



INTERDISCIPLINARY CENTER
FOR ECONOMIC SCIENCE
GEORGE MASON UNIVERSITY

Accountability and Grand Corruption

Cesar Martinelli

November 2020

Discussion Paper

Interdisciplinary Center for Economic Science
4400 University Drive, MSN 1B2, Fairfax, VA 22030
Tel: +1-703-993-4850 Fax: +1-703-993-4851
ICES Website: <http://ices.gmu.edu>
ICES RePEc Archive Online at: <http://edirc.repec.org/data/icgmuus.html>

Accountability and Grand Corruption

Cesar Martinelli*

November 15, 2020

Abstract

We propose a model of political careers and electoral accountability, in an environment in which politicians may take bribes at different stages of their careers, and in which politicians' actions are only imperfectly observed by voters. We show that the expectation of promotion to higher office may motivate some politicians to behave worse at the latest stages of their careers, thus setting off a trade-off between providing incentives for good behavior at lower office and selecting better politicians for higher office. We also show that the optimal design of rewards for higher office has a simple bang-bang structure—optimal rewards focus either on stamping out corruption at lower office, or on improving selection at the higher office. If rewards are set optimally, a more intense competition for higher office benefits voters, but better quality of information about bribe-taking at lower levels does not unambiguously benefit voters.

1 Introduction

The possibility of bribe-taking and other forms of direct personal gain by presidents, prime ministers, and other authorities at top government positions poses grave difficulties for democratic governance. While bribe-taking by high ranking politicians is not unheard of in well-established, affluent democracies, it is bound to be a more pressing problem in newly established democracies for a variety of reasons, including poorly informed electorates, acquiescent or intimidated media, and the lack of well-established political parties and other institutions pre-selecting candidates. The problem is compounded when governments deal with large firms about potentially very lucrative contracts, like large infrastructure projects or mining concessions, creating the opportunity and the incentive for side-payments.

We provide a model depicting a representative democracy in which officials at different levels of government may be tempted to take bribes.

*Department of Economics, George Mason University, 4400 University Drive, MSN 1B2, Fairfax, VA 22030, USA. Email: cmarti33@gmu.edu.

I thank without implicating Roxana Barrantes and Elena Conterno, for bringing my attention to the topic of the paper, and John Duggan, Janneth Leyva, Susan Parker, Jaideep Roy, and Oriol Tejada, for invaluable conversations. Two anonymous referees and the associate editor provided detailed comments and suggestions.

We study the trade-off between selection at higher political offices and provision of incentives at lower ranked offices in an environment in which there are multiple contestants for the higher level position. We propose a model combining discrete strategies (accept a bribe or not) with a continuum of types (quality of politicians), which allows for a neat equilibrium characterization. We use the model to analyze the effects of changes in the quality of information and in the intensity of political competition on bribery and selection. We also characterize the optimal rewarding scheme. As discussed below, our model is consistent with findings from recent empirical literature on accountability and corruption.

In our model, a number of lower rank officials are offered bribes with some probability, in exchange for adopting policies that are not optimal for voters. Voters observe noisily those policies, and elect one of the lower rank officials to a higher rank policy position. After promotion, the higher rank politician is offered again bribes with positive probability, with the magnitude of the bribes being related to the responsibilities of the office. Politicians are motivated to reject bribes by a sense of duty, or warm-glow feeling, of adopting good policies, but they differ in the intensity of this motivation. Thus, voters attempt to promote the politician who is most likely to be duty-bound. Since politicians understand this, the perspective of promotion acts as an additional incentive to reject bribes at the lower rank.

The equilibrium of the model has a simple cutoff structure. The marginal type of politician who rejects bribes at the lower rank is the type for whom the present net loss of rejecting the bribe (that is, the monetary value of the bribe, minus the value of adopting a good policy for the politician, including the warm glow feeling) is equal to the future expected net gain caused by an increase in the probability of promotion. That is, because of the added incentive of promotion, the marginal politician who rejects bribes at the lower rank is someone who would take bribes if offered at the higher rank. Thus, the expected net gain of promotion is equal to the gains associated with office perks and likely bribes, minus the policy losses associated to the fact that the promotion of the marginal politician reduces the probability that good policies will be adopted. Though the net gain of promotion is not monotonic in type, the equilibrium is unique.

Some of the comparative statics results are intuitive. Increases in the office perks associated to the higher office and the importance of higher office lead to reductions in bribery at the lower rank at the expense of increases in bribery at the higher rank, due to worse selection. The effects of increases in the quality of information available to voters are more complicated. Better information reduces bribery at the lower rank, since it increases the impact of adopting good policies on the probability of promotion, but it has effects with opposite directions on bribery at the higher rank. On the one hand, it reduces the probability of promoting politicians who take bribes at the lower rank. On the other, it increases the probability of promoting politicians who do not take bribes at the lower rank but do so at the higher rank. Either effect can dominate. Finally, increases in the number of lower rank politicians (which is a proxy for the intensity of political competition), under fairly general conditions, increase bribery at the lower rank but reduce it at the upper rank.

We then turn to the question of the optimal (licit) reward for the higher rank politician. Since electoral incentives reduce bribery at the lower rank at the expense of increasing it at the higher rank, we could expect the optimal reward to decrease smoothly with the importance of the higher rank policy decisions. In fact, optimal office rewards have a simple bang-bang structure: for low levels of the importance of higher office, optimal rewards are set so as to stamp bribery out completely at the lower rank. At a certain critical value of the importance of higher office, optimal rewards fall discontinuously to zero. The reason for the bang-bang structure of optimal rewards is that the probability of selecting politicians who do not take bribes is convex in the cutoff employed by politicians—a simple and yet unexpected consequence of Bayes law in the context of the model.

We also consider the effects of changes in the parameters of the model on the welfare of voters when rewards are set optimally. We show that increases in the intensity of political competition cannot hurt, and possibly do help voters. Intuitively, any negative effect of political competition via reduced electoral incentives can be compensated if necessary by higher rewards to office. An equivalent, positive result about the quality of public information is not possible—even if rewards for politicians are set optimally, voters can be hurt by improvements in the quality of information. Negative effects of information quality via increased electoral incentives cannot be necessarily compensated by lower rewards to office because those may be already set equal to zero at a corner solution. This highlights an essential distortion introduced by bribery of top officials: “bad” politicians are perversely motivated for promotion by the expectation of future bribes.

Finally, we consider several variations of the model. In particular, we examine a version of the model in which incumbent politicians are not motivated by promotion but by the possibility of being reelected once, as in the standard finite-horizon accountability model. We show that in this case better quality of information cannot hurt voters. That is, if incumbent politicians are not primarily competing with each other but with unknown challengers, the importance of negative effects of better information selection is reduced. Moreover, in this version of the model, maximal rewards are optimal if the quality of information is relatively low, if there are more corruptible politicians, if the second period is relatively more important, and if (second period) bribes are relatively rare.

Our framework is consistent with findings of the recent empirical literature on accountability and corruption. Due to data availability, this literature has focused on local politicians. In their seminal work on random audits in Brazil, Ferraz and Finan (2008) find that better public information reduces corruption among mayors facing reelection incentives. Ferraz and Finan (2011) find that corruption is lower for mayors in their first period than in their second (and last, due to term limits) period. They also discuss whether local politicians respond to the probability of promotion, and conjecture that this effect is not important due to the low probability of running for higher office for mayors. That is, the situation they study resembles the version of our model in which incumbents do not compete with each other. In the same vein, de Janvry et al. (2012)

find evidence that the performance of a decentralized social program in Brazil improves when local authorities face reelection possibilities, and that good performance enhances reelection probabilities. Similarly, Bobonis et al. (2016) show that timely audits reduce corruption among local authorities in Puerto Rico in concurrent elections but not necessarily in subsequent elections. In line with our equilibrium construction, Bobonis et al. describe their evidence as suggestive that the better quality of information provided by audits “enable voters to select responsive but corruptible politicians to office.”

As pointed out by Fisman and Golden (2017), echoing Treisman (2007), evidence about the effect of political institutions on corruption “is ‘fragile’ at best, and more often contradictory.” Our framework highlights that the effects of different institutional features may depend in the level of government and the information environment. On the issue of politicians’ reward, for instance, Gagliarducci and Nannicini (2013), in a careful study of mayors in Italy, find that higher compensation brings better candidates to the politicians’ pool, but does not seem to improve the behavior of politicians once in a political career. In our framework, rewards are more effective when information is scant, which may resemble more the initial selection of politicians running for office.

Our model highlights the connection between corruption at different levels of government via the selection of politicians, which has received relatively less attention from the empirical literature. In a somewhat similar spirit to our framework, although in the context of a different model, Myerson (2006) has emphasized the advantages of federalism and decentralization for promoting the emergence of national-level politicians with better reputations. As corroborating evidence, Myerson cites Boix’s (2003) finding that federalism improves the chances for democratic survival.

Our work is linked to the formal literature on accountability, including Austen-Smith and Banks (1989), Banks and Sundaram (1993), Fearon (1999), Persson and Tabellini (2000), Besley (2006), Ashworth (2005, 2012), Acemoglu et al. (2013), Ashworth et al. (2017), and Duggan and Martinelli (2020); see the surveys by Ashworth (2012) and Duggan and Martinelli (2017). We deviate from this literature by contemplating a contest between lower rank politicians rather than an election between an incumbent and an (unknown) challenger, and by considering a setting in which actions are discrete (taking or not a bribe) but there is a continuum of types, rather than the reverse. This allows us to sidestep the technical difficulties associated with mixed strategies in a non-convex setting, as discussed by Duggan and Martinelli (2020). This is not merely a technical concern; our analysis in terms of the marginal politician would not be possible without an explicit consideration of a continuum of types.

Our work is also related to theoretical research on politician selection and political careers, including among others Caselli and Morelli (2004), Besley (2005), Mattozzi and Merlo (2008), and Galasso and Nannicini (2011). This literature puts emphasis on selecting for ability or “valence.” Instead, our model puts emphasis in the trade-off between selection at higher office positions and good incentives at lower office. Although for clarity we present our model of politician selection as a two-period frame-

work, it could be reinterpreted, like other models of careers, as an infinite horizon framework in which in every period there is a cohort of lower rank politicians and a higher rank politician belonging to a previous cohort.

The remainder of the paper is organized as follows. In section 2, we describe the model of careers and corruption. In section 3, we describe our signaling game and equilibrium notion. In section 4, we present existence and uniqueness results. We discuss comparative statics in section 5. In section 6, we characterize the optimal reward for politicians. In section 7, we explore different variations in the assumptions and alternative interpretations of the model. We gather concluding remarks in section 8. Formal proofs are collected in the appendix.

2 A model of careers and corruption

We consider a two-period model of political careers. In the first period, there are n politicians occupying lower rank positions, $i = 1, \dots, n$. Each of these politicians makes a policy decision $x_{i1} \in \{0, 1\}$. Politicians' decisions are not directly observable by voters, who instead observe for each politician i a noisy signal $s_i \in \{0, 1\}$ satisfying $\Pr(s_i = 1|x_{i1} = 1) = 1$ and $\Pr(s_i = 1|x_{i1} = 0) = 1 - q \in (0, 1)$. That is, $s_i = 0$ is conclusive evidence that the policy decision was $x_{i1} = 0$, and signals are independent of each other. We can think of q as the probability with which the policy decision 0 is detected by voters. After observing their signals, voters decide whom of the politicians to advance to a (single) higher rank position. If politician i is advanced to the higher rank, the politician makes a second-period policy decision, $x_2 \in \{0, 1\}$, and time is over.

Payoffs of voters depend on policy decisions, and are given by

$$(1 - \delta)\frac{1}{n} \sum_i x_{i1} + \delta x_2,$$

where $\delta \in (0, 1)$ is a parameter reflecting the importance of the higher rank position. We can think of the n politicians in the lower rank as located in districts with different voters, so that the expression above represent the average voter welfare. In this interpretation, all voters are equally affected by the higher rank policy decisions.

