sim N(0, \sigma_{fm})$ is an iid shock.

Individuals can save or borrow an amount b at an interest rate R. A mayor's wealth affects

[17] 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.

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 either the incumbent being elected or a challenger being elected.[18] 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, which includes the politician's unobserved ability. Because voters can adjust their expectation based on the incumbent's past decisions while in office, we also condition the expected lifetime utility on the amount of public consumption provided by the incumbent in the previous term, whether the municipality was audited and, if audited, the amount stolen. We use the same approach for challengers but without conditioning on past decisions, since challengers were not in office. The expected voter's lifetime utility also includes an election shock, $\varepsilon_{i,t} \sim N(0,1)$, that affects voters at the time of the election. The exact form of the voters' expected conditional lifetime utilities will be presented after we introduce the recursive formulation of their decisions.

Voters' decisions introduce 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 also generate selection on ability since more able mayors produce more public good, all else equal.

[18]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.

3.3 The Individual Decision Process

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

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

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

$$c_t^i+b_t^i=w_t^i\bar{h}+\mathbb{1}_t^{el}s_t^i+R_tb_{t-1}^i-\mathbb{1}_{t-1}^{au}\mathbb{1}_{t-1}^{c}g\left(s_{t-1}^i\right)\qquad\text{for each } t \text{ and } \omega, \tag{4}$$

where $\mathbb{1}_t^{run}$, $\mathbb{1}_t^{el}$, $\mathbb{1}_t^{au}$, and $\mathbb{1}_t^{c}$ are indicators for running for reelection, being elected, audited, and convicted at t, respectively.

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:

$$z_t^{pu}+s_t^i=f_t^{pu}\qquad\text{for each } t \text{ and } \omega.^{19} \tag{5}$$

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

$$\frac{Q_t}{d_t}=\left(\frac{z_t^{pu}}{d_t}\right)^{\alpha_1}\left(z_t^{pr}\right)^{\alpha_2}a_i\qquad\text{for each } t \text{ and } \omega. \tag{6}$$

At the beginning of term $t+1$, the current mayor also decides whether to run for reelection, and all individuals choose whether to vote for the incumbent.

Our corruption data contain measurement error, as only a random subsample of a municipality's expenditure is audited. To account for this, we introduce a measurement error ν_t^i such that the stealing variable we and voters observe from 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 implies

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

that voters cannot perfectly infer the mayor's ability in the case of an audit. They can only update their beliefs about 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 end of term $t-1$, the federal government audits a fraction of municipalities at random, discloses the results publicly, and convicts a fraction of the mayors who were caught stealing. At the beginning of term t, the government collects fines from the previous-term mayors who were convicted based on the audit outcomes. If allowed to run, the current mayor, who has a given ability, observes the running shock and then decides whether to run for reelection. If the incumbent chooses not to or cannot run, a challenger to the incumbent with characteristics Z is drawn from the distribution $f_m(Z)$, where Z includes the challenger's age, education, savings, and ability.[20] Elections take place between candidates who cannot credibly commit to their campaign platform. The winning candidate governs in term t. If an incumbent chooses not to run (i.e. an open election) the challenger wins the election. Public funds, private inputs, and the wage shocks are then realized. The elected mayors choose consumption, savings, the fraction of public funds to invest in public consumption, and the fraction to steal for their own personal use.

3.5 Recursive Formulation

We solve and estimate the model using its recursive formulation. It requires computing three sets of value functions: one for voters, V_V, one for former mayors, V_{FM}, and one for current mayors, V_M. We assume that once a mayor leaves office they cannot return.[21]

[20] We draw challengers from the distribution of first-term mayors as opposed to the entire challenger pool because the characteristics of first-term mayors are more likely to represent the distribution of observables among runner-up challengers.

[21] In Brazil, term-limited mayors can run again for mayor as long as they sit out a term. But in the data, only 3 percent of ex-mayors ran for election after leaving office for at least one term.

Voters

To determine the voting decisions of residents, we have to derive their value function conditional on the incumbent winning reelection and conditional on the selected challenger becoming the new mayor. Voters then decide for whom to vote by comparing the two value functions. We abstract from the decision to become a candidate for mayor because, in the data, the probability that a citizen runs for election is below 1 percent.

Let $S_t^{V,C}$ denote the set of voter's deterministic state variables conditional on the challenger being elected, whose values are known prior to their decisions. This set includes the municipality's population size (d_t), the probability that the elected mayor is audited (p_t^{au}) and convicted (p_t^c), the voter's age (age_t^V), education (e_t^V), and savings (b_{t-1}^V), and the challenger's observable characteristics: age (age_t), education (e_t), and savings (b_{t-1}). We will denote with $S_t^{V,-C}$ the same set when the challenger's observables are excluded. Then, the voters' value function at the beginning of period t conditional on the challenger winning the election, $V_{V,C}^i\left(S_t^{V,C}\right)$, is the solution to the following problem:

$$V_{V,C}^i\left(S_t^{V,C}\right) = \max_{c_t^i, b_t^i} \; E\left[u^i\left(c_t^i, \bar{Q}_t\right) + \beta V_V^i\left(S_{t+1}^V\right)\middle|\, S_t^{V,C}\right] \tag{7}$$

$$s.t. \;\; c_t^i + b_t^i = w_t^i \bar{h} + R_t b_{t-1}^i,$$

where $V_V^i\left(S_{t+1}^V\right)$ denotes the voter's value function at the beginning of term $t+1$ and S_{t+1}^V is the corresponding set of deterministic state variables. We will derive the value function $V_V^i\left(S_{t+1}^V\right)$ and describe its state variables later in this section. For reference, we list the entire set of state variables used to compute all our value functions in Table 3.

Since voters compute the value function $V_{V,C}^i\left(S_t^{V,C}\right)$ before $\bar{Q}_t$ is realized, they have to compute the expectation over the wage shocks ϵ_t^w and all possible realizations of $\bar{Q}_t$. Embedded in this calculation are the stochastic state variables that affect the production of public consumption: the amount of public funds (f_t^{pu}) and private inputs (z_t^{pr}) the municipality receives, whether the elected challenger will be audited at the end of the term, and their ability (a).[22]

[22] When computing $E\left[V_V^i\left(S_{t+1}^V\right)\middle|\, S_t^{V,C}\right]$ in problem (7), voters should also take expectations over whether the elected challenger will be convicted and the fine shock. In the estimation, we discretize the elected challenger's savings to allow for its effects on the voter's value function to be more non-parametric. Because of this discretization,

The voter's value function conditional on the incumbent's winning the election $V^i_{V,I}$ is similar to problem (7) except that voters observe the amount of public consumption produced by the municipality in the previous term (Q_{t-1}), whether the incumbent was audited $\mathbb{1}^{au}_{t-1}$, and the amount stolen at $t-1$ if audited ($\mathbb{1}^{au}_{t-1}s_{t-1}$). The voter can use these additional outcomes to calculate the expected value function for this scenario, as they provide information about the incumbent's ability. Specifically,

$$V^i_{V,I}\left(S^{V,I}_t, Q_{t-1}, \mathbb{1}^{au}_{t-1}, s^i_{t-1}\mathbb{1}^{au}_{t-1}\right) = \max_{c^i_t, b^i_t} E\left[u^i\left(c^i_t, \bar{Q}_t\right) + \beta V^i_V\left(S^V_{t+1}\right) \middle| \, S^{V,I}_t, Q_{t-1}, \mathbb{1}^{au}_{t-1}, s^i_{t-1}\mathbb{1}^{au}_{t-1}\right]$$

$$s.t. \;\; c^i_t + b^i_t = w^i_t \bar{h} + R_t b^i_{t-1}, \tag{8}$$

where $S^{V,I}_t$ contains the same variables as $S^{V,C}_t$ except that the challenger's observable characteristics (age, education, and savings) are replaced with the incumbent's observables. The expectation is taken over the same stochastic state variables used in problem (7), but with the variables associated with the elected challenger replaced with those of the elected incumbent.

A citizen of a municipality votes for the incumbent if:

$$V^i_{V,I}\left(S^{V,I}_t, Q_{t-1}, \mathbb{1}^{au}_{t-1}, s^i_{t-1}\mathbb{1}^{au}_{t-1}\right) > V^i_{V,C}\left(S^{V,C}_t\right) + \varepsilon_{i,t},$$

where $\varepsilon_{i,t} \sim N(0,1)$ is an iid electoral shock. To make the model feasible to estimate, we assume that there is only one representative voter deciding the election that corresponds to a citizen of median age (33), education (high school completion), and wealth ($R\$47,387$) in 2005.

Let $S^e_t = \left\{S^{V,I}_t, S^{V,C}_t, Q_{t-1}, \mathbb{1}^{au}_{t-1}, s^i_t\mathbb{1}^{au}_{t-1}\right\}$ be the set of state variables that affect the voter's decision. The probability that the incumbent wins is

$$p\left(S^e_t\right) = \Phi\left(V^i_{V,I}\left(S^{V,I}_t, Q_{t-1}, \mathbb{1}^{au}_{t-1}, s^i_{t-1}\mathbb{1}^{au}_{t-1}\right) - V^i_{V,C}\left(S^{V,C}_t\right)\right), \tag{9}$$

where Φ is the standardized normal cdf. If an incumbent cannot run because of term limits, we set $p\left(S^e_t\right)$ to zero and the selected challenger becomes the new mayor.

To solve problem (7), we need to compute the expectation of the voter's value function at the

we need to assume for computational reasons that voters believe that future convictions and fines of the elected challenger do not affect their value function at $t+1$. Under this assumption, voters do not take expectations over those two variables.

beginning of period $t+1$, $E\left[V_V^i\left(S_{t+1}^V\right)\middle|\, S_t^{V,C}\right]$. Similarly, to find the solution of problem (8), we must derive $E\left[V_V^i\left(S_{t+1}^V\right)\middle|\, S_t^{V,I}, Q_{t-1}, \mathbb{1}_{t-1}^{au}, s_{t-1}^i \mathbb{1}_{t-1}^{au}\right]$. Let $p_R\left(S_{t+1}^M\right)$ denote the probability that an incumbent will run for reelection after entering term $t+1$. We will derived $p_R\left(S_{t+1}^M\right)$ and the corresponding state variables S_{t+1}^M at the end of this section.