If politician i enacts the policy 1 in period 1 (respectively, in period 2), the politician experiences a ‘warm glow’ satisfaction of having done the right thing equal to $(1 - \delta)\theta_i$ (respectively, $\delta\theta_i$), where θ_i is distributed uniformly on the interval $[0, 1]$ and is observed privately by the politician at the beginning of period 1. We refer to θ_i as the politician’s type. Before making a policy decision in period $t = 1, 2$, a politician receives privately a bribe offer with probability $\lambda \in (0, 1)$. If a politician is offered a bribe in period 1 (respectively, in period 2), the politician obtains a monetary payment of $(1 - \delta)B$ (respectively, δB) if the politician enacts the policy 1, and a payment of 0 otherwise, where $B > 0$. Probabilities of receiving a bribe offer are independent across periods and across politicians. If a politician is elected for the higher rank, the politician obtains a rent of $\delta R > 0$ regardless of type. Finally, the politician payoff depends as well on her policy decision in the lower rank, and from whatever decision is taken

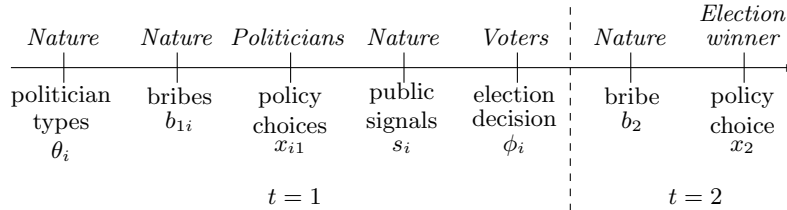


Figure 1: Political careers and corruption

in the higher rank. In the multi-district interpretation, the politician lives in her own district.

Summarizing, the payoff of politician i is given by

$$(1 - \delta)(x_{i1} + (1 - x_{i1})b_{i1}B + x_{i1}\theta_i) + \delta(x_2 + \phi_i(R + (1 - x_2)b_2B + x_2\theta_i)),$$

where $b_{i1} = 1$ if the politician is offered a bribe in period 1 and zero otherwise, $\phi_i = 1$ if the politician is promoted and zero otherwise, and $b_2 = 1$ if the politician is offered a bribe in period 2 and zero otherwise. The first and second line in the expression above are, respectively, the first and second period payoff of politician i . Note that politicians are citizens as well, and bribes and rewards are obtained in addition to the policy payoff. Note also that we assume that potential bribes and warm-glow feelings scale up with the importance of the position. We discuss in section 7 variations in some of these assumptions.

Figure 1 summarizes the model. Politicians' types as well as bribe opportunities are exogenously determined at the beginning of the first period. Politicians then undertake first period policy decisions, and public signals are released to voters. Voters decide whom among the politicians to promote. Bribe opportunity for the election winner is exogenously determined at the beginning of the second period. The election winner then undertake the second period policy decision, and payoffs are realized.

3 Signaling equilibrium

Next, we define strategies, beliefs, and equilibrium for the extensive game of incomplete information played by politicians and voters. The information set available to politician i before making a policy choice¹ in period 1 is characterized by the politician's own type and bribe opportunity

$$(\theta_i, b_{i1}) \in [0, 1] \times \{0, 1\} \equiv \mathcal{I}_1.$$

If politician i is promoted, the information set available before making a policy choice in period 2 is characterized the politician type, bribe opportunity, and policy choice in the first period, the public signal about

¹That is, the set of histories that are indistinguishable for the politician.

the politician, the vector of public signals for other politicians, and the bribe opportunity in the second period, that is $(\theta_i, b_{i1}, x_{i1}, s_i, (s_j)_{j \neq i}, b_2) \in [0, 1] \times \{0, 1\}^{n+3} \equiv \mathcal{I}_2$.

A (behavioral) *strategy for politician i* is a pair of mappings $\sigma_i = (\sigma_{i1}, \sigma_{i2})$,

$$\sigma_{i1} : \mathcal{I}_1 \rightarrow [0, 1] \quad \text{and} \quad \sigma_{i2} : \mathcal{I}_2 \rightarrow [0, 1]$$

specifying a probability of adopting the policy decision 1 as a function of the information held by the politician in period 1 and in period 2.²

We treat voters as a unitary player. A *strategy for voters* is a mapping

$$\nu : \{0, 1\}^n \rightarrow \Delta^{n-1} \in \mathfrak{R}^n$$

from the set of possible vectors of public signals to the $n - 1$ -dimensional simplex, where $\nu(s_1, \dots, s_n) = (\nu_i(s_1, \dots, s_n))$ specifies for every vector of public signals the probability $\nu_i(s_1, \dots, s_n)$ that i will be promoted for $i = 1, \dots, n$. A *belief system for voters* is an n -tuple of mappings $\beta = (\beta_1, \dots, \beta_n)$,

$$\beta_i : \{0, 1\} \rightarrow \mathcal{F},$$

where \mathcal{F} is the family of distribution functions with support $[0, 1]$, and $\beta_i(s_i)$ specifies a distribution function over the set of types for politician i .³⁴

A (perfect Bayesian) *equilibrium* is a profile of strategies for politicians (σ_i) , a strategy for voters ν , and a belief system for voters β such that

- (i) The voters' strategy ν is optimal given voters' beliefs β ,
- (ii) Voters' beliefs β are consistent with politicians' strategies; i.e. for each politician i , β_i is derived from prior beliefs and the politician i 's strategy using Bayes' rule for every signal with positive probability given the politician's strategy, and
- (iii) Politicians' strategies (σ_i) are optimal given other politicians' strategies and the voters' strategy.

Under our payoff assumptions, a promoted politician will adopt policy decision 1 if and only if either not being offered a bribe, or if offered a bribe, but having a "warm-glow" reward strong enough to induce her to relinquish the bribe, that is if $\theta_i + 1 > B$. We assume throughout that

$$1 < B < 2. \tag{A1}$$

This condition guarantees that politicians with low enough types adopt bad policies in the second period if offered a bribe, while politicians with high enough types do not.

²As customary, we assume that strategies are measurable functions. In what follows, we treat strategies for politicians that differ on a set of types of measure zero as identical.

³That is, $\beta_i(s_i)[z] : [0, 1] \rightarrow [0, 1]$ is a weakly increasing, right-continuous function satisfying $\beta_i(s_i)[0] = 0$ and $\beta_i(s_i)[1] = 1$.

⁴For simplicity, we assume that beliefs about politicians depend only on their own signals and not on other politicians' signals. Given the requirement of consistency (see below), this only matters for profiles of signals that have zero probability in equilibrium. (This is the 'no-signaling-what-you-don't-know' condition on beliefs of Fudenberg and Tirole (1991).)

Under assumption A1, in equilibrium, voters will promote the politician whom they believe is more likely to have a type above $B - 1$. That is, for every profile of public signals s_1, \dots, s_n , voters will promote i only if $\beta_i(s_i)[B - 1] \leq \min_{j \neq i} \beta_j(s_j)[B - 1]$. The following condition on equilibrium requires voters to interpret a good signal about a politician as favorable evidence regarding the politician's type. We say that an equilibrium is a *signaling equilibrium* if, in addition to (i)–(iii), the following condition holds:⁵

$$(iv) \quad \beta_i(0)[B - 1] > \beta_i(1)[B - 1] \quad \text{for } i = 1, \dots, n.$$

We say that a strategy for voters is *neutral* if for all signal profiles (s_1, \dots, s_n) ,

$$\beta_i(s_i)[B - 1] = \beta_j(s_j)[B - 1] \quad \text{implies} \quad \nu_i(s_1, \dots, s_n) = \nu_j(s_1, \dots, s_n).$$

We say that a signaling equilibrium is neutral if voters play a neutral strategy and if $s_i = 0$ has zero probability for all i , then $\beta_i(0)[B - 1] = \beta_j(0)[B - 1]$ for all i, j .⁶ As it turns out, equilibria satisfying this equal-treatment property have a very simple structure.

4 Equilibrium analysis

We say that a strategy σ_i for politician i is a *cutoff strategy* if there is some $\bar{\theta}_i \in [0, B - 1]$ such that

$$\sigma_{i1}(\theta_i, b_{i1}) = \begin{cases} 0 & \text{if } \theta_i < \bar{\theta}_i \text{ and } b_{i1} = 1 \\ 1 & \text{if } \theta_i \geq \bar{\theta}_i \text{ or } b_{i1} = 0 \end{cases},$$

and

$$\sigma_{i2} = \begin{cases} 0 & \text{if } \theta_i < B - 1 \text{ and } b_{i2} = 1 \\ 1 & \text{if } \theta_i \geq B - 1 \text{ or } b_{i2} = 0 \end{cases}.$$

Focusing on cutoff strategies for politicians is without loss of generality:⁷

Lemma 1. *In every signaling equilibrium, politicians play cutoff strategies.*

(See the proof for this and other results in the appendix.)

We say that a strategy profile (σ_i) for politicians is *symmetric* if $\sigma_{j1} = \sigma_{k1}$ for all $j, k = 1, \dots, n$.

We refer to the following strategy for voters as the *monotonic* strategy:

$$\nu_i(s_1, \dots, s_n) = \begin{cases} 1/|\{j : s_j = 1\}| & \text{if } s_i = 1 \\ 1/n & \text{if } s_j = 0 \text{ for all } j = 1, \dots, n \\ 0 & \text{otherwise} \end{cases}.$$

We have:

⁵Non signaling equilibria are possible. For instance, if politicians care enough about the next period payoff, there is a perverse “counter-signaling” equilibrium in which all politicians adopt bad policies in the first period and voters, who are indifferent, reward politicians with bad public signals. That equilibrium violates condition (iv).

⁶This restriction on (extreme) out-of-equilibrium beliefs allows us to describe equilibrium compactly when all politicians adopt good policies with probability one (see footnote 14).

⁷See footnote 2.

Lemma 2. *In every neutral signaling equilibrium, politicians play a symmetric cutoff strategy profile, and voters play the monotonic strategy.*

Lemma 2 simplifies the search for an equilibrium. In particular, let

$$\begin{aligned}
U(\theta, \bar{\theta}) &= -(1 - \delta)(B - 1 - \theta) \\
&\quad + \delta q \left(\frac{1 - (\lambda q \bar{\theta})^{n-1}}{n(1 - \lambda q \bar{\theta})} \right) \times [R + (1 - \lambda)\theta + \lambda B] \\
&\quad - \delta q \left(\frac{1 - (\lambda q \bar{\theta})^{n-1}}{n(1 - \lambda q \bar{\theta})} - \frac{n-1}{n} (\lambda q \bar{\theta})^{n-1} \right) \left[\frac{\lambda(2 - B)}{1 - \lambda q \bar{\theta}} \right]. \quad (1)
\end{aligned}$$

As shown in the appendix, $U(\theta, \bar{\theta})$ is the expected payoff gain for a politician of adopting the good policy instead of the bad policy in period 1 when the politician's type is $\theta \in [0, B - 1]$, every other politician is playing a cutoff strategy with cutoff $\bar{\theta} \in [0, B - 1]$, and voters play the monotonic strategy.

The first line in the right-hand side of equation 1 is the first period payoff difference between implementing the good policy, and enjoying the policy payoff 1 and the warm-glow satisfaction θ , or collecting the bribe B . The second and the third line in the right-hand side of equation 1 represent the anticipated second period net payoff gain of implementing the good rather than the bad policy in period 1. The parameter q at the beginning of the second and third line conveys the increase in the probability of getting a good signal by implementing a good rather than a bad policy, which is the link connecting decisions in the first period with second period payoffs.

The term in parenthesis in the second line is the increase in the probability of promotion in case of a good rather than a bad signal, that is $p_g(\bar{\theta}) - p_b(\bar{\theta})$, where

$$p_g(\bar{\theta}) = \sum_{\omega=0}^{n-1} \frac{1}{n-\omega} \binom{n-1}{\omega} (\lambda q \bar{\theta})^\omega (1 - \lambda q \bar{\theta})^{n-1-\omega} = \frac{1 - (\lambda q \bar{\theta})^n}{n(1 - \lambda q \bar{\theta})}$$

and

$$p_b(\bar{\theta}) = \frac{1}{n} (\lambda q \bar{\theta})^{n-1}$$

are obtained from voters' monotonic strategy and (other) politicians' cutoff strategy, as detailed in the appendix.