If the mayor at t can run for reelection in term $t+1$, the voter's value function at the beginning of period t, $V_V^i\left(S_{t+1}^V\right)$, is defined by:

$$
\begin{aligned}
& V_V^i\left(S_{t+1}^V\right) \\
= \;& p_R\left(S_{t+1}^M\right) E\left[p\left(S_{t+1}^e\right) V_{V,I}^i\left(S_{t+1}^{V,I}, Q_t, \mathbb{1}_t^{au}, s_t^i \mathbb{1}_t^{au}\right) + \left(1 - p\left(S_{t+1}^e\right)\right) V_{V,C}^i\left(S_{t+1}^{V,C}\right)\middle|\, S_{t+1}^{V,I}, Q_t, \mathbb{1}_t^{au}, s_t^i \mathbb{1}_t^{au}\right] \\
+ \;& \left(1 - p_R\left(S_{t+1}^M\right)\right) E\left[V_{V,C}^i\left(S_{t+1}^{V,C}\right)\middle|\, S_{t+1}^{V,-C}\right],
\end{aligned}
$$

where the expectations are taken over all possible challengers since the value function is computed before a challenger is drawn. The set S_{t+1}^V includes Q_t, $\mathbb{1}_t^{au}$, $s_t^i \mathbb{1}_t^{au}$, and all the variables in $S_{t+1}^{V,I}$ and S_{t+1}^M.[23] The equation states that the mayor at t will run for reelection in term $t+1$ with probability $p_R\left(S_{t+1}^M\right)$. Conditional on running, this mayor will win reelection at $t+1$ against a specific challenger with probability $p\left(S_{t+1}^e\right)$. Conditional on the incumbent winning, the voter's expected lifetime utility at $t+1$ is equal to $V_{V,I}^i\left(S_{t+1}^{V,I}, Q_t, \mathbb{1}_t^{au}, s_t^i \mathbb{1}_t^{au}\right)$. With probability $1 - p\left(S_{t+1}^e\right)$ the challenger wins the election, in which case the voter's expected lifetime utility is equal to $V_{V,C}^i\left(S_{t+1}^{V,C}\right)$. Since voters compute $V_V^i\left(S_{t+1}^V\right)$ before a challenger is selected, they take the expectation over all possible challengers. With probability $1 - p_R\left(S_{t+1}^M\right)$, the incumbent chooses not to run and one of the challengers becomes the new mayor.

If the mayor at t cannot run for reelection in term $t+1$, one of the challengers becomes the new mayor and $V_V^i\left(S_{t+1}^V\right)$ simplifies to

$$
V_V^i\left(S_{t+1}^V\right) = E\left[V_{V,C}^i\left(S_{t+1}^{V,C}\right)\middle|\, S_{t+1}^{V,-C}\right]. \tag{10}
$$

Thus, in problem (7), $E\left[V_V^i\left(S_{t+1}^V\right)\middle|\, S_t^{V,C}\right]$, the voter's expected value function in period $t+1$ conditional on the challenger having won in period t, takes the following form:

[23] The variables in $S_t^{V,C}$ are either included in $S_t^{V,I}$ or integrated over when we take the expectation.

$$
\begin{aligned}
& E\left[V_V^i\left(S_{t+1}^V\right) \middle| \, S_t^{V,C}\right] \\
= \; & E\left[p_R\left(S_{t+1}^M\right) E\left[p\left(S_{t+1}^e\right) V_{V,I}^i\left(S_{t+1}^{V,I}, Q_t, \mathbb{1}_t^{au}, s_t^i \mathbb{1}_t^{au}\right) + \left(1 - p\left(S_{t+1}^e\right)\right) V_{V,C}^i\left(S_{t+1}^{V,C}\right) \middle| \, S_{t+1}^{V,I}, Q_t, \mathbb{1}_t^{au}, s_t^i \mathbb{1}_t^{au}\right]\right. \\
+ \; & \left.\left(1 - p_R\left(S_{t+1}^M\right)\right) E\left[V_{V,C}^i\left(S_{t+1}^{V,C}\right) \middle| \, S_{t+1}^{V,-C}\right] \middle| \, S_t^{V,C}\right].
\end{aligned}
$$

In problem (8), we need $E\left[V_V^i\left(S_{t+1}^V\right) \middle| \, S_t^{V,I}, Q_{t-1}, \mathbb{1}_{t-1}^{au}, s_{t-1}^i \mathbb{1}_{t-1}^{au}\right]$, the voter's expected value function in period $t+1$ conditional on the incumbent having won in term t. Since the incumbent won the election in term t and in our setting a politician can only hold office for two consecutive terms, the elected mayor will be term-limited at the end of t. Consequently, we can use equation (10) to compute $E\left[V_V^i\left(S_{t+1}^V\right) \middle| \, S_t^{V,I}, Q_{t-1}, \mathbb{1}_{t-1}^{au}, s_{t-1}^i \mathbb{1}_{t-1}^{au}\right]$, which takes the form:

$$
\begin{aligned}
E\left[V_V^i\left(S_{t+1}^V\right) \middle| \, S_t^{V,I}, Q_{t-1}, \mathbb{1}_{t-1}^{au}, s_{t-1}^i \mathbb{1}_{t-1}^{au}\right] &= E\left[E\left[V_{V,C}^i\left(S_{t+1}^{V,C}\right) \middle| \, S_{t+1}^{V,-C}\right] \middle| \, S_t^{V,I}, Q_{t-1}, \mathbb{1}_{t-1}^{au}, s_{t-1}^i \mathbb{1}_{t-1}^{au}\right] \\
&= E\left[E\left[V_{V,C}^i\left(S_{t+1}^{V,C}\right) \middle| \, S_{t+1}^{V,-C}\right] | S_t^S\right], \qquad (11)
\end{aligned}
$$

where S_t^S is the subset of $S_t^{V,I}$ that excludes the incumbent's observables.

Approximation of $V_{V,I}^i$. To compute the voter's value function if the incumbent is elected, $V_{V,I}^i$, we need the distribution of the stochastic state variables, including the incumbent's unobserved ability, conditional on Q_{t-1}, $\mathbb{1}_{t-1}^{au}$, and $s_{t-1}^i \mathbb{1}_{t-1}^{au}$. Given the complexity of our model, this distribution is unknown. To overcome this issue, we approximate the expectation of the voter's value function using the following specification:

$$
V_{V,I}^i\left(S_t^{V,I}, Q_{t-1}, \mathbb{1}_{t-1}^{au}, s_{t-1}^i \mathbb{1}_{t-1}^{au}\right) \approx \tilde{V}_{V,I}^i\left(S_t^{V,I}\right) + \delta_1 + \delta_2 f\left(Q_{t-1}\right) + \delta_3 \mathbb{1}_{t-1}^{au} \mathbb{1}_{\{s_{t-1}^i = 0\}} + \delta_4 s_{t-1}^i \mathbb{1}_{t-1}^{au}, \qquad (12)
$$

where $\tilde{V}_{V,I}^i\left(S_t^{V,I}\right)$ is the solution to the voter's problem (8) without conditioning on $t-1$ variables, $\mathbb{1}_{\{s_{t-1}^i=0\}}$ is an indicator equal to one if the mayor chose not to steal and $f\left(Q_{t-1}\right)$ is a function capturing how the representative voter transforms observed public consumption into information about the incumbent's ability.[24] In the estimation, we set $f\left(Q_{t-1}\right) = \ln \dfrac{Q_{t-1}}{d_{t-1}} - \ln E\left[\dfrac{Q}{d}\right]$, where

[24] Alternatively, one could assume a conditional distribution for ability and integrate the value function over this distribution. We chose not to use this approach because our approximation method offers potentially greater

$E\left[\frac{Q}{d}\right]$ is expected per-capita public consumption produced in the municipality by a challenger. We can then write the probability that the incumbent wins reelection as follows:

$$p\left(S_t^e\right)=\Phi\left(\tilde{V}_{V,I}^i\left(S_t^{V,I}\right)-V_{V,C}^i\left(S_t^{V,C}\right)+\delta_1+\delta_2 f\left(Q_{t-1}\right)+\delta_3 \mathbb{1}_{t-1}^{au}\mathbb{1}_{\{s_{t-1}^i=0\}}+\delta_4 s_{t-1}^i\mathbb{1}_{t-1}^{au}\right).$$

In the estimation of the model, we use an indicator equal to 1 if the incumbent was audited and stole a positive amount in place of $s_{t-1}^i\mathbb{1}_{t-1}^{au}$, because the effects on voters' decisions of the amount stolen if audited is not statistically significant.

Former Mayors

To compute the value function of a current mayor, we must first determine the value function of former mayors who were in office in the previous term but are no longer in power, $V_{FM}^i\left(S_t^{FM}\right)$, as it influences the current mayor's decision to run for reelection.

This value function at the beginning of term t can be computed using the voters' derivations. It is equal to the expectation of the voter's value function $V_V^i\left(S_t^V\right)$ calculated when the budget constraints of problems (7) and (8) are adjusted to account for potential fines the former mayor may incur if caught stealing. Specifically, $V_{FM}^i\left(S_t^{FM}\right)$ is equal to $E\left[V_{V,C}^i\left(S_t^{V,C}\right)|S_t^S\right]$ in equation (11), as the former mayor just left office; the budget constraints become,

$$c_t^i+b_t^i=w_t^i\bar{h}+R_t b_{t-1}^i-\mathbb{1}_{t-1}^{au}\mathbb{1}_{t-1}^{c}g\left(s_{t-1}^i\right); \tag{13}$$

and the set of deterministic state variables S_t^{FM} must be augmented to include whether the former mayor was audited $(\mathbb{1}_{t-1}^{au})$, convicted $(\mathbb{1}_{t-1}^{c})$, and the amount stolen (s_{t-1}), which are needed to compute the fine amount $g\left(s_{t-1}^i\right)$.

Current Mayor

To derive the value function of a current mayor at the beginning of term t, V_M^i, we need to consider the decision to run for reelection. If the mayor decides to run, they win with probability $p\left(S_t^e\right)$, which was derived in equation (9). Let the value function of a winning incumbent be V_{WM}^i, which

flexibility in how voters learn about the incumbent's ability.

can be derived by solving the following problem:

$$V^i_{WM}\left(S^{WM}_t\right) = \max_{c^i_t, b^i_t, s^i_t} \; E\left[u^i\left(c^i_t, \bar{Q}_t\right) + \rho - \kappa + \beta V^i_M\left(S^M_{t+1}\right) \middle| \, S^{WM}_t\right]$$

$$s.t. \; (4), \; (5), \; \text{and} \; (6),$$

where S^{WM}_t denotes the set of deterministic state variables. It includes the population size, the probability of an audit, the probability of conviction, the mayor's age, education, savings, and ability, whether they were audited, convicted, and the amount stolen at $t-1$. The expectation is taken over the stochastic state variables: the municipality's public funds and private inputs, whether the mayor will be audited and convicted, and the fine shocks. A winning mayor who is term-limited after t solves the same problem except that $V^i_M\left(S^M_{t+1}\right)$ is replaced by $V^i_{FM}\left(S^{FM}_{t+1}\right)$.

With probability $1 - p\left(S^e_t\right)$ the challenger wins the election. Let $V^i_{LM}\left(S^{LM}_t\right)$ denote the value function of a mayor who ran but lost the election. This value function corresponds to the value function of a former mayor that just left office conditional on the challenger being the new mayor, after we subtract from it the cost of running κ. We can calculate it by solving the voter's problem (7) using the budget constraint for a mayor that just left office (13). The set S^{LM}_t therefore contains the variables in S^{VC}_t plus whether the former mayor was audited, convicted, and the amount stolen in the previous term.

The value function of an incumbent that chooses to run for reelection, V^i_{RM}, is:

$$V^i_{RM}\left(S^M_t\right) = E\left[p\left(S^e_t\right) V^i_{WM}\left(S^{WM}_t\right) + \left(1 - p\left(S^e_t\right)\right) V^i_{LM}\left(S^{LM}_t\right)\right],$$

where the expectation is taken over all possible challengers since the mayor decides to run before a challenger is drawn. Therefore, S^M_t includes all the variables in S^{WM}_t, S^{LM}_t, and S^e_t except the challengers' observables. If the mayor decides to forgo reelection, the corresponding value function is equal to that of a former mayor who just left office, $V^i_{FM}\left(S^{FM}_t\right)$.

A mayor eligible for reelection will choose to run if

$$V^i_{RM}\left(S^M_t\right) \geq V^i_{FM}\left(S^{FM}_t\right) + \epsilon^R_{i,t}, \tag{14}$$

where $\epsilon_{i,t}^{R}$ is the running shock that accounts for events occurring before the election that we do not explicitly model (e.g. positive or negative news coverage about a candidate). If the mayor decides against running for reelection, the selected challenger will become the new mayor. The value function of a current mayor before observing the shock is:

$$V_{M}^{i}\left(S_{t}^{M}\right)=E\left[\max\ \left\{V_{RM}^{i}\left(S_{t}^{M}\right),V_{FM}^{i}\left(S_{t}^{FM}\right)+\epsilon_{i,t}^{R}\right\}\right],$$

where the expectation is taken over $\epsilon_{i,t}^{R}$. We can also compute the probability of running for reelection, which is needed to determine the voters' and former mayors' value functions at the beginning of period t, $V_{V}^{i}\left(S_{t}^{V}\right)$. It takes the following form

$$p_{R}\left(S_{t}^{M}\right)=\Phi\left(V_{RM}^{i}\left(S_{t}^{M}\right)-V_{FM}^{i}\left(S_{t}^{FM}\right)\right).$$

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. Therefore, we normalize κ to zero and estimate μ_R. The normalization is without loss of generality because we do not need to know κ separately from μ_R in our policy evaluations.

4 Estimation

In this section, we discuss the identification and estimation of the model's parameters. We estimate all the model's parameters except for 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)).[25] For the probability of being convicted conditional on being audit and caught stealing, we use the estimates presented in Avis, Ferraz, and Finan (2018) and compute it to be equal to 0.756.[26]

[25]The value $\delta=2$ and $\beta=0.98$ are generally used for ordinary individuals in a population and not for politicians. The implicit assumption is that politicians have the same curvature for the utility for private consumption and same discount factor.

[26]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

Production Function and Ability Parameters: α_1, α_2**,** μ_a**,** σ_a**.** We can estimate and identify the parameters of the production function and the ability distribution from our sample of first-term mayors. Let $\bar{a}_i = \log a_i - \mu_a$ denote demeaned log ability. The log of the production function (2) for first-term mayors is:

$$\log \frac{Q_t}{d_t} = \alpha_1 \log \frac{z_t^{pu}}{d_t} + \alpha_2 \log z_t^{pr} + \mu_a + \bar{a}_i, \tag{15}$$

where the public inputs z_t^{pu} depend on the mayor's stealing and thus ability. Let y_t be a vector of variables that are mean independent of $\bar{a}_i$, but are correlated with z_t^{pu}, and let $Y_t = [z_t^{pr}, y_t]$, where z_t^{pr} is the private sector inputs. Since $E\left[\bar{a}_i | Y_t\right] = E\left[\bar{a}_i\right] = 0$, we can estimate the parameters of (15) by two-stage least square regression, where we instrument the endogenous variable $\log \frac{z_t^{pu}}{d_t}$ with population thresholds corresponding 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.[27] The use of instrumental variables also controls for the measurement error in stealing. We employ the residuals of this regression to estimate the standard deviation of ability.

The estimation of the production function parameters requires information on the amount diverted by mayors. Thus, we have to restrict the sample to mayors that were audited. This sample restriction does not invalidate the identification result because municipalities were audited at random.

Wages of Former Mayors Parameters: $\gamma_0, \ldots, \gamma_5$ **and** σ_{fm}**.** The wage process of former mayors is assumed to be linear and independent of past stealing. We can estimate the parameters

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, $Pr(Conv|Audit = 1\&Stealing = 1) = 0.072/(0.13 * 0.73) = 0.756$.

[27] 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)).

$\gamma_0, \ldots, \gamma_5$ and σ_{fm} of the wage equation (3) by running an OLS regression of wages of former mayors on education, age, and municipality size.

Fine Parameters: μ_τ, σ_τ. In Section 3.2, we define the fine variable, τ, as a multiple of the amount stolen and assume it to be distributed as $\log \tau \sim N(\mu_\tau, \sigma_\tau)$. Since we observe actual fines and amounts stolen, we can estimate the parameters μ_τ and σ_τ using the sample mean and standard deviation of the log of the observed variable τ.

Distribution of Mayor's Characteristics: $f_m(Z)$. The vector of mayor's characteristics Z includes the mayor's age, education, savings, and ability. We assume that the mayor's ability is independent of the remaining characteristics. Since we observe age, education, and savings, we draw them from their joint empirical distribution.

Distribution of Funds and Private Inputs: $f_{pu}(f^{pu})$, $f_{pr}(f^{pr})$. We observe public funds and private inputs. We initialize these variables by assigning the observed values to each municipality. In subsequent periods, we draw them from their empirical joint distribution conditional on municipality size.

Remaining Parameters

There are ten remaining parameters: the relative taste for public consumption parameter, θ, the rivalry parameter η, and the utility from being in power parameter, ρ, all in the utility function (1); the parameter accounting for prosecutorial delays δ_τ introduced in Section 3.2; the electoral parameters $\delta_1 - \delta_4$ in the voter's value function (12); the mean and variance of the shock to run for reelection, μ_R and σ_R, in inequality (14); and the standard deviation of the measurement error σ_ν introduced in Section 3.3. We estimate these parameters using dynamic programming and Simulated Method of Moments (Gourieroux and Monfort (1996)), which we discuss in more detail in the appendix. Below, we describe the moments we use in the estimation and why they help identify our parameters of interest.

Relative Taste for Public Consumption. In our model, an increase in θ reduces stealing among second-term mayors all else equal. To understand why, consider the three factors that 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. Conditional on the first two factors, with a higher θ, second-term mayors will reduce their stealing because they care more about public consumption. Thus, to identify this parameter, we use average stealing among second-term mayors.

Prosecutorial Delays. An increase in δ_τ reduces stealing among first-term mayors. We can think of stealing in the first term 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 $\mu_\tau = \bar{\mu}_\tau + \delta_\tau$ and standard deviation σ_τ. An increase in δ_τ raises the negative return on this asset and thus reduces its demand. To identify this parameter, we therefore use average stealing among first-term mayors.

Electoral Parameters. The parameter δ_1 measures incumbency advantage.[28] Thus, larger values of δ_1 lead to higher probabilities of reelection. The parameter δ_2 determines how observed per-capita public consumption influences the incumbents' likelihood of reelection. In our model, voters adjust favorably their expected value function conditional on the incumbent's being reelected if they observe a per-capita public consumption that is greater than the amount produced by an expected challenger. The favorable adjustment and, thus, the probability of reelection increases with δ_2. The electoral parameter δ_3 accounts for changes in voter's expected lifetime utility conditional on the incumbent being reelected due to a clean audit. This parameter has a positive effect on their probability of being reelected if a mayor was audited and did not steal. The parameter δ_4 plays a similar role to δ_3 but for stealing. It determines the effect of stealing on the incumbent's reelection probability. Specifically, the election probability of an incumbent who stole declines the more negative the parameter is. In sum, we use the following moments to identify the electoral parameters: the probability an incumbent is reelected, the probability an incumbent is reelected after producing more than the 75^{th} percentile of per-capita public consumption, the probability of being reelected conditional on having a clean audit, and the probability of being reelected conditional on having been audited and caught stealing.

[28] It captures aspects of the incumbency advantage that we do not explicitly model (e.g. the incumbent's political organization) or any differential selection between challengers and first-term mayors.

Degree of Rivalry. In our model, average stealing among mayors increases with the rivalry parameter η. Moreover, the rate of increase in average stealing with η is larger for municipalities with larger population sizes. To see why, recall that a 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. To identify this parameter, we use the difference in average stealing between small and large municipalities.

Utility from Being in Power and Running Shock Parameters. A higher ρ increases the difference in the probability of running between incumbents with a clean audit and incumbents that were found to be corrupt; a rise in μ_R increases the probability of running of audited mayors that did not steal; and an increase in σ_R reduces the unconditional probability of running. To see why, note that incumbents running for reelection draw shocks from a distribution with mean μ_R and standard deviation σ_R regardless of the value of being reelected. This is not the case for the utility from being in power: incumbents who have a higher probability of winning are more likely to run and more likely to enjoy ρ. This is the case for audited incumbents who were not found to be corrupt. All else equal, they have a higher probabilities of winning and, thus, of running and enjoying ρ than audited mayor who were found to be corrupt. An increase in ρ has therefore a positive effect on the likelihood of running for both groups, but the effect is larger for incumbents with a clean audit. Thus, the difference in the probability of running increases, independent of μ_R and σ_R. Given ρ, we can identify the parameters μ_R and σ_R 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. In sum, we identity these three parameters using the following moments: the probability of running for reelection conditional on a clean audit, the same probability conditional on having been audited and found to be corrupt, and the unconditional probability of running for reelection.

Standard Deviation of Measurement Errors. The standard deviation of the fraction stolen is by construction increasing in the parameter σ_ν, so we use it to identify this parameter.

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 various anti-corruption policies, including policy bundles.

5.1 Model Fit

Fit of Moments Used in the SMM Estimation. We use eleven 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 a clean audit; (vi) reelection rates conditional on running and having been caught stealing; (vii) probability of running conditional on having a clean audit; (viii) probability of running conditional on having been caught stealing; (ix) unconditional probability of running; (x) difference in average stealing between small and large municipalities; (xi) the standard deviation of fraction stolen.

In panel A of Table 4, 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 first (5.6%) and second-term mayors (7.3%), which we use to identify the effect of delays 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. But even in this case, the difference is only 0.8 percentage points.