The term in square brackets in the second line represents the non-policy related gain of promotion, that is receiving the office rent R plus the warm glow satisfaction with probability $1 - \lambda$ and the bribe with probability λ .

The term in parenthesis in the third line is equal to the increase in the probability of promotion in case of a good rather than a bad signal against a rival or rivals with good signals. The term in square brackets in the third line represents the expected *loss* in policy payoff in case of promotion against a rival with a good signal. The reason there is a policy-related loss associated with promotion is that the marginal politician knows that she will adopt a good policy only if not offered a bribe, while there is some probability that a rival with a good signal in fact has a type above

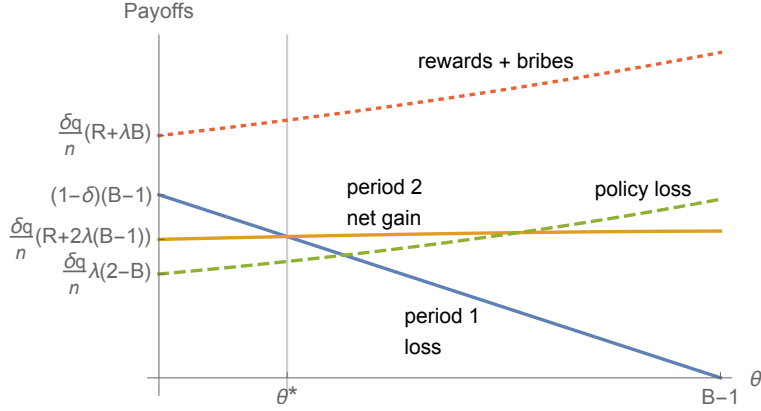


Figure 2: Equilibrium cutoff

$B-1$. That probability is $(2-B)/(1-\lambda q\bar{\theta})$, and is obtained from Bayesian updating. That is, the marginal politician knows that she is worse than the average politician with a good signal.

The difference between the term in parenthesis in the second and the third line represents the event in which every other politician has a bad signal, which happens with probability $(\lambda q\bar{\theta})^{n-1}$. In that case, the increase in the probability of winning the election when having a good rather than a bad signal is $1 - 1/n = (n-1)/n$. There is no policy-related loss in that case, since the marginal politician can infer that all the rivals are as bad as herself.

Note that $U(\theta, \bar{\theta})$ is continuous and strictly increasing in θ , and for every $0 \leq \bar{\theta} \leq B-1$, $U(B-1, \bar{\theta}) > 0$. Hence, the politician strategy is a best response to other politicians adopting the cutoff strategy $\bar{\theta}$ if and only if it is a cutoff strategy with cutoff 0 if $U(0, \bar{\theta}) \geq 0$ and otherwise with cutoff given by the unique solution $\hat{\theta} \in (0, B-1)$ to $U(\hat{\theta}, \bar{\theta}) = 0$.

Let

$$G(\theta) \equiv U(\theta, \theta).$$

From the preceding argument, there is a neutral signaling equilibrium with cutoff θ^* if and only if $\theta^* \in [0, B-1)$ solves

$$G(\theta^*) \geq 0 \text{ with equality if } \theta^* > 0. \quad (2)$$

We show in the appendix that equation 2 has a unique solution. We illustrate the equilibrium construction in figure 2. The equilibrium cutoff θ^* is determined when the first period loss due to giving up the bribe is equal to the expected second period net gain due to the increased probability of promotion after adopting a good policy. The first period loss, $(1-\delta)(B-1-\theta)$, is declining linearly with the intensity of warm-glow feelings for the marginal politician. The second period net gain is subject to two forces. On the one hand, the probability of promotion is increasing in the cutoff adopted by other politicians, in turn increasing the lure of office earnings and expected bribes (the second line in equation 1). On the

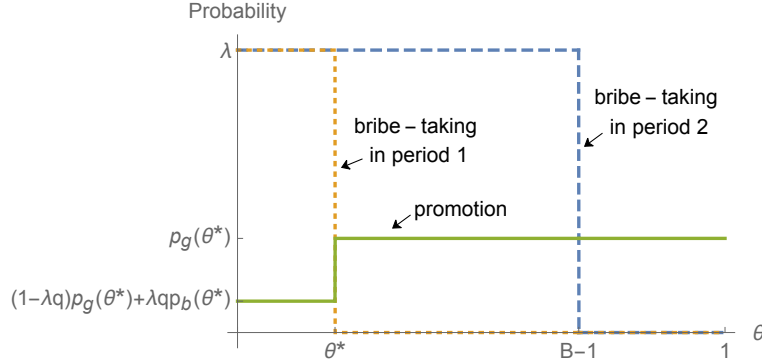


Figure 3: Equilibrium partition

other hand, as the cutoff adopted by other politicians rises, the quality of a rival with a good signal increases, leading to a policy loss associated to victory (the third line in equation 1). The second period gain curve is not necessarily monotonic, but, as shown in the proof of the theorem, it can only cross the first period loss curve from below.

We have:

Theorem 1. *There is a neutral signaling equilibrium and it is unique. In equilibrium, politicians play a symmetric cutoff strategy profile with cutoff θ^* given by the solution to equation 2.*

Ignoring knife-edge cases, equilibrium determines a partition of politician types into those who take bribes in either period, if offered ($0 \leq \theta < \theta^*$), those who do not take bribes in period 1 but take them in period 2, if offered ($\theta^* < \theta < B-1$), and those who do not take bribes ($\theta > B-1$). The probabilities of bribe-taking and promotion for the different politician types are illustrated in figure 3.

The equilibrium probabilities of bribery in period 1 and period 2 from the viewpoint of society can be calculated using the equilibrium partition, and are, respectively,

$$\Pr(x_{i1} = 0) = \lambda\theta^* \text{ and}$$

$$\Pr(x_2 = 0) = \lambda \left(1 - \left[(1 - (\lambda q \theta^*)^n) \frac{2-B}{1 - \lambda q \theta^*} \right] \right). \quad (3)$$

In particular, the term in brackets is the probability of promoting a politician who does not take bribes in the second period, and it is equal to the probability of promoting a politician with a good signal, $1 - (\lambda q \theta^*)^n$, multiplied by the probability that a politician with a good signal does not take bribes in the second period, $(2-B)/(1 - \lambda q \theta^*)$.

Bribery can be more frequent in the first or in the second period, as illustrated by examples in the next section. In the example calculated⁸ in order to draw figure 2, we set $n = 5$, $\delta = 0.6$, $B = 1.3$, $\lambda = 0.9$, $q = 0.9$,

⁸All computations are performed using Wolfram Mathematica 12.

and $R = 0.3$. The equilibrium cutoff is about 0.0685. That is, 6.85% of politicians take bribes whenever they can, 23.15% of politicians do not take bribes in the lower rank but intend to take bribes if promoted, and 70% of politicians do not take bribes in any circumstance. The probability of observing bribery is about 6.16% in the first period and about 23.3% in the second period. Intuitively, there is a better selection of politicians in the second period, but some politicians takes bribes more aggressively after promotion, since electoral incentives cease to operate.

5 Comparative statics

Some of the comparative statics results of the model accord to intuition. Increases in the office perks associated to the higher office and the importance of higher office reduce the incentive to take bribes at the lower rank, but worsen selection. We focus in this section on the effects of changes in the quality of information and in the intensity of political competition on politician selection and bribery.

If the quality of the public signals increase, the second period gain curve depicted in figure 2 moves upward, but the first period loss curve remains constant. As a result, the equilibrium cutoff must decrease. Thus, bribery in the first period decreases. The effect of the quality of information on bribery in the second period is more involved. On the one hand, better scrutiny reduces the frequency of promotion of unreformed politicians (those who adopt bad policies whenever offered a bribe). On the other, because of changes in the equilibrium cutoff, better scrutiny increases the frequency of promotion of dissembling politicians (those who adopt good policies in the lower rank but bad policies in the higher rank if offered a bribe). Either effect can predominate.

As an illustration, let $n = 5$, $\delta = 0.6$, $B = 1.3$, $\lambda = 0.9$, and $R = 0.3$, and vary q from 0 to 1, as depicted in figure 4. In this example, improvements in the quality of information provide incentives for better behavior at the lower rank, but have a non monotonic effect in the selection of the politician who gets promoted to the upper rank.

We summarize this discussion as follows:

Corollary 1. *If $\theta^* > 0$ in the initial situation, an increase in the quality of the public signal q reduces bribery in the first period and has ambiguous effects on bribery in the second period.*

We also look at changes in the number of competitors n . We can interpret this parameter as representing the degree of political competition. Intuitively, as the number of competitors increases, the probability of winning by obtaining a good signal is reduced, which reduces the incentive to adopt good policies in period 1.⁹ Selection of the higher rank politician improves, though, so the probability of observing bribes in the second period declines. We have:

⁹There is a possible countervailing effect. The increase in the probability of winning because of having a good rather than a bad signal when everyone else has a bad signal is $(n - 1)/n$, and increases with n . The qualification in the statement of corollary 2 below makes sure that the probability that everybody else has a bad signal is small enough for this effect not to predominate.

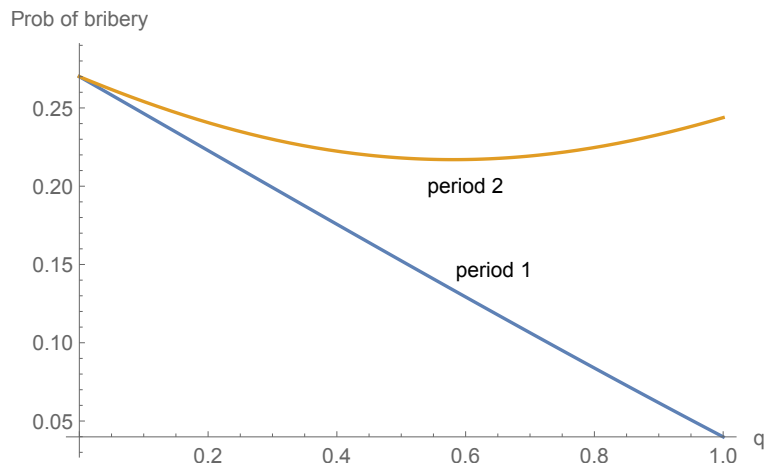


Figure 4: Quality of information and bribery

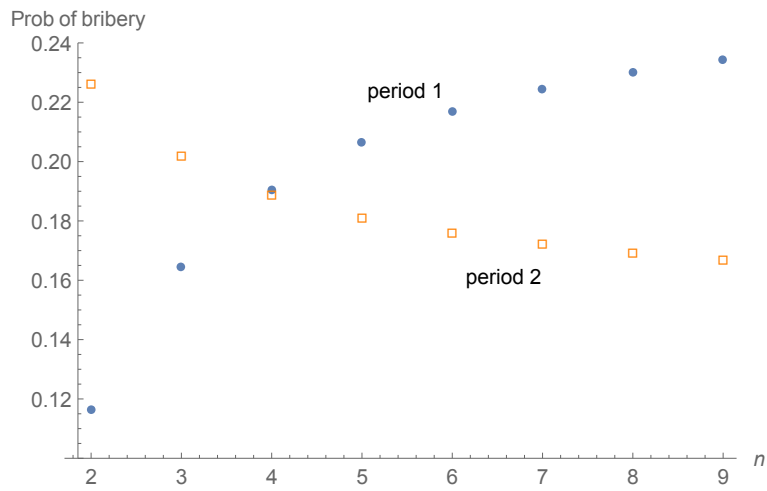


Figure 5: Number of lower rank politicians and bribery

Corollary 2. *If $\theta^* > 0$ and $1 - 1/n^2 \geq \lambda q(B - 1)$ in the initial situation, then an increase in the number of lower rank officers increases bribery in the first period, and reduces bribery in the second period.*

Figure 5 illustrates the effects of changes in the number of lower rank politicians on the probability of bribery in either period. We set $\delta = 0.4$, $q = 0.6$, $B = 1.3$, $\lambda = 0.9$, and $R = 0.3$, and vary n from 2 to 9. As the number of competing lower rank politicians increases, the probability of bribery in the first period goes from about 12% to 23%, while the probability of bribery in the second period goes from about 23% to 17%. As n increases unboundedly, the limits are respectively 0.27% and 0.15%.