Fit of Moments Not Used in the Estimation. In panel B of Table 4, we assess how well our model matches moments not used in the estimation. We compare seven moments and do quite well in general, with the largest difference being in the 90^{th} percentile of the fraction of funds stolen (0.171 versus 0.196). 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. Second-term mayors, however, are a selected subset of first-term mayors. Thus, the model has to generate the right amount of electoral selection to match the mean log-ability of second-term mayors in the data. We find that the model does relatively well in matching this moment, with a simulated mean of -0.40 that is slightly lower than the mean in the data of -0.42. This is an indication that the type of selection generated by the model is consistent with the patterns observed in the data.

In Section 2, we document 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 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, as documented in panel B of Table 4.

5.2 Parameter Estimates

Preference Parameters. We present estimates of our model parameters in Table 5. We estimate a relative taste for public consumption equal to 0.069. This implies that individuals value the utility from private consumption about 10.4 times more than the utility from public consumption at the mean of their public and private consumption. Our rivalry parameter ($\eta = 0.94$) suggests that public consumption is characterized by a relatively high degree of rivalry, but it still generates significantly higher levels of welfare compared to a fully private good.[29] Mayors also enjoy high levels of utility from being in power. We estimate the utility from being in power to be 1.3, which corresponds to 63% 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 0.34. To interpret its magnitude, recall that the probability of getting reelected is 57.3%. If we re-simulated

[29] For example, if a mayor were to produce \$1,000,000 of public goods in a municipality with 100,000 inhabitants, then our estimate implies that the citizens enjoy twice the level of consumption compared to a fully rivalrous public good.

the model with this parameter set to zero, this probability decreases to 50.4%, a 12.0% decline. Voters also punish mayors found to be corrupt. If we re-simulate the model with the coefficient on having been caught stealing ($\delta_4 = -0.24$) set to zero, the probability of reelection conditional on stealing increases by 9.1%, and stealing among first-term mayors increases by 5.1%. Similarly, voters reward mayors who were audited and not found to be corrupt. If we set δ_3 to zero, the probability of running conditional on not stealing declines slightly from 78.7% to 77.1%, but the probability of reelection conditional on not stealing goes down from 65% to 58.3%. The log difference between the incumbent's per-capita public consumption and the expected challenger's also affects electoral decisions. If we set the corresponding parameter (δ_2) to zero, the probability of winning reelection conditional on producing an amount of per-capita public consumption that is higher than the 75th percentile (the moment we use to identify δ_2) declines from 61.3% to 57.1% and stealing in the first term increases from 5.6% to 6.2%.

Production Function. Panel C of Table 5 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: if we re-simulate our model after increasing the mean of the ability distribution by 0.1 standard deviations, public consumption increases by 4.5 percent. Similarly, if we re-simulate the model by setting the mean of the ability distribution equal to the estimated mean ability of second-term mayors, per-capita public consumption increases by 3 percent.

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, we allow for a reduction in the mean of the fine distribution, which we estimate at -0.789. Overall, these estimates suggest that corrupt mayors, on average, have to repay the original amount and pay a fine equal to 50% of the amount stolen.

Wage Process Parameters. In Panel E of Table 5, we report the estimated coefficients for the wage process of former mayors. As expected, wages are positively associated with having a college degree and exhibit an inverted u-shape with respect to age. Former 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 or reelection incentive effect*. Second, elections allow voters to select more able politicians, which is referred to as a *selection effect*. As the literature has emphasized, these two effects can explain why, in the data, second-term mayors steal more than first-term mayors but 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.[30]

In Table 6, 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.[31]

Second-term mayors steal about 24.5 percent more than first-term mayors.[32] Public consumption is also higher among second-term mayors by 3.9%, which indicates positive selection. We can see this more clearly in the subsequent rows. Second-term mayors are, for example, wealthier,

[30] In our model, corruption differences between first-term and second-term mayors are generated by electoral incentives, selection, and optimal decisions. The differences could also be generated by learning how to engage in corruption over the course of the first term. However, Ferraz and Finan (2008) do not find any evidence of learning, even when they compared second-term mayors to first-term mayors who had held a previous political office. Therefore, we have decided to abstract from this consideration.

[31] As we mentioned above, the audit program started in 2003, which corresponds to the middle of a mayor's 2001-2004 term. Given that a unit of time in our model refers to a 4-year term, we decided to apply the 16.8% audit probability to the entire 2001-2004 term. While the timing is imperfect, our assignment does allow for the possibility that any public discourse preceding the program's implementation may have increased mayors' subjective probabilities on being audited.

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

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 now compute the relative importance of reelection incentives, selection effects, and wealth effects for the differences in stealing and public consumption between first and second-term mayors. 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 from wealth effects.

To isolate the effect of each channel, we simulate the model under different environments. 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 compare two groups with the same incentives, the same distribution of observable and unobservable traits, but different savings levels. In Table 7, 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 20.4%, an effect approximately four times the size of the selection effect (40.3% vs 9.1%). The wealth effect accounts for approximately half of the difference we see in stealing between first and second-term mayors. This finding stems from the fact that second-term mayors are significantly wealthier than first-term mayors and given the concavity of the utility function, wealthier mayors engage in less corruption.

As for per-capita public consumption, we see that most of the difference between first and second-term mayors comes from selection effects, which accounts for more than 89.9% of the total effects. This result makes sense given that ability has a direct effect on the production of public goods and the positive selection of second-term mayors relative to first-term mayors. Reelection incentives account for an additional 4.1% of the total effects, whereas wealth effects contribute the remaining 6%. In contrast to 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. These results suggest that policies that can enhance reelection incentives and allow voters to better screen politicians might provide effective instruments for combating corruption and improving voter's welfare. 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 anti-corruption policies. The first policy increases the audit probability, which, based on previous research (e.g., Olken (2007)), is expected to reduce overall corruption. The second policy is Brazil's Clean Record Act, introduced in 2010, which deters corruption by barring politicians convicted of corruption from running for office. The third policy extends the maximum number of mayoral terms from two to three. Since term-limited mayors often engage in more corruption (e.g., Ferraz and Finan (2011)), this policy could reduce corruption by introducing additional re-election incentives and allowing voters to better assess incumbents. The final policy doubles mayors' wages. According to Becker and Stigler (1974), increasing the value of a position can discourage stealing, provided there is a credible threat of punishment. We compare these policies to a base case where the audit probability is set at 5% for all periods. In the next subsection, we show how combining some of these policies can further reduce corruption.

Brazil's Audit Program (Audit). In 2003, Brazil's Federal Government introduced an audit program that substantially increased the audit probabilities for mayors 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 left y-axis the average fraction stolen for different audit probabilities. As expected, stealing declines as the probability of an audit increases. As we go from a 6% audit probability to 20%, average stealing decreases by almost 40%, from 9.9% to 5.9%. For the case of Brazil's audit program, our simulations suggest that the program has reduced corruption by about 3.7 percentage points, which represents a 36.2% 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.[33] From 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 the audit probability adopted by Brazil's Federal Government 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 right 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 15%. At this point, the willingness to pay is

[33] Note that this estimate only includes transportation costs, the auditors' salaries, and per diem. It excludes any overhead costs. If we, instead, divide the program's annual budget ($25 million) by the number of yearly audits, we would estimate an average cost of $150,000 per audit.

maximal at 0.60, suggesting that the representative citizen would be willing to pay 0.60% 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 0.58.

Besides increasing the probability of an audit, we can also induce deterrence effects further 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 log-normal 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 the 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. However, even with a probability of conviction equal to 1, stealing is still significant since the audit probability is only 5%. Such a policy, if feasible, would only decrease stealing by 6%.

Clean Record Act (CRA). The Clean Record Act bans individuals convicted of corruption from holding public office for eight years. This law increases the implicit cost of engaging in corruption, thereby incentivizing political candidates to avoid such behavior. In our model, we simulate the effects of this policy by prohibiting incumbents who have been convicted of stealing from participating in future elections.[34]

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 15.1%, relative to the base case. The policy reduced corruption by 32.3% among first-term mayors, but increased corruption among second-term mayors by 7.4% relative to the base case. This pattern is expected because the CRA

[34] 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.

policy works via electoral incentives, which exist only for first-term mayors. In contrast, some second-term mayors may engage in increased corruption to make up for refraining from stealing during their first term. Under the CRA policy, these reelected mayors actually start their second term with on average lower savings than they would have under the base case, prompting them to steal relatively more. Despite its adverse effect on second-term mayors, citizens are willing to pay 0.98% of their 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. We have documented the importance of electoral incentives and selection effects in explaining the levels of stealing and public consumption observed in the data. Since both of these effects are a byproduct of elections, allowing mayors to serve more terms could help reduce corruption. To evaluate this policy, we simulate the effects of allowing mayors to be reelected for a third term.[35] Under this scenario, corruption reduces by 9.7% relative to the base case, and citizens are willing to pay 0.50% of their income for such a policy (see Figure 2). The largest reduction in stealing comes from second-term mayors who now have reelection incentives and reduce their stealing by 21.57% 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 example, the average ability among third-term mayors is 5.5 and 2.6 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 by only 9.7%. 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 increases the incentives for being

[35] To simulate this policy, we assume that the effect of Q_{t-1} on the voter's expectation at time t that the incumbent runs at $t+1$ is negligible because for the 3-term policy, the expectation in equation (11) is conditional on Q_{t-1}.

reelected, which should discourage mayors from stealing. In Figure 2a, we see that a doubling of wages decreases corruption by 12.8%. Figure 2c reports the willingness to pay for the policy when we account for its cost. In spite of the cost, citizens are willing to pay 0.62% of their annual income to adopt the wage policy.[36]

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 reduces corruption substantially but is also 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 8, 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 and welfare.

Among our individual policies, the Clean Records Act produces 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 34.0%, compared to 15.1% (for the CRA) or 9.7% (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 27%. Importantly, citizens are willing to pay 1.17% of their annual income for this policy bundle, which is 19.3 basis points higher than their willingness to pay for just the CRA.

We find similar improvements for a policy bundle that combines the 3-term with a 13% probability of being audited. We set the audit probability to 13% because it is the audit probability

[36]We have also computed the increase in wages that maximizes the average willingness to pay. It requires wages to be 4.4 times those observed in the data. Citizens are willing to pay on average 0.9 of their income for this policy.

that maximizes the willingness to pay for this combined policy. As shown in Figure 3a, combining the 3-term with a 13% audit probability reduces corruption by 34.9%. While this is an impressive reduction, it is similar to the reduction one could achieve by implementing just the (optimal) audit policy. Nevertheless, citizens would prefer this policy bundle to the individual optimal audit policy (WTP for bundle = 0.82% versus WTP for audit = 0.60%). 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 a bundle that combines the CRA policy with an audit policy that has a welfare-maximizing audit probability of 9%. 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 also be quite effective at reducing corruption (34.9% relative to baseline), with most of the reduction again coming from first-term mayors (60%). The willingness to pay for the CRA-audit policy bundle is higher than the 3-term+CRA policy bundle because of the lower optimal audit probability and its effect on the policy costs.

Citizens' highest willingness to pay combines all three policies: CRA, 3-term, and a welfare-maximizing 9% audit probability. For this policy bundle, citizens are willing to pay 1.28% of their annual income (see Figure 3c). The policy also reduces average stealing by an impressive 56.4% relative to the base case. It is also equally effective among first-term mayors (59.9% reduction relative to the base case) and second-term mayors (50.9% reduction relative to the base case).

Given that audits are the most effective tool for last-term politicians but are also expensive, we might do better if we only audit mayors in their last term. Thus, we also consider a policy bundle that includes the CRA, 3-term, and a welfare-maximizing 87% audit probability for third term mayors. This policy reduces corruption significantly (33.8% relative to the base case) and voters value it with a willingness to pay of 1.24% of their income. Nevertheless, with its high audit probabilities among third-term mayors and the relatively moderate reductions in corruption, voters would still slightly prefer the previous policy bundle.

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 12.8% under the base case probability of an audit of 5%. When combining these two policies, the optimal audit probability is 10%, and stealing falls by 27.3% relative to the base case (see Figure 3a). If we compare this reduction to the effect of each policy separately – 10% audit policy = 17.8% reduction in corruption and wage policy = 12.8% reduction in corruption – we see some degree of substitutability in the effects of the combined policy. 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 0.50% to 0.21%, a 58% decline. Overall, these results suggest that selection on ability is an important channel for generating the effects of these policies on voter's welfare.

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 will become more positively selected as we increase the probability of an audit or extend term limits. Although we cannot directly test this assumption, we do provide three pieces of evidence to suggest that either the effects of the policies on the candidate pool are likely to be minimal or that moderate changes to the candidate pool do not affect our general findings.

In Appendix Table 10, we test whether eligibility into the audit program affects the type of candidates that run for election. Recall that when the program was first introduced in 2003, only municipalities with a population below 450,000 were eligible for the program. Because the program significantly increased the probability of an audit for eligible municipalities, we can test the extent to which higher audit probabilities change the candidate pool by estimating a difference-in-differences regression. 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 is eligible for a program audit post 2003. The regression also controls for time and municipality fixed effects, with standard errors clustered at the municipality level. As the table reports, we do not find any evidence that the program affected either the size or composition of the candidate pool, as measured by their gender, education, or age.

In Appendix Table 11, we test whether the introduction of the CRA policy affected characteristics of the candidate pool, by regressing an attribute of the candidate pool on an indicator for whether the municipal election was held after the introduction of the CRA policy. Because this policy was introduced nationwide, time trends may confound any changes in the candidate pool we observe over time. To account for time trends, we consider two different specifications: one with linear time trends (odd columns) and one with quadratic time trends (even columns). All the regressions also include municipal fixed effects, with standard errors clustered at the municipality level. As one can see from the table, we do not find robust evidence that the CRA policy affected the candidate pool. The only stable finding is on the average age of the candidate pool, which only increased by half a year post CRA.

Finally, we also re-simulate all 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 12. 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.

6 Conclusions

In this paper, we develop and estimate a dynamic model of 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. By combining policies that enhance reelection incentives with an audit policy that has a lower audit probability, we could reduce corruption at a lower cost. Our findings suggest that residents in Brazil are willing to pay 1.28% of their annual income for such a multi-pronged approach, which is more than 30% of what they would be willing to pay for any individual policy.

Our estimates and policy analysis are clearly applicable to Brazil. However, our framework is quite general and can be used to understand corruption in any context where politicians manage large public budgets and are elected representatives within a democracy with strong electoral protections. In our model, the effectiveness of anti-corruption policies hinges on three main mechanisms: electoral incentives, disciplinary actions, and selection. The extent to which our policy findings can be extrapolated to other settings depends on the strength of these three channels in those contexts.

In Brazil, corrupt politicians are held accountable by voters as well as the legal system, both of which generate significant selection effects. Therefore, we would expect our policies to have similar effects in countries where these electoral and disciplinary forces against corruption have also been empirically documented, such as in Colombia (Garbiras-Díaz and Montenegro (2022)), India (Banerjee et al. (2020)), Italy (Chang, Golden, and Hill (2010)), Mexico (Larreguy, Marshall, and Snyder (2020)), and Uganda (Buntaine et al. (2018)).

In settings lacking effective disciplinary mechanisms, many of our proposed policies may be less effective. For instance, corruption audits are effective in Brazil because they enable voters to remove corrupt officials and allow the judiciary to convict and penalize them. Without a functioning judiciary, an audit policy may require a higher audit probability, leading to a reduced willingness to pay for such measure. In contexts with strong electoral incentives but weak disciplinary mechanisms, policies that foster strong selection effects through voting – such as extending term limits – may be more welfare enhancing.

Ultimately, assessing whether our recommended policies are applicable to other settings is an important empirical question. We hope our approach serves as a blueprint for how to analyze various policies and combinations of policies within different contexts.

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.

Banerjee, Abhijit, Nils T Enevoldsen, Rohini Pande, and Michael Walton. 2020, April. "Public Information is an Incentive for Politicians: Experimental Evidence from Delhi Elections." Working paper 26925, National Bureau of Economic Research.

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.

Buntaine, Mark T, Ryan Jablonski, Daniel L Nielson, and Paula M Pickering. 2018. "SMS texts on corruption help Ugandan voters hold elected councillors accountable at the polls." *Proceedings of the National Academy of Sciences* 115 (26): 6668–6673.

Chang, Eric C. C., Miriam A. Golden, and Seth J. Hill. 2010. "Legislative Malfeasance and Political Accountability." *World Politics* 62 (2): 177–220.

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.

Garbiras-Díaz, Natalia, and Mateo Montenegro. 2022. "All Eyes on Them: A Field Experiment on Citizen Oversight and Electoral Integrity." *American Economic Review* 112 (8): 2631–68 (August).

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).

Kennan, John. 2006. "A Note on Discrete Approximations of Continuous Distributions." *Manuscript.*

Larreguy, Horacio, John Marshall, and Jr. Snyder, James M. 2020. "Publicising Malfeasance: When the Local Media Structure Facilitates Electoral Accountability in Mexico." *The Economic Journal* 130 (631): 2291–2327 (05).

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
Ran for reelection in 2004	3,307	0.718	0.450
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
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	Log GDP per capita	Ran 2004	Reelected 2004	Log Wage
	(1)	(2)	(3)	(4)	(5)
Second term	0.016*	0.241			
	(0.009)	(0.256)			
Log of Amount Stolen		-0.024**			
		(0.010)			
Federal Funds	-0.003	0.475***			
	(0.009)	(0.072)			
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)
Change in GDP per capita 2001-2004			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 Squared					-0.107***
					(0.019)
Number of observations	474	474	3254	2333	3389
R^2	0.02	0.65	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: Deterministic and Stochastic State Variables

Description	Variable	$V^i_{V,C}$	$V^i_{V,I}$	V^i_{V}	V^i_{FM}	V^i_{WM}	V^i_{LM}	V^i_{RM}	V^i_{M}
Panel A: Deterministic State Variables									
Population Size	d_t	✓	✓	✓	✓	✓	✓	✓	✓
Probability of an Audit	p^{au}_t	✓	✓	✓	✓	✓	✓	✓	✓
Probability of a Conviction	p^c_t	✓	✓	✓	✓	✓	✓	✓	✓
Voters:									
Observable Characteristics	age_t, e_t, b_{t-1}	✓	✓	✓					
Former Mayor's:									
Observable Characteristics	age_t, e_t, b_{t-1}			✓	✓		✓	✓	✓
Audited	$\mathbb{1}_{\{au,t-1\}}$			✓	✓		✓	✓	✓
Convicted	$\mathbb{1}_{\{c,t-1\}}$			✓	✓		✓	✓	✓
Stealing	s_{t-1}			✓	✓		✓	✓	✓
Incumbent's:									
Observable Characteristics	age_t, e_t, b_{t-1}		✓	✓		✓		✓	✓
Ability	a_t			✓		✓		✓	✓
Public Consumption at $t-1$	Q_{t-1}		✓	✓					
Audited at $t-1$	$\mathbb{1}_{\{au,t-1\}}$		✓	✓		✓		✓	✓
Convicted $t-1$	$\mathbb{1}_{\{c,t-1\}}$			✓		✓		✓	✓
Stealing	s_{t-1}		✓	✓		✓		✓	✓
Challenger's:									
Observable Characteristics	age_t, e_t, b_{t-1}	✓					✓		
Panel B: Stochastic State Variables									
Amount of Public Funds	f^{pu}_t	✓	✓		✓	✓	✓	✓	
Private Inputs	z^{pr}_t	✓	✓		✓	✓	✓	✓	
Running Shocks	$\epsilon^R_{i,t}$								✓
Wage Shock	ϵ^w_t	✓	✓		✓		✓	✓	
Fine Shocks	τ_t					✓		✓	
Incumbent's at t:									
Observable Characteristics	age_t, e_t, b_{t-1}								
Ability	a_t		✓						
Audited	$\mathbb{1}_{\{au,t\}}$		✓			✓		✓	
Convicted	$\mathbb{1}_{\{c,t\}}$					✓		✓	
Challenger's at t:									
Observable Characteristics	age_t, e_t, b_{t-1}			✓	✓			✓	
Ability	a_t	✓		✓	✓		✓	✓	
Audited		✓			✓		✓	✓	

Table 4: Comparing Data Moments with Model Moments

	Model	Data	Corresponding Parameter
<u>Panel A</u>: Moments Used in the Estimation			
Average Stealing in the First Term	0.056	0.056	δ_τ
Average Stealing in the Second Term	0.073	0.073	θ
Pr[Reelection]	0.573	0.573	δ_1
Pr[Reelection$\|$Public Cons. $>$ 75^{th}perc.]	0.613	0.612	δ_2
Pr[Reelection$\|$Audited and Stealing $=0$]	0.649	0.650	δ_3
Pr[Reelection$\|$Audited and Stealing >0]	0.513	0.515	δ_4
Pr[Running$\|$Audited and Stealing $=0$]	0.787	0.779	ρ
Pr[Running$\|$Audited and Stealing >0]	0.689	0.684	μ_R
Pr[Running]	0.718	0.718	σ_R
Difference in Stealing, Small Vs Large Municipalities	0.240	0.240	η
Standard Deviation of Fraction Stolen	0.103	0.103	σ_ν
<u>Panel B</u>: Moments Not Used in the Estimation			
Fraction of Mayors Stealing	71.6	73.3	
25th Percentile Fraction of Funds Stolen	0.00	0.00	
Median Fraction of Funds Stolen	0.034	0.021	
75th Percentile Fraction of Funds Stolen	0.087	0.076	
90th Percentile Fraction of Funds Stolen	0.171	0.196	
Mean of Log Ability, Second term	-0.40	-0.42	
Per-capita consumption	4.997	4.958	

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 100 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 5: Parameter Estimates

	Parameter	Estimate	Standard Error
Panel A: Preference Parameters			
Relative Value of Public Consumption	θ	0.069	0.003
Rivalry parameter	η	0.938	0.015
Utility of being in power	ρ	1.299	0.022
Panel B: Electoral Parameters			
Informational Value of Incumbency	δ_1	0.343	0.025
Difference in public consumption (incumbent vs challenger)	δ_2	0.308	0.014
Indicator if Audited and No Stealing	δ_3	0.335	0.026
Indicator if Audited and Stealing	δ_4	-0.240	0.004
Mean of Running Shock	μ_R	-0.055	0.002
Std Dev. of Running Shock	σ_R	0.474	0.033
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	0.248
Mean Log Ability Distribution Second term	μ'_a	-0.424	0.258
Standard Deviation Log Ability Distribution	σ_a	0.439	0.027
Std. Dev. Measurement Error in Stealing	σ_ν	0.620	0.134
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
Prosecutorial Delays Parameter	δ_τ	-0.789	0.055
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_{fm}	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 6: Political Outcomes & Selection – Baseline Model

	Full Sample (1)	First-term mayors (2)	Second-term mayors (3)
Fraction Stolen	0.103	0.094	0.117
Per-Capita Public Cons.	4.987	4.912	5.104
Age	49.482	50.295	48.195
Initial Wealth	441.996	339.856	603.619
College Education	0.367	0.406	0.304
Ability	0.717	0.704	0.737

Notes: This table presents simulated moments using the baseline model with a 5% audit probability in all terms, based on 100 simulations for each municipality.