The expected welfare of voters can go up or down with the number of lower rank politicians, depending on the importance of the high rank office. For instance, in the example in figure 5, the expected welfare of voters declines with the number of lower rank politicians, going from about 0.84 to about 0.79 if the number of lower rank politicians vying for higher office goes from 2 to 9.

6 Optimal reward

In this section, we study the optimal choice of the high ranked politician's reward from the viewpoint of voter. In terms of interpretation, the politician's reward includes the salary as well as the value the politician puts in the prestige, power, and perks associated to office and that are not linked directly to policy decisions. Following Rogoff (1990), it is common to refer to the reward of holding office per se as ego-rents. To some extent, this reward can be manipulated by the choice of salaries associated to different positions.¹⁰

To find the office reward R that maximizes the expected welfare of voters, let $\Theta(R)$ to be the equilibrium cutoff as a function of $R \geq 0$. Using the specification of payoffs in section 2 and the equilibrium probabilities of bribery, we can write voters' expected welfare as a function of R as $W(R) \equiv \tilde{W}(\Theta(R))$, where

$$\tilde{W}(\theta) \equiv 1 - (1 - \delta)\lambda\theta - \delta\lambda \left(1 - (1 - (\lambda q\theta)^n) \frac{2 - B}{1 - \lambda q\theta} \right).$$

We define the *optimal reward* R^* for the higher rank politician as

$$R^* = \inf \arg \max_{R \geq 0} W(R).$$

We take the infimum in the definition of optimal reward to convey the idea that rewards to politicians may have costs that are lexicographically less important than the potential gains in voter welfare.

¹⁰The line between the licit value of holding office and what constitutes bribery may be thin. A politician may benefit from holding office because patronizing the politician's private businesses is a way to curry favor, because of access to privileged information, etc.

Let

$$\begin{aligned}
F(\theta, R) &= -(1 - \delta)(B - 1 - \theta) \\
&+ \delta q \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)} \right) \times [R + (1 - \lambda)\theta + \lambda B] \\
&- \delta q \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)} - \frac{n-1}{n} (\lambda q \theta)^{n-1} \right) \left[\frac{\lambda(2 - B)}{1 - \lambda q \theta} \right]. \quad (4)
\end{aligned}$$

F represents the net gain from adopting a good policy in period 1 for a politician with type θ , after being offered a bribe, when every other politician is using a cutoff θ , and the rewards for office are set at R . The expression for F is obtained from equation 1.

We can show

Lemma 3. $\Theta(R)$ is equal to zero if

$$R \geq \bar{R} \equiv \max \left\{ (B - 1) \left[\frac{n(1 - \delta)}{\delta q} - 2\lambda \right], 0 \right\},$$

and otherwise $\Theta(R)$ is given implicitly by $F(\Theta(R), R) = 0$. Moreover, $\Theta(R)$ is continuous and it satisfies $\Theta(0) < B - 1$ and $\Theta'(R) < 0$ for $0 < R < \bar{R}$.

Lemma 3 provides an upper bound to the optimal reward. In particular, if the expression in brackets in the statement of the lemma is not positive, the optimal reward for the higher rank politician is zero.

To understand lemma 3, note that every politician adopts the good policy in period 1 if a politician with warm-glow feelings equal to zero gains more from adopting the good rather than the bad policy, when every other politician adopts the good policy with probability 1. For such politician, the first period loss of relinquishing the bribe is $(1 - \delta)(B - 1)$, while the second period gain of adopting good policies is equal to the importance of the second period (δ) times the probability of obtaining a good rather than a bad signal (q), times the increase in the probability of winning ($1/n - 0$), times the payoff in case of winning ($R + \lambda B + 1 - \lambda$) minus the payoff in case of losing ($1 - \lambda(B - 1)$). The bound \bar{R} is obtained by equating first period losses and second period gains.

Equipped with lemma 3, we can characterize the optimal reward. The optimal reward for politicians has a simple bang-bang structure: it is equal to 0 if voters care more about selecting a higher rank politician, and it is equal to $\bar{R} > 0$ if voters care more about providing incentives for lower rank politicians. We have:

Theorem 2. $R^* \in \{0, \bar{R}\}$, with $R^* = \bar{R} > 0$ if and only if $\delta < \bar{\delta}$, where $\bar{\delta} \in (0, 1)$ is the unique solution to

$$\frac{1 - \bar{\delta}}{\bar{\delta}} = \lambda q \frac{1 - (\lambda q \Theta(0))^{n-1}}{1 - \lambda q \Theta(0)} (2 - B). \quad (5)$$

To provide an intuition, suppose that in the initial situation the reward to politicians is below \bar{R} , so that the cutoff is $\tilde{\theta}$. By increasing the reward of politicians to \bar{R} , the cutoff is driven to zero. The welfare gain in period 1 due to the reduction in bribery is

$$(1 - \delta)\lambda\tilde{\theta},$$

and the welfare loss in period 2 due to the reduction in the probability of promoting an incorruptible politician is

$$\delta\lambda \left[\frac{1 - (\lambda q \tilde{\theta})^n}{1 - \lambda q \tilde{\theta}} (2 - B) - (2 - B) \right],$$

where the first term within brackets is the probability of promoting an incorruptible politician when the cutoff is $\tilde{\theta}$ (just an application of Bayes law) and the second term is the probability of promoting an incorruptible politician when the cutoff is zero. Thus, it is convenient for voters to drive down bribery in period 1 to zero if

$$(1 - \delta)\lambda\tilde{\theta} > \delta\lambda \left[\frac{1 - (\lambda q \tilde{\theta})^n}{1 - \lambda q \tilde{\theta}} - 1 \right] (2 - B),$$

or equivalently, if

$$1 - \delta > \delta\lambda q \frac{1 - (\lambda q \tilde{\theta})^{n-1}}{1 - \lambda q \tilde{\theta}} (2 - B).$$

The right-hand side in the inequality above is strictly increasing in $\tilde{\theta}$, so that if it pays for voters to marginally reduce the cutoff by introducing positive rewards for promotion when $\tilde{\theta} = \Theta(0)$, then it makes sense for them to increase rewards until the cutoff is driven to zero.

In the proof of theorem 2, we show that equation 5 has a unique solution, and that it is optimal for voters to provide maximal rewards ($R = \bar{R}$) if the second period is not important enough (i.e. low δ). This is not trivial since $\Theta(0)$ depends on the parameters of the model including δ . Equation 5 becomes a very simple condition if $n = 2$. In that case, maximal rewards are optimal if the quality of information is low (low q), if bribery is infrequent (low λ), and if there are not that many incorruptible politicians ($2 - B$ close to 0).

In figure 6, we show the optimal reward to the higher rank politician as δ goes from zero to one. Below $\bar{\delta}$, voters focus on providing incentives for the lower rank politicians to provide good policies. For very low values of δ , the reward to the higher rank politician must be set very high, since lower rank politicians do not care much about the next period. As δ increases, the necessary reward to keep lower rank politicians adopting good policies declines, but so does the interest of voters in keeping lower rank politicians adopting good policies. At the critical value $\delta = \bar{\delta}$, it is no longer convenient for voters to keep lower rank politicians in good behavior, and the optimal reward jumps discontinuously down to zero. The optimal cutoff is equal to zero below $\bar{\delta}$, and it jumps up to $\Theta(0) > 0$ at the critical value. The value of $\Theta(0)$ declines with δ , and it is equal to zero for $\delta \geq \bar{\delta}$ given by $(1 - \bar{\delta})/\bar{\delta} = 2\lambda q/n$.

If rewards are set optimally, we can rule out perverse effects of increased political competition on the welfare of voters. We have:

Corollary 3. *If $1 - 1/n^2 \geq \lambda q(B - 1)$ in the initial situation, and the rewards to politicians are set optimally, then an increase in the number of lower rank officers strictly increases the welfare of voters, and it increases strictly the welfare of voters if in addition in the initial situation $R^* = 0$ and $\Theta(0) > 0$.*

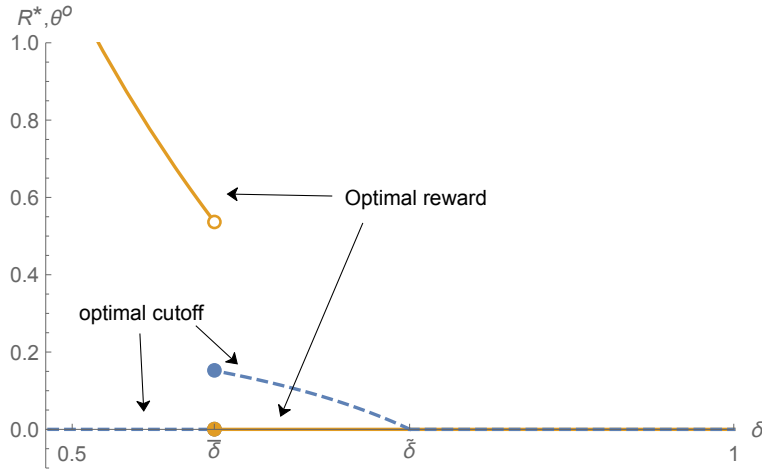


Figure 6: Optimal reward and cutoff as a function of δ

Corollary 3 follows straightforwardly from previous results. From corollary 2, $\Theta(0)$ increases with n , so that the range of cutoffs that can be attained by setting different rewards increases with the number of lower rank politicians. Using the probabilities in equation 3, expected welfare of voters remains constant or increases with n if the cutoff is kept constant by adjusting rewards accordingly. As a consequence, the expected welfare of voters cannot decrease if rewards are readjusted optimally.

Similar positive results cannot be obtained for the quality of public information. In particular, if $R^* = 0$ and $\Theta(0) > 0$ in the initial situation, an increase in the quality of public information reduces $\Theta(0)$, reducing bribery in the present but potentially increasing it in the future, in a situation in which voters care more about the latter at the margin. A countervailing effect is that although the mix of politicians who adopt good policies (and hence are more likely to be promoted) worsens, better information allows to screen out better those politicians who choose bad policies. If $n = 2$, increasing information quality makes the voters better off for any value of R (see Appendix), but for higher values of n this is no longer true. Using the example in figure 6, suppose that initially $n = 5$, $\delta = 0.7$, $q = 0.9$, $\lambda = 0.9$, and $B = 1.3$. Increasing the quality of public information to near 1 reduces voters' expected welfare slightly from 0.8173 to 0.8157.

The mechanism at work in the previous paragraph illustrates what goes wrong with accountability when bribery of the highest rank politician becomes likely: even if rewards after promotion are driven down to zero, “bad” politicians are perversely motivated for promotion by the expectation of future bribes.

The underlying problem for voters is that, although lower rank politicians are subject to the discipline of accountability through the expectation of promotion, the higher rank politician is unaccountable except to their own feelings of accomplishment—unless of course there is some an-

anticipated probability of ending up in jail. This suggests a path to reform: if responsibilities between the levels of government can be reallocated (that is, if it is possible to manipulate δ), then as much responsibility as possible should be delegated to the lower rank, increasing the licit reward of promotion (the perks and symbolism) but reducing the responsibility of higher office. The aftermath of the Odebrecht S.A. scandal, including judicial processes in several Latin American countries, suggests a different path to reform: increase accountability at the top level through vigorous judicial prosecution of bribe-taking by former heads of state, cabinet members, etc. In terms of the model, such path, if sustainable, also calls to increase the licit reward of promotion to keep good behavior at the lower rank.