Table 7: Relative Importance of Incentives, Selection, and Savings for the Difference in Stealing between First and Second-term Mayors

	$\%\Delta\{2^{nd} - 1^{st}\}$ mayors (1)	Reelection Incentives (2)	Selection Effect (3)	Wealth Effect (4)
Fraction Stolen	24.5%	40.25%	9.06%	50.69%
Public Consumption	3.91%	4.10%	89.93%	5.97%

Notes: This table presents simulated moments based on 100 simulations for each municipality.

Table 8: 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 affect 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 create stronger incentives (e.g., CRA or audits)
Double Wages	Reduces corruption	Expensive	Combine with audit policy to increase threat of punishment

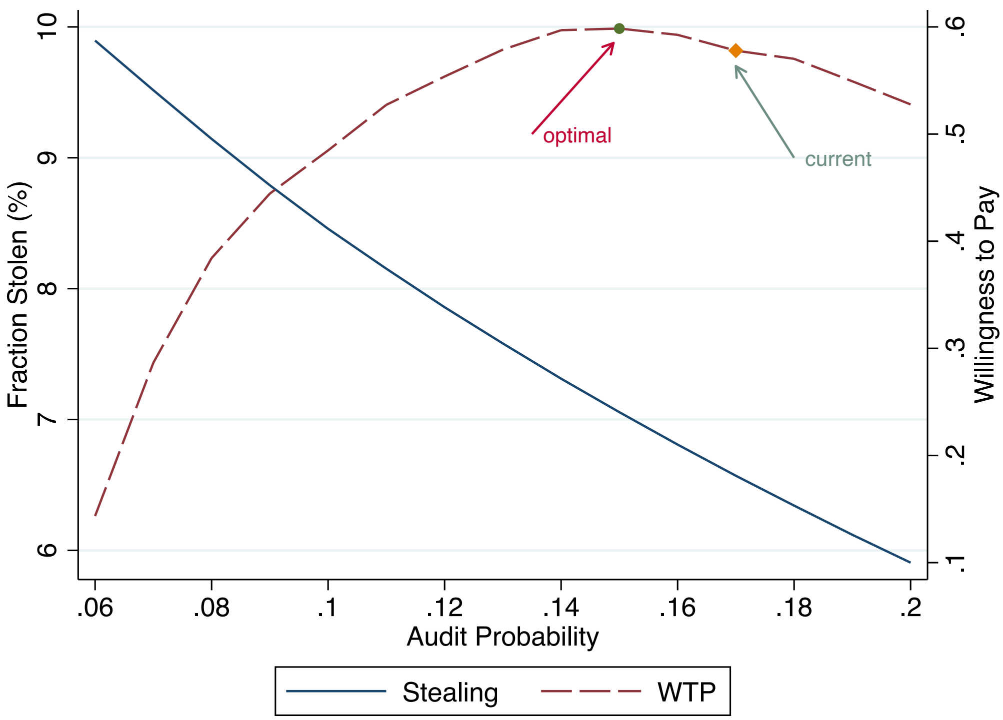
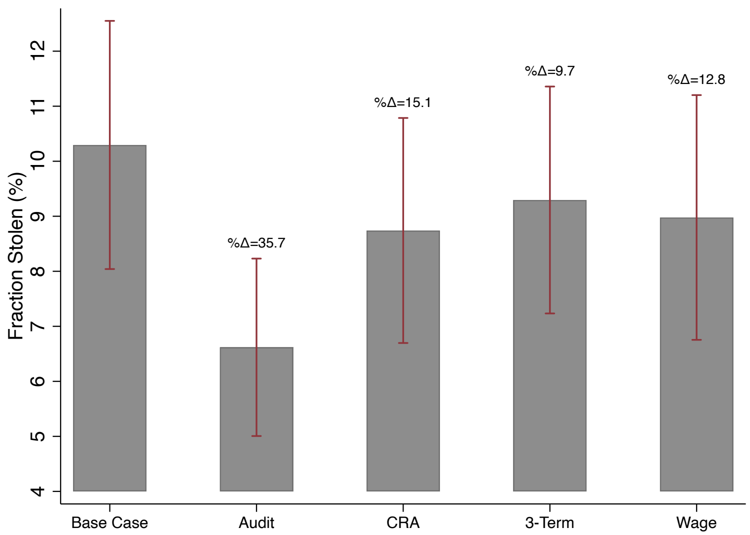
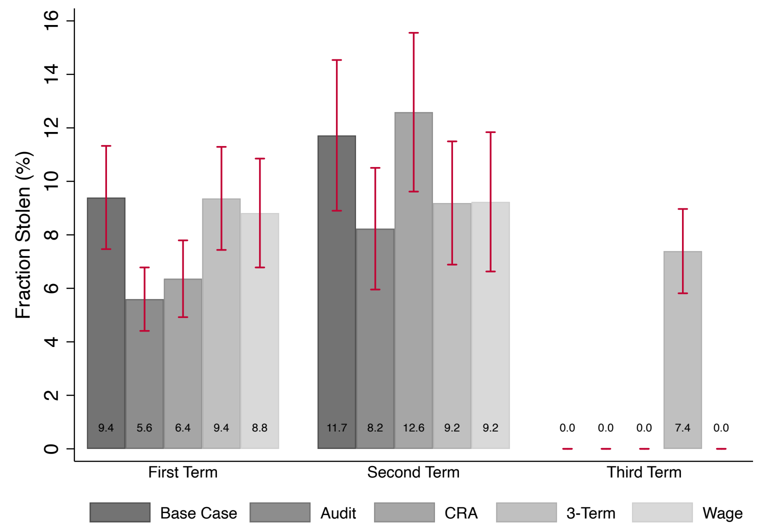


Figure 1: Optimal Audit Policy

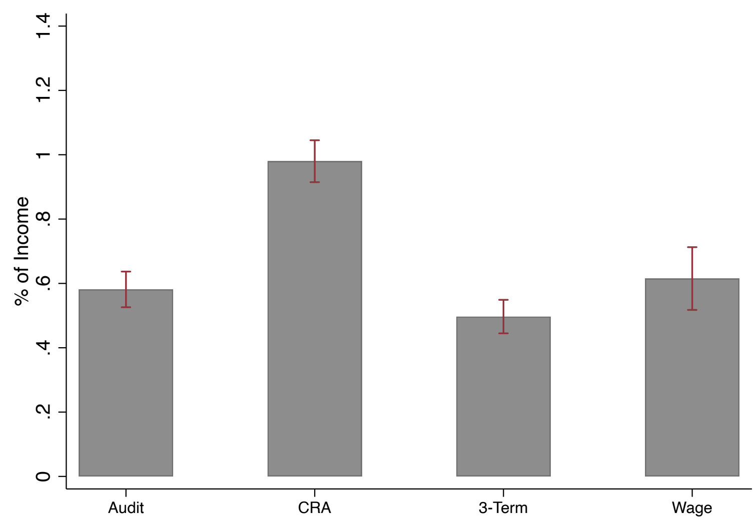
Notes: This figure presents results based on 500 simulations for each municipality.