7 Variations and extensions

In this section, we study the effect of some variations in our assumptions, as well as alternative interpretations of the model. The main version of the model considered so far assumes that lower rank politicians who do not get promoted leave politics. Political careers that are not simply up or out can be introduced by assuming that there are several prizes or promotion positions. In particular, the standard accountability model, i.e. a two-period situation in which the incumbent faces a challenger, can be thought of as a version of the model in which there are as many prizes as positions. This is a useful representation of a situation in which competition with other incumbents for promotion to a higher level of government is not an important motivation for politicians.

We can adapt our framework to the standard accountability model. If the distribution of challenger types is equal to the ex ante distribution of incumbent types, then in a signaling equilibrium, every politician that receives a good signal gets promoted with probability one. We can derive an equivalent to equation 1, that is the expected payoff gain for a politician of adopting the good policy instead of the bad policy in period 1 when the politician's type is $\theta \in [0, B - 1]$:

$$U(\theta) = -(1 - \delta)(B - 1 - \theta) + \delta q[R + (1 - \lambda)\theta + \lambda B] - \delta q\lambda(B - 1).$$

The equilibrium cutoff as a function of q and R is

$$\bar{\theta} = \Theta(q, R) = \max \left\{ 0, \frac{(1 - \delta)(B - 1) - \delta q(R + \lambda)}{1 - \delta + \delta q(1 - \lambda)} \right\}. \quad (6)$$

The expected welfare of voters is now

$$W = 1 - (1 - \delta)\lambda\bar{\theta} - \delta\lambda \left[\lambda q\bar{\theta}(B - 1) + (1 - \lambda q\bar{\theta}) \left(1 - \frac{2 - B}{1 - \lambda q\bar{\theta}} \right) \right]. \quad (7)$$

The second term in this equation is the welfare loss due to bribe taking in the first period; the probability that a bribe is offered and taken in period 1 is, as before, $\lambda\bar{\theta}$. The third term is the welfare loss due to bribe taking in the second period. The first term within the brackets is the probability that the incumbent loses reelection, $\lambda q\bar{\theta}$, multiplied by the probability

that a challenger would take a bribe if offered, $B - 1$. The last term within the brackets is the probability that the incumbent wins reelection, $1 - \lambda q \bar{\theta}$, multiplied by the updated probability that the incumbent has a type below $B - 1$.

Using equations 6 and 7, we can show that the expected welfare of voters is nondecreasing in the quality of information, q , for any value of the reward to politicians, R (see the Appendix for this and other claims). That is, if promotion to a higher level of government is not an important motivation for politicians, better quality of information cannot hurt voters. The optimal reward to politicians still has a bang-bang structure. In particular, maximal rewards are optimal if the quality of information is relatively low, if there are more corruptible politicians, if the second period is relatively less important, and if bribes are relatively rare, exactly the same conditions we found in the main model when $n = 2$.

We briefly consider other variations of the model. The equilibrium structure is built on combining discrete strategies with a continuum of politician types due to an idiosyncratic sense of duty. The assumption that politicians care directly about policies (other and above the warm-glow payoff), though it fits naturally with the interpretation of politicians as citizens, is not essential for the results. If politicians do not care about policies, there is no policy loss associated to victory, so the cutoff structure is preserved. In terms of figure 2, the second period gain of not taking bribes at the lower rank would be simply the dashed line representing future expected office rewards plus bribes. Somewhat paradoxically, politicians who do not care about policy are more keen on promotion.

If there are sources of heterogeneity other than the sense of duty, preserving the cutoff structure depends on whether the order in the desirability of promotion is preserved. Suppose, for instance, that less duty-bound politicians are more likely to receive bribe offers. In that case, in some circumstances, there may be a double-cutoff equilibrium with both the worst politicians and the (relatively) best politicians refusing to take bribes at the lower rank—the worst motivated essentially by future expected bribes, the best by a mixed motive of bribes and expected warm glow feelings.

As an example, let $n = 5$, $\delta = 0.9$, $B = 1.9$, $q = 0.9$, and $R = 0.1$, and suppose that the probability of receiving bribes is a function of the type, as follows:

$$\lambda(\theta) = \begin{cases} 1 & \text{if } \theta \in [0, \tilde{\theta}) \\ 0.1 & \text{if } \theta \in [\tilde{\theta}, 1] \end{cases} .$$

For $\tilde{\theta}$ near zero, there is an equilibrium in which types below $\tilde{\theta}$ and above $\bar{\theta} \approx 0.177$ adopt good policies in period 1, though with very different expectations conditional on promotion.

A version of the model can be used to study electoral campaigns.¹¹ In this alternative interpretation, the first period corresponds to parties formulating policy platforms during the electoral campaign. Investing effort in the platform is a binary choice. Investing no effort in the platform (e.g. a blank endorsement of the leader) is detected by voters with some

¹¹This research idea was suggested by an anonymous referee.

probability. Policy platforms may have some value in themselves for voters, as a contribution to policy debate, and may serve as signals of warm glow feelings regarding policy once in office. In this version of the model, selection would be more important than incentives. We leave this possible line for future work.

8 Final remarks

Grand corruption has motivated a large literature, including the survey by Ogden and Pande (2012), comprehensive works by Rose-Ackerman and Palifka (2016) and Fisman and Golden (2017), and the recent collected volume by Basu and Cordella (2018). A focus of attention has been the measurement of corruption. We contribute to this literature by offering a formal model of the impact of corruption on accountability and political careers that may help organize and motivate empirical research.

As pointed out by Rose-Ackerman and Palifka (2016, p.12), “Many heads of state (presidents, prime ministers, etc.) have stolen government funds throughout history. [...] In somewhat more institutionalized settings, the head of state may derive illicit gains by playing a direct role in public procurement or in the approval of foreign direct investments (FDI) projects by, for example, charging a ‘consulting fee’ for every contract approved.”

Peru’s recent political history offers a stark contrast between both types of grand corruption. Under the nearly autocratic administration of President Fujimori, government officials used embezzled money or money extorted from drug dealing, to bribe media owners, judges, opposition politicians, etc. to secure the regime longevity (see e.g. Mcmillan and Zoido, 2004). Under the subsequent democratic administrations, bribes have flowed instead from large companies, like Odebrecht S.A., to government officials, who had no longer access to other forms of enrichment. We provide a model depicting an “institutionalized setting” in which open kleptocracy is no longer feasible, but exploitation of office for private benefit is.

The Odebrecht S.A. scandal provides an opportunity to observe payments to different levels of government in the countries that were involved. The plea agreement of the Brazilian construction company Odebrecht S.A. with the US Department of Justice, led in 2016, provides details about US 718 million in bribes in ten Latin American and two African countries, paid to office holders at different levels, like federal legislators, state-level officials, and top executives at publicly owned enterprises, involved in procurement of new major infrastructure projects or renegotiation of terms of contract.¹² Campos et al. (2019) connect the bribes to nearly ninety projects awarded to the company, though they do not elaborate on the recipients of the bribes, as they focus on the effects of bribery on concession contracts. Additional empirical work on the basis of subsequent investigations (see e.g. ICIJ, 2019) to understand the effects of bribery on politician selection would be helpful.

¹² *United States vs Odebrecht, S.A.*, December 15, 2016, US District Court, E.D.N.Y.

The potential of bribes to alter career incentives was clearly not lost to politicians at the time of Odebrecht’s bribery schemes. As former President of Peru Alan García allegedly explained to a journalist and aspiring politician who complained about the lackluster presidential salary, “No seas cojudo, la plata llega sola” (don’t be stupid, the money comes on its own) (Bayly, 2010).

Our research indicates that increasing accountability at the top rung of the political ladder through the use of the judiciary alleviates the distortions on political careers occasioned by bribery. Avis et al. (2018) find that a “legal disciplining effect” may have had a larger role in the reduction of corruption than electoral incentives after random audits in Brazil. From a broader perspective, however, the behavior of the judiciary is hardly exogenous; judges can themselves be influenced by bribes or used for political purposes. The Odebrecht S.A. scandal itself may be an “out-of-equilibrium” event, warning potentially corrupt politicians and firms of the dangers of overreach.

Political economy models like the one developed in this paper highlight that corruption is an equilibrium phenomenon, not just the effect of some ill-functioning specific political institution. Corruption, of course, is a multifaceted phenomenon, and any one particular angle can only provide a partial view. We hope our analysis of the relation between careers and corruption provides a useful angle to understanding the overall picture of corruption in politics.

A Appendix

A.1 Proof of lemma 1

To characterize the behavior of politicians, we proceed backwards, starting from the decision in the second period of the politician who gets promoted. The second-period strategy σ_{i2} ,

$$\sigma_{i2} = \begin{cases} 0 & \text{if } \theta_i < B - 1 \text{ and } b_2 = 1 \\ 1 & \text{if } \theta_i \geq B - 1 \text{ or } b_2 = 0 \end{cases} \quad (8)$$

for all $i = 1, \dots, n$ follows directly from the payoffs specified in the model.

Using equation 8, the interim (that is, before knowing the second period bribe) expected second period payoff of a politician that is promoted is $\delta(R + V^2(\theta_i))$, where

$$V^2(\theta_i) = \begin{cases} (1 - \lambda)(1 + \theta_i) + \lambda B & \text{if } \theta_i \leq B - 1 \\ 1 + \theta_i & \text{if } \theta_i \geq B - 1 \end{cases} .$$

An implication of equation 8 is that for every signal profile, voters should promote with positive probability only the politicians whom voters believe have the highest probability of having a type above $B - 1$. That is,

$$\nu_i(s_{11}, \dots, s_{n1}) > 0 \text{ only if } \beta_i(s_i)[B - 1] \leq \min_j \beta_j(s_j)[B - 1]. \quad (9)$$

For any given politician i , consider any profile of signals for the other politicians, $s_{-i} = (s_j)_{j \neq i}$, and let $\pi_{-i}(s_{-i})$ be the probability of such signal profile as induced by the strategies of the other politicians. Note that for any such profile that has positive probability, the beliefs of voters about each politician $j \neq i$ are pinned down by prior beliefs and the strategy of j , using Bayes law. Using other politicians' behavior and voters' behavior as described by equations 8 and 9, the expected second period payoff for politician i , conditional on signal i , can be written as

$$\begin{aligned} \delta V_i(s_i, \theta_i) = & \delta \sum_{s_{-i}} \pi_{-i}(s_{-i}) \left[v_i(s_i, s_{-i})(R + V^2(\theta_i)) \right. \\ & \left. + (1 - v_i(s_i, s_{-i}))(1 - \lambda \min_{j \neq i} \beta_j(s_j)[B - 1]) \right]. \end{aligned}$$

Taking into account the probabilities of signals $s_i = 1$ and $s_i = 0$ conditional on policy decisions of i , the expected payoff of adopting the policy $x_{i1} = 1$ can be written as

$$(1 - \delta)(1 + \theta_i) + \delta V_i(1, \theta_i),$$

while the payoff of adopting the policy $x_{i1} = 0$ can be written as

$$(1 - \delta)Bb_{i1} + \delta q V_i(0, \theta_i) + \delta(1 - q)V_i(1, \theta_i).$$

The net gain of adopting policy 1 (and rejecting the bribe) for a politician of type θ_i and bribe opportunity b_{i1} can then be written as

$$U_i(\theta_i, b_{i1}) = (1 - \delta)(1 + \theta_i - Bb_{i1}) + \delta q(V_i(1, \theta_i) - V_i(0, \theta_i)). \quad (10)$$

We claim that $V_i(1, \theta_i) - V_i(0, \theta_i)$ is nonnegative and increasing in θ_i . To see this, using equation 9 we get

$$\begin{aligned} V_i(1, \theta_i) - V_i(0, \theta_i) = & \\ & \sum_{s_{-i} \in S_{-i}} \pi_{-i}(s_{-i})(\nu_i(1, s_{-i}) - \nu_i(0, s_{-i})) \\ & \times (R + V^2(\theta_i) - 1 + \lambda \min_{j \neq i} \beta_j(s_j)[B - 1]) \quad (11) \end{aligned}$$

where

$$S_i \equiv \{s_{-i} : \beta_i(0)[B - 1] \leq \min_{j \neq i} \beta_j(s_j)[B - 1] \leq \beta_i(1)[B - 1]\}.$$

Since $\nu_i(1, s_{-i}) \geq \nu_i(0, s_{-i})$, and $V^2(\theta_i)$ is larger than one and strictly increasing in θ_i , the claim follows. Using the claim and equation 10, $U_i(\theta_i, b_{i1})$ is strictly increasing in θ_i . Since

$$U_i(0, 0) \geq 0 \quad \text{and} \quad U_i(B - 1, 1) \geq 0,$$

it follows that if $b_{i1} = 0$, politician i adopt the policy 1, and there is some $\bar{\theta}_i \leq B - 1$ such that if $b_{i1} = 1$, politician i adopt the policy 1 if and only if $\theta_i \geq \bar{\theta}_i$, as required by the definition of cutoff strategies. \square

A.2 Proof of lemma 2

We proceed in two steps.

A.2.1 In a neutral signaling equilibrium, politicians play the same cutoff strategy.

From lemma 1, politicians play cutoff strategies with cutoffs $\bar{\theta}_i$ below $B - 1$. Thus, in equilibrium, using Bayes law and consistency, if $\bar{\theta}_i > 0$,

$$\beta_i(1)[B - 1] = 1 - \frac{2 - B}{1 - \lambda q \bar{\theta}_i} \quad \text{and} \quad \beta_i(0)[B - 1] = 1,$$

and if $\bar{\theta}_i = 0$,¹³

$$\beta_i(1)[B - 1] = B - 1 \quad \text{and} \quad B - 1 < \beta_i(0)[B - 1] \leq 1.$$

Suppose, by way of contradiction, that two politicians i and j choose cutoff strategies with cutoffs $0 \leq \bar{\theta}_i < \bar{\theta}_j \leq B - 1$. Without loss of generality, suppose $\bar{\theta}_i = \min_k \bar{\theta}_k$ and $\bar{\theta}_j = \max_k \bar{\theta}_k$. Note that, by definition, $U_i(\bar{\theta}_i, 1) \geq 0$, with equality if $\bar{\theta}_i > 0$, and $U_j(\bar{\theta}_j, 1) = 0$. Hence, since $U_j(\bar{\theta}_j, 1)$ is strictly increasing in θ , it follows that $U_j(\bar{\theta}_i, 1) < 0$.

We claim that for all $\theta \in [0, 1]$,

$$V_j(1, \theta) - V_j(0, \theta) \geq V_i(1, \theta) - V_i(0, \theta).$$

¹³The extra degree of freedom for $\beta_i(0)[B - 1]$ when $\bar{\theta}_i = 0$ is a consequence of out-of-equilibrium beliefs; the lower bound is due to constraint (iv).

To see this, consider any signal profile s_{-i} for politicians other than i such that $\nu_i(1, s_{-i}) > 0$. For any such profile we must have $s_k = 0$ for all $k \neq i$, since

$$\beta_k(1)[B-1] < \beta_i(1)[B-1].$$

Note that $\pi_{-i}(0, \dots, 0) = (\lambda q)^{n-1} \prod_{k \neq i} \bar{\theta}_k$. But then for the signal profiles $s_{-j} = (s_i, (s_k)_{k \neq i, j}) = (1, 0, \dots, 0)$ and $s'_{-j} = (0, 0, \dots, 0)$ we have $\nu_j(1, s_{-j}) = \nu_j(1, s'_{-j}) = 1$. Moreover,

$$\pi_{-j}(s_{-j}) + \pi_{-j}(s'_{-j}) = (\lambda q)^{n-2} \prod_{k \neq i, j} \bar{\theta}_k \geq \pi_{-i}(s_{-i}).$$

Note that for $s_{-i} = (0, \dots, 0)$ we have $\min_{k \neq i} \beta_k(s_k)[B-1] = 1$, while for $s_{-j} = (s_i, 0, \dots, 0)$ we have $\min_{k \neq j} \beta_k(s_k)[B-1] = 1$ for $s_i = 1$ and for $s_i = 0$ if $\bar{\theta}_i > 0$ (that is for every signal for i that has positive probability). That is, politician j 's expected payoff after a signal of 1 is greater than or equal to the politician i 's after a signal of 1.

Similarly, consider any signal profile s_{-j} such that $\nu_j(0, s_{-j}) > 0$. For any such profile we must have $s_k = 0$ for all $k \neq j$, since

$$\beta_k(1)[B-1] < \beta_j(0)[B-1] = 1.$$

Note that $\pi_{-j}(0, \dots, 0) = (\lambda q)^{n-1} \prod_{k \neq j} \bar{\theta}_k$. But then for the signal profile $s_{-i} = (0, \dots, 0)$ we have that neutrality implies $\nu_i(0, s_{-i}) \geq \nu_j(0, s_{-j})$ (with strict equality if $\bar{\theta}_i > 0$). Moreover, $\pi_{-i}(s_{-i}) = (\lambda q)^{n-1} \prod_{k \neq i} \bar{\theta}_k \geq \pi_{-j}(s_{-j})$. That is, politician i 's expected payoff after a signal of 0 is greater than or equal to the politician j 's after a signal of 0.

The claim follows by inspection of equation 11. From the claim, using equation 10, we get $U_j(\bar{\theta}_i, 1) \geq U_i(\bar{\theta}_i, 1) \geq 0$, which contradicts $U_j(\bar{\theta}_i, 1) < 0$.

A.2.2 In a neutral signaling equilibrium, voters play the monotonic strategy.

From the previous step, all politicians use the same cutoff $0 \leq \bar{\theta} \leq B-1$. If $\bar{\theta} > 0$, then

$$\beta_i(1)[B-1] = \beta_j(1)[B-1] < \beta_i(0)[B-1] = \beta_j(0)[B-1] = 1$$

for all i, j .

If politicians use the cutoff $\bar{\theta} = 0$, we get from Bayes law

$$\beta_i(1)[B-1] = \beta_j(1)[B-1] = B-1$$

for all i, j , and from equilibrium condition (iv) we get

$$\beta_i(0)[B-1] > B-1$$

for all i . Note that in the public signal vector $(0, \dots, 0)$, every signal has zero probability given the strategy of politicians. From neutrality of equilibrium,

$$\beta_i(0)[B-1] = \beta_j(0)[B-1] \text{ for all } i, j.$$

The monotonic strategy follows from neutrality in either case.¹⁴ \square

¹⁴If we did not impose any condition on “extreme” out-of-equilibrium beliefs, we would

A.3 Proof of theorem 1

We proceed in several steps. We first show that the net gain of adopting policy 1 and rejecting a bribe for a politician of type θ when every other politician adopts a symmetric cutoff $\bar{\theta}$ and voters play the monotonic strategy is equal to $U(\theta, \bar{\theta})$ as defined in the main text. Next we use U to characterize best-response behavior by politicians. We then show that a (signaling neutral) equilibrium cutoff for politicians $\bar{\theta}$ is equal to 0 if $U(0, 0) \geq 0$ and otherwise solves the equation $U(\bar{\theta}, \bar{\theta}) = 0$. We finally show that such cutoff exists and is unique.

A.3.1 Derivation of $U(\theta, \bar{\theta})$.

Consider politician i and suppose every other politician adopts the cutoff $\bar{\theta} \in [0, B - 1]$, and voters play the monotonic strategy. We construct the net gain of adopting policy 1 as in equation 11.

Consider any profile s_{-i} such that $\omega \in \{1, \dots, n - 2\}$ politicians other than i obtain a signal of 0 and the remainder obtain a signal of 1. For every such profile, we have $\nu_i(1, s_{-i}) = 1/(n - \omega)$ and $\nu_i(0, s_{-i}) = 0$. The probability of any such profile is $\pi_{-i}(s_{-i}) = (\lambda q \bar{\theta})^\omega (1 - \lambda q \bar{\theta})^{n-1-\omega}$, and there are $\binom{n-1}{\omega}$ such profiles. Using Bayes law, for any such profile

$$\min_{j \neq i} \beta_j(s_j)[B - 1] = \beta_j(1)[B - 1] = 1 - \frac{2 - B}{1 - \lambda q \bar{\theta}}.$$

We can derive the net gain of winning rather than losing the election for any such profile as in equation 11 to be

$$\delta \left[R + V^2(\theta) - 1 + \lambda \left(1 - \frac{2 - B}{1 - \lambda q \bar{\theta}} \right) \right],$$

which, using $\theta \leq B - 1$, simplifies to

$$\delta \left[R + (1 - \lambda)\theta + \lambda \left(B - \frac{2 - B}{1 - \lambda q \bar{\theta}} \right) \right].$$

Similarly, consider the profile $s_{-i} = (1, \dots, 1)$. The probabilities of promotion are $\nu_i(1, s_{-i}) = 1/n$ and $\nu_i(0, s_{-i}) = 0$, and the probability of the profile is $(1 - \bar{\theta} \lambda q)^{n-1}$, which is equal to 1 when $\bar{\theta} = 0$. The net gain of winning rather than losing the election is the same as in the previous paragraph.

Finally consider the profile $s_{-i} = (0, \dots, 0)$. The probabilities of promotion are $\nu_i(1, s_{-i}) = 1$ and $\nu_i(0, s_{-i}) = 1/n$, and the probability of the profile is $(\lambda q \bar{\theta})^{n-1}$. Using Bayes law, for such profile

$$\min_{j \neq i} \beta_j(s_j)[B - 1] = \beta_j(0)[B - 1] = 1.$$

need to relax the definition of monotonic strategy when the signal profile is $(0, \dots, 0)$ and such profile requires every politician to deviate from equilibrium play. (This relaxation of monotonicity does not affect equilibrium construction, since the signal profile $(0, \dots, 0)$ only occurs after joint deviations.)

We can derive the net gain of winning rather than losing the election for such profile as

$$\delta [R + (1 - \lambda)\theta + \lambda B].$$

Putting the three cases together as in equation 11, we obtain

$$\begin{aligned} U(\theta, \bar{\theta}) &= (1 - \delta)(-B + 1 + \theta) \\ &+ \delta q \sum_{\omega=0}^{n-2} \binom{n-1}{\omega} (\lambda q \bar{\theta})^\omega (1 - \lambda q \bar{\theta})^{n-1-\omega} \\ &\quad \times \frac{1}{n - \omega} \left[R + (1 - \lambda)\theta + \lambda B - \frac{\lambda(2 - B)}{1 - \lambda q \bar{\theta}} \right] \\ &+ \delta q (\lambda q \bar{\theta})^{n-1} \times \frac{n-1}{n} [R + (1 - \lambda)\theta + \lambda B] \end{aligned}$$

if $\bar{\theta} \in (0, B - 1]$, and

$$U(\theta, 0) = (1 - \delta)(-B + 1 + \theta) + \delta q \frac{1}{n} [R + (1 - \lambda)\theta + 2\lambda(B - 1)].$$

Note that

$$\begin{aligned} &\sum_{\omega=0}^{n-2} \binom{n-1}{\omega} (\lambda q \bar{\theta})^\omega (1 - \lambda q \bar{\theta})^{n-1-\omega} \frac{1}{n - \omega} + (\lambda q \bar{\theta})^{n-1} \times \frac{n-1}{n} \\ &= \sum_{\omega=0}^{n-1} \binom{n-1}{\omega} (\lambda q \bar{\theta})^\omega (1 - \lambda q \bar{\theta})^{n-1-\omega} \frac{1}{n - \omega} - \frac{1}{n} (\lambda q \bar{\theta})^{n-1} \\ &= \frac{1}{n(1 - \lambda q \bar{\theta})} \sum_{\omega=0}^{n-1} \binom{n}{\omega} (\lambda q \bar{\theta})^\omega (1 - \lambda q \bar{\theta})^{n-\omega} - \frac{1}{n} (\lambda q \bar{\theta})^{n-1} \\ &= \frac{1 - (\lambda q \bar{\theta})^n}{n(1 - \lambda q \bar{\theta})} - \frac{1}{n} (\lambda q \bar{\theta})^{n-1} = \frac{1 - (\lambda q \bar{\theta})^{n-1}}{n(1 - \lambda q \bar{\theta})}. \end{aligned}$$

Thus,

$$\begin{aligned} U(\theta, \bar{\theta}) &= (1 - \delta)(-B + 1 + \theta) \\ &+ \delta q \left(\frac{1 - (\lambda q \bar{\theta})^{n-1}}{n(1 - \lambda q \bar{\theta})} - (\lambda q \bar{\theta})^{n-1} \frac{n-1}{n} \right) \\ &\quad \times \left[R + (1 - \lambda)\theta + \lambda B - \frac{\lambda(2 - B)}{1 - \lambda q \bar{\theta}} \right] \\ &+ \delta q (\lambda q \bar{\theta})^{n-1} \times \frac{n-1}{n} [R + (1 - \lambda)\theta + \lambda B]. \end{aligned}$$

Equation 1 in the text follows.