(a) Average Stealing



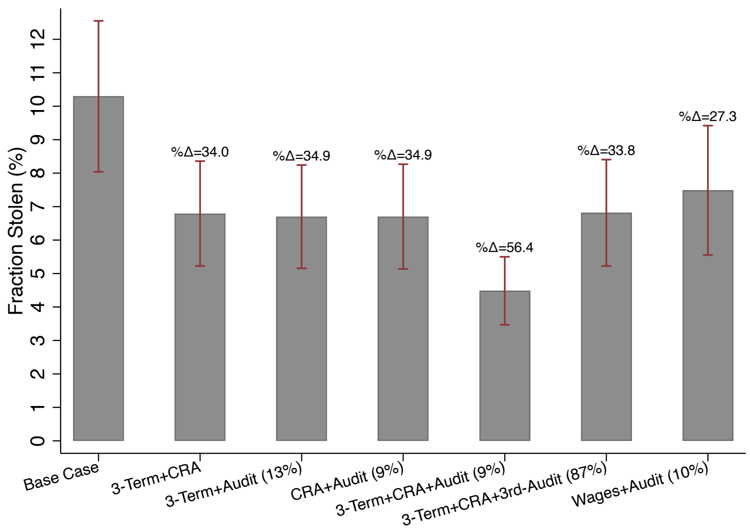
(b) Average Stealing by Term



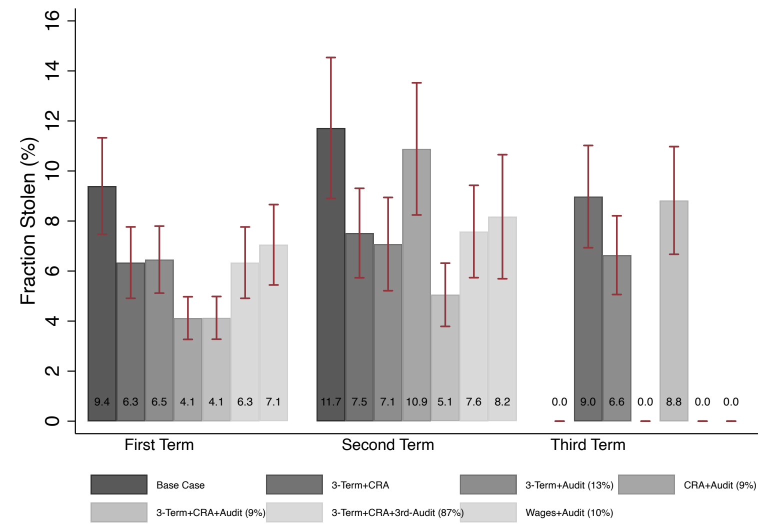
(c) Willingness to Pay

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

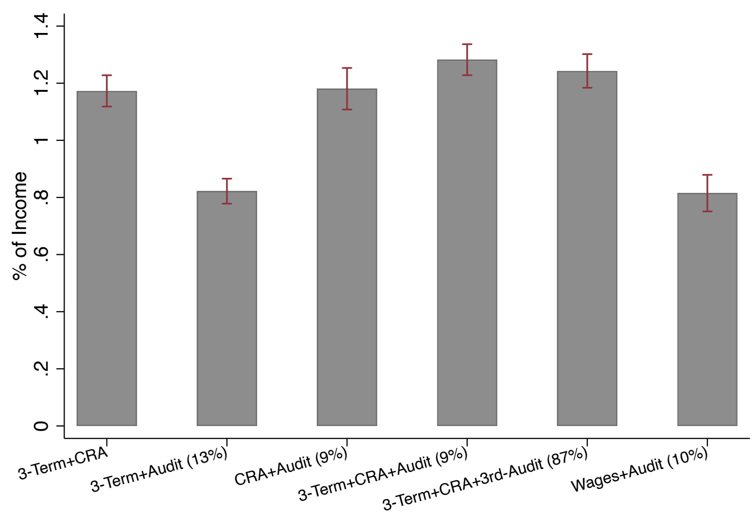
Notes: Each reported statistic is computed based on 100 simulations for each municipality. The 95% confidence intervals are computed by drawing the model parameters from their asymptotic distribution 256 times, calculating the relevance statistics for each draw of the parameters, and computing the confidence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

Combating Political Corruption with Policy Bundles

ONLINE APPENDIX

Frederico Finan and Maurizio Mazzocco

October 2024

A SMM 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.

We now describe the calculation of the value functions in more detail. As discussed in the paper, the computation is complex because we need to recover three distinct value functions: one for mayors, one for former mayors, and one for voters. These value functions are interconnected through the public good, which is produced by current mayors and benefits all citizens in the municipality. Moreover, current mayors influence the value functions of both former mayors and voters solely through their impact on the provision of the public good. By leveraging these interconnections within the model, we can simplify the computational complexity.

In the simulations, one period corresponds to a term of 4 years and we consider individuals from age $t = 1$ to age $t = T = 10$, which correspond to 40 years. We assign individuals in the simulations to one of 10 age bins depending on their actual age. The first age bin contains individuals with ages between 30 and 35, which includes the youngest mayors in our data. The last age bin includes individuals that are between 70 and 75. We therefore compute value functions for individuals from the age of 30 to the age of 75.

We start the simulations with the last period T. We first solve the problem of current mayors for each point of the discrete state space (continuous variables are discretized at this point). This is the most time-consuming part of the computation because the current mayors have the choice of stealing, which in turn affects the production of the public good. Importantly, the solution of this problem gives us the optimal amount of public goods provision for each point of the state space. We then multiply the utility at the solution by the probability of observing the stochastic variables of the corresponding point of the state space and add up the outcomes to derive the winning mayor's value function (V^i_{WM}) for each point of the deterministic state space.

We then move to solve the period-T problem of a voter in a municipality where an incumbent won the election. We do this for each point of the state space, which includes the observable characteristics of the incumbent. We can solve this problem, because for each set of characteristics of the newly elected mayor, we know the optimal production of the public good from solving the current mayor's problem. By multiplying the utility at the solution by the probability of observing the stochastic variables of the corresponding point of the state space and adding up, we can derive the value functions of voters conditional on the incumbent being reelected ($V^i_{V,I}$) for each point of the deterministic state space. Note that we can use the solution to this problem for former mayors who were in power two or more terms ago, as their problem is identical.

Next, we solve the voters' problem conditional on the challenger winning the election. We do this for each point of the state space, which includes the observable characteristics of the challenger that just won the election. As in the previous case, we can solve this problem because we know the optimal amount of public goods produced by mayors, including newly elected challengers. The solution to this problem gives us also (ii) the solutions for losing mayors conditional on a particular challenger (V^i_{LM}) and (ii) the solution for someone that enters the period as the mayor and chooses not to run for reelection conditional on a particular challenger, as they are former mayors that were in power the previous period. We only need to adjust their budget constraint. We multiply the utility at the solution by the probability of the stochastic state variables of the corresponding point of the state space to recover the value function conditional on the challenger being elected ($V^i_{V,C}$) for each point of the deterministic state space. By integrating over all possible challengers the derived value functions for someone that enters the period as a mayor and chooses not to run, we compute the value function of not running for reelection (V^i_{FM}) at the beginning of the term.

We now know the voters' value functions conditional on the incumbent and challenger winning the election for each point of the deterministic state space. Then, for any incumbent, we can compute the probability that they win the election against any possible challenger, $p\left(S^e_t\right)$, using equation (9).

Given these election probabilities, we can compute the value function of a mayor who chooses to run for reelection (V^i_{RM}) for each point in the deterministic state space. We do this by multiplying the value function of a winning mayor (V^i_{WM}) by the probability of winning the election, plus the value function of a losing mayor (V^i_{LM}) multiplied by the probability of losing the election, and by integrating over possible challengers.

Given V^i_{RM} and V^i_{FM}, we can compute the value function of a current mayor at the beginning of the term V^i_M and the probability that an incumbent runs for reelection $p_R\left(S^M_t\right)$ for each point of the state space.

Lastly, given $p\left(S^e_t\right)$, $p_R\left(S^M_t\right)$, $V^i_{V,I}$ and $V^i_{V,C}$, we can compute $E\left[V^i_V\left(S^V_{t+1}\right)\middle| S^{V,C}_t\right]$ and $E\left[V^i_V\left(S^V_{t+1}\right)\middle| S^{V,I}_t, Q_{t-1}, \mathbb{1}^{au}_{t-1}, s^i_{t-1}\mathbb{1}^{au}_{t-1}\right]$, the expected value function of voters and former mayors at the beginning of $t+1$ outlined in the main text.

Now that we know all the relevant value functions for the discrete state space for term T, we can compute them for the entire state space by regressing the values of the value functions obtained for each point of the discretized state space on a polynomial of the discretized state variables.

Next, we move to term $T-1$ and repeat the steps using as inputs the known value functions for T, which give us the value functions for $T-1$. We can then proceed to previous terms until we reach $t=1$. Once we know the value functions for each term, we can simulate the data moving forward.

In the derivation of the voters' and former mayors' value functions, we must deal with a particular computational issue. In each term t, we only know the optimal amount of public good produced by mayors of age greater than or equal to t. Thus, we can only compute the value functions of voters and former mayors conditional on having a current mayor of age $\geq t$. But to solve our model, we need to compute their value functions conditional on having a current mayor of any age. To address this issue, we compute all value functions from $t = 1$ to $t = T$ twice. The first time we use the steps described above and save the optimal amount of public consumption for all possible mayors, including all possible ages. The second time, for each t' we compute the voters' and former mayors' value functions for mayors of all ages using the optimal amount of public consumption computed in the second iteration for $t \geq t'$ and the optimal amount of public consumption computed in the first iteration for $t < t'$. We then use the value functions obtained in the second iteration in our simulations. Although, in principle, this process could be repeated for more than two iterations, we found that doing so resulted in only negligible differences. Due to the time costs associated with each iteration, we decided to limit the process to just two iterations.

As mentioned earlier, in the first step of computing the value functions, we discretize the continuous state variables. We use the following grids: a ten-point grid for savings, $[-500000, -100000, 0, 50000, 100000, 250000, 500000, 750000, 1000000, 1250000]$; a three-point grid for ability, which is optimally derived given the mean and standard deviation of the ability distribution following Kennan (2006); a four-point grid for the fine distribution, which is also optimally computed following Kennan (2006) using the estimated mean and standard deviation; a three-point grid for public funds, $[669618.8, 1018890, 1803933]$, which correspond to the 25th, 50th, and 75th percentiles in the data; a one-point grid for private inputs, 2.553, which corresponds to the 50th percentile of the private input index in the data; a 6-point grid for the fraction stolen, $[0.0, 0.025, 0.05, 0.1, 0.2, 0.5]$; a three-point grid for log per-capital public consumption for each of the three municipality sizes, $[1.281881, 2.409436, 3.694281]$ for small municipalities, $[1.036066, 1.851396, 3.383907]$ for medium-sized municipalities, and $[2.00786, 3.720078, 5.450025]$ for large municipalities, which correspond to the 25th, 50th, and 75th percentiles in the data; a one-point grid at their mean for the running and wage shocks. In the second step, we approximate the value function for each value of the continuous deterministic state variables using non-parametric methods. In this step, it is crucial to approximate well the value function over savings. Regressing a polynomial in savings does not work well because a second order polynomial is not flexible enough and higher order polynomials create numerical issues as the value functions can be non-increasing in savings. For this reason, we have adopted the following non-parametric method. For each discrete value of savings, we compute the average value function over all the discrete values of the other continuous deterministic state variables. Then for each discrete value of savings and other variables, we compute residuals by subtracting the computed value function averages from the actual value function. We then regress the residuals on polynomials of the continuous deterministic state variables that are not savings. The value function is then approximated as a piece-wise linear function that joins the value function averages for savings with the addition of the estimated polynomials for the remaining deterministic continuous variables. For the latter, we use the following polynomial order: order two for ability and order one for public consumption and stealing.

Computationally, our model is quite demanding. To estimate it, we use Message Passing Interface (MPI) on a high-performance cluster that uses between 64 and 128 core processors.

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.

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 9: 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 10: The Effects of the Audits on the Candidate Pool

	Number of Candidates	Share Male	Share College	Average Age
	(1)	(2)	(3)	(4)
Eligible for the Audit Program	0.170	-0.022	0.007	-1.513
	(0.134)	(0.022)	(0.033)	(1.031)
Number of observations	22182	22182	22182	21792
R^2	0.53	0.43	0.56	0.44

Notes: This table presents output from a regression, in which we regress the dependent variable indicated by the columns on an indicator for whether the municipality is eligible for the audit program post 2003. The regression also controls for time and municipality fixed effects. The unit of observation is a municipality by election year pair. The sample is over 2000-2012 period and excludes capital cities. Robust standard errors clustered at the municipality level are reported in parentheses.