A.3.2 Best response behavior.

Since $U(\theta, \bar{\theta})$ is strictly increasing in θ , and $U(B - 1, \bar{\theta}) > 0$, it follows that a cutoff strategy $\bar{\theta}_i$ is a best response for politician i to every other politician using the cutoff strategy $\bar{\theta}$ if either $\bar{\theta}_i = 0$ and $U(0, \bar{\theta}) \geq 0$ or $0 < \bar{\theta}_i < B - 1$ and $U(\bar{\theta}_i, \bar{\theta}) = 0$. Note that for every cutoff strategy $\bar{\theta} \in [0, B - 1]$ played by other politicians, there is a unique best response cutoff strategy $\bar{\theta}_i$ for politician i .

A.3.3 Necessary and sufficient condition for equilibrium.

From lemma 2, we know that a (neutral signaling) equilibrium is necessarily symmetric. Hence every politician plays a best response to other politicians if and only if each adopts a policy $\hat{\theta}$ such that either $\hat{\theta} = 0$ and $U(0, 0) = G(0) \geq 0$ or $0 < \hat{\theta} < B - 1$ and $U(\hat{\theta}, \hat{\theta}) = G(\hat{\theta}) = 0$. Equation 2 in the main text summarizes this condition.

A.3.4 Equilibrium existence and uniqueness.

Since G is continuous and $G(B - 1) > 0$, it follows there is a solution θ^* to the equilibrium condition given by equation 2. To show that the equilibrium is unique, we claim that $G(\theta) = 0$ with $\theta \in (0, B - 1)$ implies $G'(\theta) > 0$ so that there is at most one zero-crossing of G . To see this, from equation 1,

$$\begin{aligned} G(\theta) &= -(1 - \delta)(B - 1 - \theta) \\ &\quad + \delta q \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)} \right) \times [R + (1 - \lambda)\theta + \lambda(B - 1)] \\ &\quad + \delta q \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)} \right) \times \frac{\lambda(B - 1 - \lambda q \theta)}{1 - \lambda q \theta} \\ &\quad + \delta q (\lambda q \theta)^{n-1} \times \frac{n-1}{n} \left(\frac{\lambda(2 - B)}{1 - \lambda q \theta} \right), \end{aligned}$$

or, after some algebra,

$$\begin{aligned} G(\theta) &= -(B - 1 - \theta) \left[1 - \delta - \delta q \lambda \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)^2} \right) \right] \\ &\quad + \delta q \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)} \right) \times (R + (1 - \lambda)\theta + \lambda(B - 1)) \\ &\quad + \delta q \left(\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)} \right) \times \frac{\lambda \theta (1 - \lambda q)}{1 - \lambda q \theta} \\ &\quad + \delta q (\lambda q \theta)^{n-1} \times \frac{n-1}{n} \left(\frac{\lambda(2 - B)}{1 - \lambda q \theta} \right). \end{aligned} \tag{12}$$

Note that $G(\theta)$ is strictly increasing if the expression in brackets is strictly positive. In particular, taking derivatives, we can show that

$$\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)^2} \quad \text{and} \quad \frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)}$$

are increasing in θ for $\theta < B - 1 < 1$.

If $G(\hat{\theta}) = 0$ for some $\hat{\theta} \in [0, B - 1)$,

$$\begin{aligned} (B - 1 - \hat{\theta}) \left[1 - \delta - \delta q \lambda \left(\frac{1 - (\lambda q \hat{\theta})^{n-1}}{n(1 - \lambda q \hat{\theta})^2} \right) \right] &= \\ \delta q \left(\frac{1 - (\lambda q \hat{\theta})^{n-1}}{n(1 - \lambda q \hat{\theta})} \right) \times \left(R + (1 - \lambda)\hat{\theta} + \lambda(B - 1) + \frac{\lambda \hat{\theta} (1 - \lambda q)}{1 - \lambda q \hat{\theta}} \right) & \\ + \delta q (\lambda q \hat{\theta})^{n-1} \times \frac{n-1}{n} \left(\frac{\lambda(2 - B)}{1 - \lambda q \hat{\theta}} \right). & \end{aligned}$$

The right-hand side of the equation above is strictly positive. Hence, the left-hand side must be strictly positive. This, in turn, implies that the term in brackets is strictly positive, implying that $G(\theta)$ is strictly increasing around $\theta = \hat{\theta}$. \square

A.4 Proof of corollary 1

(In the main text.)

A.5 Proof of corollary 2

Note that

$$\frac{1 - (\lambda q \theta)^{n-1}}{n(1 - \lambda q \theta)}$$

is decreasing in n . Similarly,

$$(\lambda q \theta)^{n-1} \times \frac{n-1}{n}$$

is strictly decreasing in n if and only if $1 - 1/n^2 > \lambda q \theta$. Recall that in equilibrium $\theta^* < B - 1$. Using these observations and equation 12, we have that θ^* must be strictly increasing in n under the conditions in the statement of the corollary. This directly implies that expected bribery in the first period increases with n under the conditions in the statement of the corollary.

With respect to expected bribery in the second period, note that

$$\frac{1 - (\lambda q \theta^*)^n}{1 - \lambda q \theta^*}$$

in the expression in 3 is strictly increasing in n and in θ^* . Using the result in the previous paragraph, it follows that expected bribery in the second period increases with n under the conditions in the statement of the corollary. \square

A.6 Proof of lemma 3

Equation 4 can be obtained from equation 1. Note that F is strictly increasing in R . Following the steps of the proof of theorem 1 (section A.3.4), F is strictly increasing in θ if $F(\theta, R) \geq 0$. Moreover, from theorem 1, $\hat{\theta}$ is an equilibrium cutoff if it solves

$$F(\hat{\theta}, R) \geq 0 \text{ with equality if } \hat{\theta} > 0.$$

Manipulating F , we get that the equilibrium cutoff is equal to zero if and only if $R \geq \bar{R}$. For smaller values of R , then, the equilibrium cutoff is strictly decreasing in R and satisfies $F(\hat{\theta}, R) = 0$. The maximum value of the equilibrium cutoff is obtained by setting $R = 0$; it is smaller than $B - 1$ since $F(B - 1, 0) > 0$. Using the implicit function theorem, since

F is continuously differentiable, then $\Theta(R)$ satisfying $\Theta(R) = \hat{\theta}$ such that $F(\hat{\theta}, R) = 0$ is a continuously differentiable function for $0 < R < \bar{R}$. Since $\Theta(R)$ is strictly decreasing, it follows that $\Theta'(R) < 0$ for $0 < R < \bar{R}$. Since $\Theta(R)$ is continuously differentiable on $(0, \bar{R})$, it is continuous on $[0, \bar{R}]$. \square

A.7 Proof of theorem 2

We proceed in two steps. We first characterize the optimal reward R^* , and then show that there is a critical value $\bar{\delta}$ of the importance of higher office such that $R^* > 0$ if and only if $\delta < \bar{\delta}$.

A.7.1 Characterization of optimal reward R^*

From lemma 3, by choosing $R \in [0, \bar{R}]$, it is possible to induce an equilibrium cutoff equal to any value in the interval $[0, \Theta(0)]$, where

$$\Theta(0) > 0 \quad \text{if and only if} \quad (1 - \delta)/\delta > 2\lambda q/n.$$

In particular, let $\tilde{\Theta} : [0, \bar{R}] \rightarrow [0, \Theta(0)]$ be given by $\tilde{\Theta} = \Theta$ on the restricted domain, so that the function $\tilde{\Theta}$ admits an inverse. Then we can write the voters' welfare problem as setting $R^* = \tilde{\Theta}^{-1}(\theta^\circ)$, where θ° solves

$$\theta^\circ = \sup \arg \max_{\theta \in [0, \Theta(0)]} \tilde{W}(\theta).$$

Note that

$$\tilde{W}'(\theta) = -(1 - \delta) + \delta\lambda q(2 - B) \left[\frac{1 - n(\lambda q\theta)^{n-1} + (n-1)(\lambda q\theta)^n}{(1 - \lambda q\theta)^2} \right],$$

where the term in brackets is equal to one if $\theta = 0$, and it is constant for $n = 2$ and strictly increasing for $n \geq 3$. Hence, $\tilde{W}(\theta)$ is either linear or strictly convex, so that

$$\theta^\circ = \begin{cases} 0 & \text{if } \tilde{W}'(0) > \tilde{W}'(\Theta(0)) \\ \Theta(0) & \text{if } \tilde{W}'(0) \leq \tilde{W}'(\Theta(0)) \end{cases}.$$

Since $R = \bar{R}$ is the minimal value of R inducing the cutoff $\theta = 0$, and $R = 0$ induces the cutoff $\Theta(0)$,

$$R^* = \begin{cases} \bar{R} & \text{if } \tilde{W}'(0) > \tilde{W}'(\Theta(0)) \\ 0 & \text{if } \tilde{W}'(0) \leq \tilde{W}'(\Theta(0)) \end{cases}.$$

Using

$$\tilde{W}(0) = 1 - \delta\lambda + \delta\lambda(2 - B)$$

and

$$\tilde{W}(\Theta(0)) = 1 - (1 - \delta)\lambda\Theta(0) - \delta\lambda + \delta\lambda(1 - (\lambda q\Theta(0))^n) \frac{2 - B}{1 - \lambda q\Theta(0)},$$

we get

$$\tilde{W}(0) > \tilde{W}(\Theta(0))$$

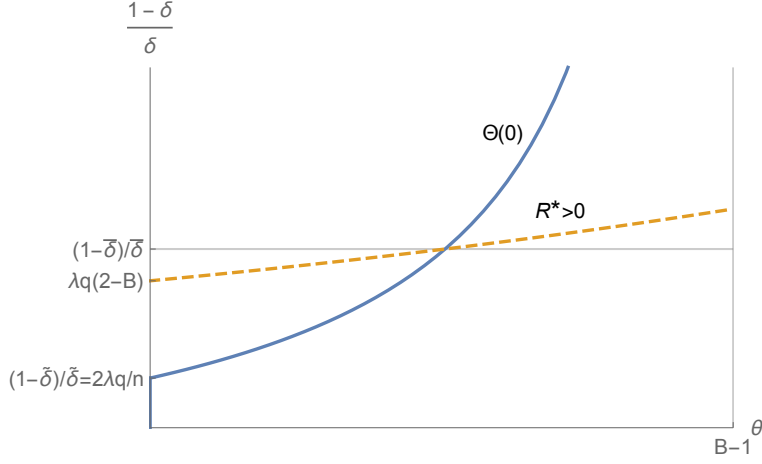


Figure 7: Critical values of δ

if and only the inequality

$$\frac{1-\delta}{\delta} > \lambda q (1 - (\lambda q \theta)^{n-1}) \frac{2-B}{1-\lambda q \theta}.$$

is satisfied for $\theta = \Theta(0)$.

We represent this inequality by the area above the dashed line in figure 7, where the dashed line is given by equation 13 in the main text. Note that the right-hand side of equation 13 takes the value $\lambda q(2-B)$ for $\theta = 0$ and it is bounded above for $\theta = B-1$. (We calculated figure 7 using the parameters of the example in section 6 for illustration purposes.)