Table 11: The Effects of the CRA on the Candidate Pool

	Number of Candidates		Share Male		Share College		Average Age		Wealth	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Post CRA	0.126***	-0.168***	0.017***	0.006	-0.002	0.001	0.495***	0.420**	-4.189	-4.183
	(0.025)	(0.029)	(0.005)	(0.005)	(0.007)	(0.008)	(0.161)	(0.190)	(17.682)	(17.671)
Time trend	0.020***	-22.092***	-0.004***	-0.822***	0.008***	0.235	0.113***	-5.585	6.632**	-675.772
	(0.002)	(1.244)	(0.000)	(0.199)	(0.001)	(0.296)	(0.014)	(6.833)	(3.261)	(1682.674)
Time trend (sq)		0.006***		0.000***		-0.000		0.001		0.169
		(0.000)		(0.000)		(0.000)		(0.002)		(0.418)
Number of observations	27332	27332	27332	27332	27332	27332	27332	27332	21823	21823
R^2	0.52	0.53	0.38	0.38	0.50	0.50	0.40	0.40	0.28	0.28

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 12: 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	0.616	0.059	0.051	0.071	0.000
CRA	0.998	0.078	0.057	0.110	0.000
3-term	0.540	0.083	0.086	0.079	0.065
Wages	0.690	0.078	0.079	0.077	0.000
3-term+Audit	0.869	0.060	0.059	0.061	0.058
CRA+Audit	1.207	0.060	0.037	0.096	0.000
3-term+CRA+Audit	1.326	0.040	0.037	0.044	0.078
3-term+CRA+3^{rd}-Audit	1.297	0.061	0.057	0.066	0.000
Wage+Audit	0.879	0.065	0.063	0.068	0.000
25% Decrease in ability					
Audit	0.557	0.074	0.061	0.095	0.000
CRA	0.943	0.097	0.070	0.142	0.000
3-term	0.470	0.103	0.102	0.106	0.084
Wages	0.557	0.102	0.097	0.110	0.000
3-term+Audit	0.782	0.075	0.071	0.081	0.076
CRA+Audit	1.158	0.075	0.046	0.123	0.000
3-term+CRA+Audit	1.250	0.050	0.046	0.058	0.100
3-term+CRA+3^{rd}-Audit	1.183	0.076	0.070	0.086	0.000
Wage+Audit	0.759	0.085	0.078	0.098	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.

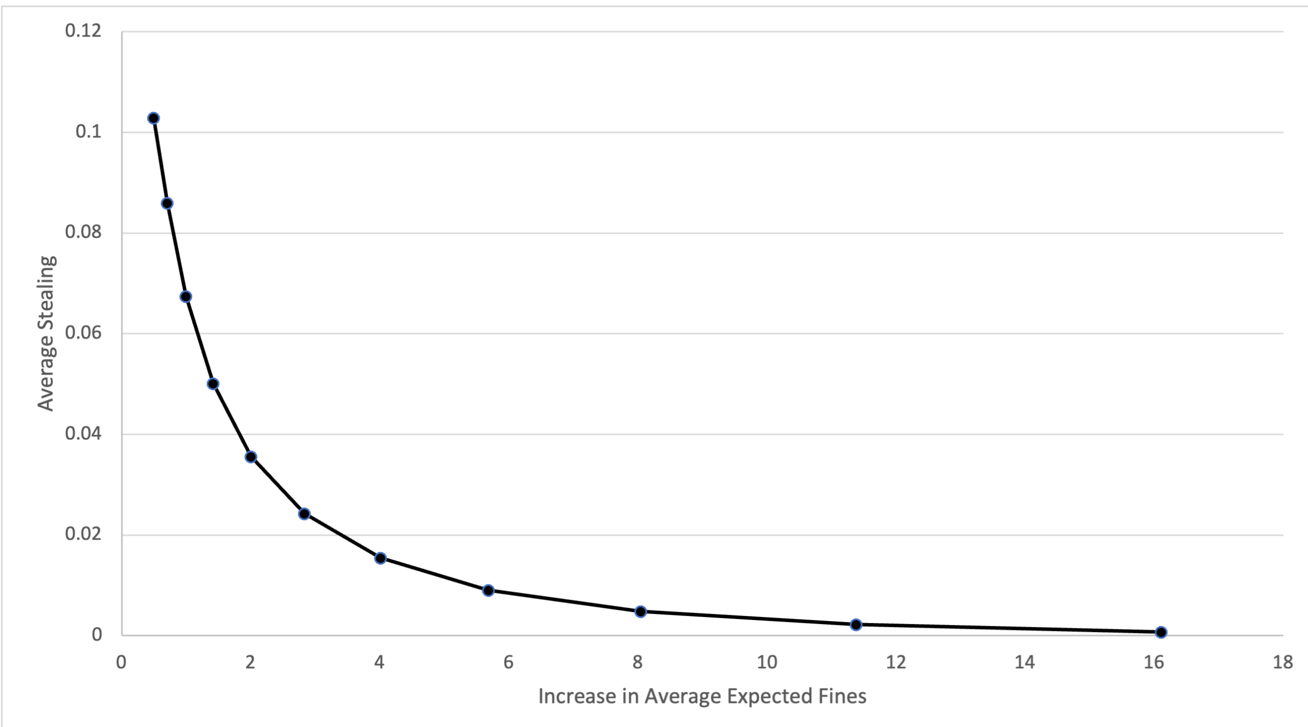


Figure 4: The Effects of Fines on Stealing

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

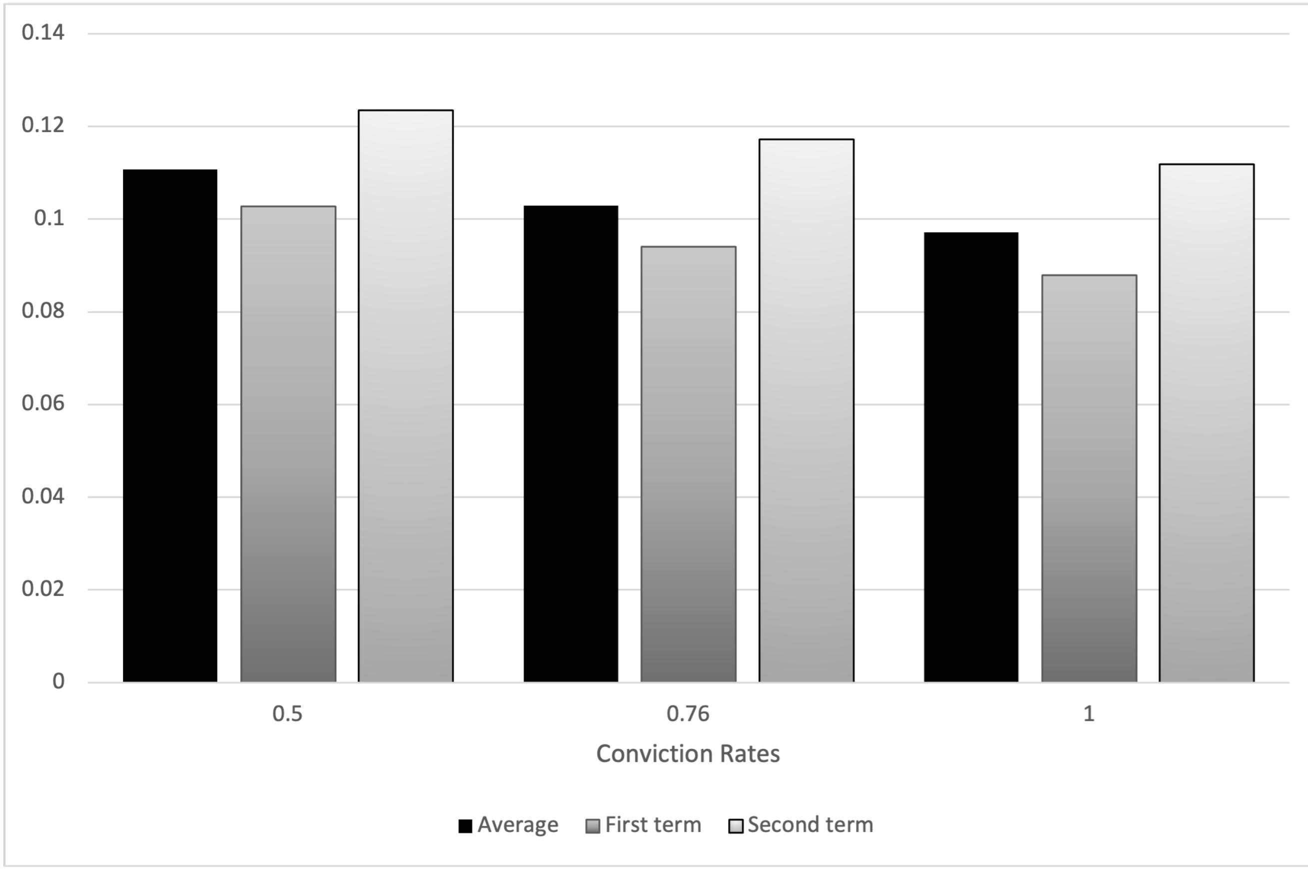


Figure 5: The Effects of Conviction Rates on Average Stealing

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

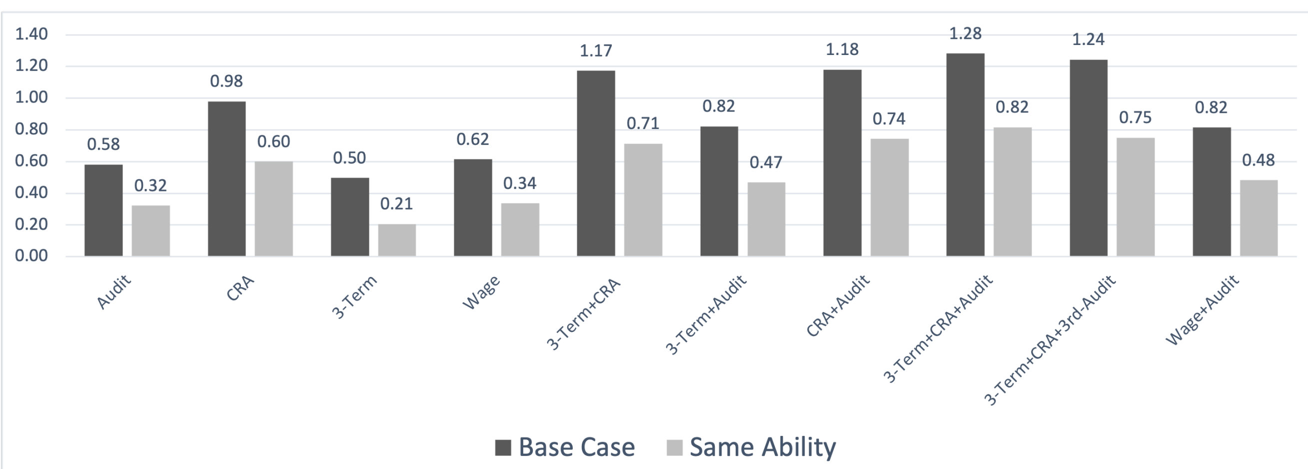


Figure 6: Willingness to Pay: The Effects of Ability

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