A.7.2 Existence of a critical value $\bar{\delta}$

It remains to be shown that equation 5 has a unique solution $\bar{\delta}$, so that $\tilde{W}(0) > \tilde{W}(\Theta(0))$ if and only if $\delta < \bar{\delta}$. As depicted in figure 7, there is a unique solution to equation 5 because the equation characterizing the region in which positive rewards are optimal (the dashed line),

$$\frac{1-\delta}{\delta} = \lambda q \frac{1 - (\lambda q \theta)^{n-1}}{1 - \lambda q \theta} (2-B), \quad (13)$$

crosses exactly once the equation characterizing $\Theta(0)$ as a function of δ (the continuous line), obtained from $F(\Theta(0), 0) = 0$.

To see this, using equation 4, we have that $\Theta(0) = 0$ if $(1-\delta)/\delta \leq 2\lambda q/n$, and otherwise it satisfies the equation

$$\begin{aligned} \frac{1-\delta}{\delta} = \frac{q/n}{B-1-\theta} & \left[\frac{1 - (\lambda q \theta)^{n-1}}{1 - \lambda q \theta} ((1-\lambda)\theta + \lambda B) \right. \\ & \left. - \left(\frac{1 - (\lambda q \theta)^{n-1}}{1 - \lambda q \theta} - (n-1)(\lambda q \theta)^{n-1} \right) \frac{\lambda(2-B)}{1 - \lambda q \theta} \right]. \quad (14) \end{aligned}$$

We represent this equation by the continuous line in figure 7. Note that the right-hand side takes the value $2\lambda q/n$ for $\theta = 0$ and grows unboundedly as θ approaches $B - 1$. Note also that $2\lambda q/n < \lambda q(2 - B)$, so that equations 13 and 14 must cross at least once. We claim that equations 13 and 14 cross exactly once, so that there is some critical value $\bar{\theta}$ with the desired property.

To prove the claim, note that if equations 13 and 14 cross at some θ , then θ must satisfy

$$\begin{aligned} \lambda q \frac{1 - (\lambda q \theta)^{n-1}}{1 - \lambda q \theta} (2 - B)(B - 1 - \theta) = \\ (q/n) \left[\frac{1 - (\lambda q \theta)^{n-1}}{1 - \lambda q \theta} ((1 - \lambda)\theta + \lambda B) \right. \\ \left. - \left(\frac{1 - (\lambda q \theta)^{n-1}}{1 - \lambda q \theta} - (n - 1)(\lambda q \theta)^{n-1} \right) \frac{\lambda(2 - B)}{1 - \lambda q \theta} \right]. \end{aligned}$$

After some simplification, this is equivalent to

$$\begin{aligned} (B - 1 - \theta)(2 - B) = \frac{1}{n\lambda} \left[(1 - \lambda)\theta + \lambda B - \lambda \frac{2 - B}{1 - \lambda q \theta} \right. \\ \left. + \frac{(n - 1)(\lambda q \theta)^{n-1} \lambda(2 - B)}{1 - (\lambda q \theta)^{n-1}} \right]. \end{aligned}$$

Re-arranging,

$$\begin{aligned} (B - 1 - \theta)(2 - B) = \frac{1}{n\lambda} \left[(1 - \lambda)\theta + \lambda(B - 1) + \lambda \frac{B - 1 - \lambda q \theta}{1 - \lambda q \theta} \right. \\ \left. + \frac{(n - 1)(\lambda q \theta)^{n-1} \lambda(2 - B)}{1 - (\lambda q \theta)^{n-1}} \right]. \end{aligned}$$

Or equivalently,

$$\begin{aligned} (B - 1 - \theta) \left[2 - B - \frac{1/n}{1 - \lambda q \theta} \right] \\ = \frac{1}{n\lambda} \left[(1 - \lambda)\theta + \lambda(B - 1) + \frac{\lambda \theta(1 - q)}{1 - \lambda q \theta} \right. \\ \left. + \frac{(n - 1)(\lambda q \theta)^{n-1} \lambda(2 - B)}{1 - (\lambda q \theta)^{n-1}} \right]. \end{aligned}$$

The right-hand side is positive for any $\theta \in [0, B - 1]$; hence, the term in square brackets in the left-hand side must be positive. Note that the expression in the left-hand side is strictly decreasing in θ , while the expression in the right-hand side is strictly increasing in θ . Hence, there is at most one solution to the equation above. \square

A.8 Proof of corollary 3

(In the main text.)

A.9 Proof of claims in section 6

We claim that when $n = 2$, increasing the quality of information cannot hurt voters. To see this, we can use equation 4 to obtain

$$\Theta(R) = \max \left\{ 0, \frac{(B-1)(1-\delta-\delta q\lambda) - \delta q R/2}{1 - \delta\delta q(1-\lambda)/2} \right\}.$$

Hence, if $\Theta(R) > 0$, we have $q < (1-\delta)/\delta\lambda$. The welfare of voters is then

$$1 - (1-\delta)\lambda\Theta(R) - \delta\lambda \left(1 - (1 - (\lambda q\Theta(R))^2) \frac{2-B}{1 - \lambda q\Theta(R)} \right)$$

or, simplifying,

$$1 - \delta\lambda(B-1) + [-(1-\lambda) + \delta\lambda q(2-B)]\Theta(R).$$

The term in brackets is negative if $q < (1-\delta)/\delta\lambda(2-B)$ which is true if $\Theta(R) > 0$. Since $\Theta(R)$ is strictly decreasing in q , it follows that increasing the quality of information makes voters better off.

A.10 Proof of claims in section 7

Using equations 6 and 7, we get that the expected welfare of voters in the model with no promotion incentives is proportional to

$$\hat{W} = [\delta\lambda q(2-B) - (1-\delta)] \times \max \left\{ 0, \frac{(1-\delta)(B-1) - \delta q(R+\lambda)}{1-\delta + \delta q(1-\lambda)} \right\}.$$

We claim that \hat{W} is nondecreasing in q for any value of R . To see this, note that the second term in \hat{W} is equal to zero unless

$$q < \bar{q} = \frac{(1-\delta)(B-1)}{\delta(R+\lambda)},$$

and moreover, the second term is positive and strictly decreasing in q if the inequality above is satisfied. The first term is negative if

$$q < \tilde{q} = \frac{1-\delta}{\delta\lambda(2-B)}.$$

Since $\tilde{q} > \bar{q}$, the expected welfare of voters is constant in q if $q \geq \tilde{q}$ and strictly increasing in q if $q < \tilde{q}$.

We also claim that the optimal reward to politicians has a bang-bang structure. To see this, note that the second term in \hat{W} is zero for any value of R if

$$q > \underline{q} = \frac{(1-\delta)(B-1)}{\delta\lambda}$$

and otherwise is strictly decreasing in R for

$$R < \bar{R} = \lambda + \frac{(1-\delta)(B-1)}{\delta q}$$

and equal to zero if $R \geq \bar{R}$. Since $\tilde{q} > \underline{q}$, we obtain that the optimal reward is

$$R^* = \begin{cases} 0 & \text{if } q > \underline{q} \\ \bar{R} & \text{if } q \leq \underline{q} \end{cases} .$$

Note that the condition for maximal rewards to be optimal, $q \leq \underline{q}$, is satisfied if the quality of information is relatively low, if there are more corruptible politicians ($B - 1$ close to one), if the second period is more important (δ close to zero), and if bribes are relatively rare (λ close to zero).

References

- Acemoglu, D., G. Egorov, and K. Sonin (2013). A political theory of populism. *Quarterly Journal of Economics* 128, 771–805.
- Ashworth, S. (2005). Reputational dynamics and political careers. *Journal of Law, Economics, and Organization* 21, 441–466.
- Ashworth, S. (2012). Electoral accountability: Recent theoretical and empirical work. *Annual Review of Political Science* 15, 183–201.
- Ashworth, S., E. Bueno de Mesquita, and A. Friedenber (2017). Accountability and information in elections. *American Economic Journal: Microeconomics* 9(2), 95–138.
- Austen-Smith, D. and J. Banks (1989). Electoral accountability and incumbency. In Peter Ordeshook (Ed.), *Models of Strategic Choice in Politics*, pp. 121–150. University of Michigan Press.
- Avis, E., C. Ferraz, and F. Finan (2018). Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians. *Journal of Political Economy* 126(5), 1912–1964.
- Banks, J. and R. Sundaram (1993). Adverse selection and moral hazard in a repeated election model. In W. Barnett, M. Hinich, and N. Schofield (Eds.), *Political Economy: Institutions, Information, Competition, and Representation*, pp. 295–311. Cambridge University Press.
- Basu, K. and T. Cordella (Eds.) (2018). *Institutions, Governance and the Control of Corruption*. Palgrave MacMillan.
- Bayly, J. (2010, December 6). La plata llega sola. *Diario Peru* 21.
- Besley, T. (2005). Political selection. *Journal of Economic Perspectives* 19(3), 43–60.
- Besley, T. (2006). *Principled Agents? The Political Economy of Good Government*. Oxford University Press.
- Bobonis, G. J., L. R. C. Fuertes, and R. Schwabe (2016). Monitoring corruptible politicians. *American Economic Review* 106(8), 2371–2405.
- Boix, C. (2003). *Democracy and Redistribution*. Cambridge, UK: Cambridge University Press.
- Campos, N., E. Engel, R. D. Fischer, and A. Galetovic (2019). Renegotiations and corruption in infrastructure: The Odebrecht case. SSRN.
- Caselli, F. and M. Morelli (2004). Bad politicians. *Journal of Public Economics* 88, 759–782.
- de Janvry, A., F. Finan, and E. Sadoulet (2012). Local electoral incentives and decentralized program performance. *Review of Economics & Statistics* 94(3), 672–685.

- Duggan, J. and C. Martinelli (2017). The political economy of dynamic elections: A survey and some new results. *Journal of Economic Literature* 55, 916–984.
- Duggan, J. and C. Martinelli (2020). Electoral accountability and responsive democracy. *Economic Journal* 130, 675–715.
- Fearon, J. (1999). Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance. In A. Przeworski, S. Stokes, and B. Manin (Eds.), *Democracy, Accountability, and Representation*, pp. 55–97. Cambridge University Press.
- Ferraz, C. and F. Finan (2008). Exposing corrupt politicians: the effect of Brazil’s publicly released audits on electoral outcomes. *Quarterly Journal of Economics* 123, 703–745.
- Ferraz, C. and F. Finan (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review* 101(4), 1274–1311.
- Fisman, R. and M. Golden (2017). *Corruption: What Everyone Needs to Know*. New York, US: Oxford University Press.
- Fudenberg, D. and J. Tirole (1991). Perfect Bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory* 53(2), 236–260.
- Gagliarducci, S. and T. Nannicini (2013). Do better paid politicians perform better? Disentangling incentives from selection. *Journal of the European Economic Association* 11(2), 369–398.
- Galasso, V. and T. Nannicini (2011). Competing on good politicians. *American Political Science Review* 105, 79–99.
- ICIJ (2019). International Consortium of Investigative Journalists, The Bribery Division <<https://www.icij.org/investigations/bribery-division/>>. Accessed May 19, 2020.
- Mattozzi, A. and A. Merlo (2008). Political careers or career politicians? *Journal of Public Economics* 92, 597–608.
- McMillan, J. and P. Zoido (2004). How to subvert democracy: Montesinos in Peru. *Journal of Economic Perspectives* 18(4), 69–92.
- Myerson, R. (2006). Federalism and incentives for success of democracy. *Quarterly Journal of Political Science* 1, 3–23.
- Ogden, B. and R. Pande (2012). Corruption in developing countries. *Annual Review of Economics* 4, 479–509.
- Persson, T. and G. Tabellini (2000). *Political Economics: Explaining Economic Policy*. MIT Press.
- Rogoff, K. (1990). Equilibrium political business cycles. *American Economic Review* 80(1), 21–36.

Rose-Ackerman, S. and B. Palifka (2016). *Corruption and Government: Causes, Consequences, and Reform* (2 ed.). Cambridge, UK: Cambridge University Press.

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