Mackerels in the Moonlight: Corrupt Politicians and Anti–Corruption Reform in Two–Candidate Elections

Haldun Evrenk[1]

Economics Department,

Suffolk University

8 Ashburton Place,

Boston, MA

hevrenk@suffolk.edu

June 6, 2005

[1] I would like to thank my advisor Dilip Mookerjee for his help and encouragement. Gregory Besharov, Martino DeStefano, Hsueh-Ling Hyunh, Bart Lipman, Michael Manove, Zvika Neeman, Rasim Ozcan, Onur Tas, and Jorgen Weibull provided helpful comments. Errors are mine.

Abstract

This paper examines causes of the persistence of corruption among elected politicians and the effectiveness of some commonly discussed anti–corruption reforms. We study a theoretical model of competition between two candidates who differ both in ability and popularity in a probabilistic voting setup. Each candidate proposes a tax rate and a public good level. The elected candidate's ability determines the cost of producing the public good. The budget constraint implies that taxes collected must equal the sum of funds used in public good production plus funds stolen by the elected politician. We solve for the tax rate and public good level chosen by each candidate and how much each candidate decides to steal. We then identify conditions under which (i) imposing constitutional constraints such as tax rate (upper) or public good (lower) limits, (ii) increasing compensation of elected politicians, and (iii) small changes in legal penalties, will reduce corruption and increase voters' welfare. We find that the designers of a successful reform need to have information that is privately held by candidates. The redistributive effects of a reform and how that would affect the popularity of the reform is discussed as well. Finally, we argue that a welfare–improving reform that would reduce the corruption may not be supported by both corrupt and honest politicians.

1 Introduction

According to a survey conducted by the Open Society Institute, three-fourths of Lithuanians believe that either most or all of the politicians in their country are corrupt (The New York Times, November, 7, 2002). Corrupt politicians, as citizens of many other countries would agree, exist beyond the borders of Lithuania as well. John Randolph complained[1] that his Congressional colleague, Henry Clay, "... is so brilliant, so capable, and yet so corrupt that like a rotten mackerel in the moonlight, he both shines and stinks". Depending on the strength of the law enforcement, a politician as well as anyone else may decide to commit a corrupt act. The advantage of democracy over other forms of government is that any politician who wants to be reelected incorporates the effect of his actions on his support from the electorate in subsequent elections. Yet, given voters' dislike of corruption and politicians' desire for reelection, it seems paradoxical that corrupt politicians not only survive in politics, but also win repeatedly. In light of recent findings on the negative impact of corruption on economic growth, the need to understand the role of political institutions in deterring corruption is especially crucial. In this paper we examine conditions under which politicians engage in corrupt behavior, analyze the effectiveness of some commonly discussed anti-corruption reforms, and discuss willingness of politicians to support such reforms.

The argument for the persistence of corruption in democracy is based on the nature of political competition. We formalize the idea that candidates can be differentiated from one another in terms of dimensions other than corruption, e.g., with respect to their ability or popularity with voters. A candidate that is more able or popular than his rival can engage in greater corruption and still remain competitive. This is captured by a model of electoral competition with probabilistic voting, in which voters evaluate candidates in terms of the policies they offer, as well as their intrinsic loyalties. Loyalties may be subject to random, unpredictable swings, implying that even

[1] Quoted in Ehrenhalt (2002).

candidates identical in ability and *ex ante* popularity can afford to engage in corruption and yet be reelected with positive probability. In the model, candidates propose fiscal policy platforms, where the amount they steal from the public treasury is implicitly defined by the difference between revenues and public good costs. Candidates thus choose the amount they steal along with the tax rates they propose. Corruption in equilibrium is increasing in heterogeneity among candidates with respect to their popularity and ability, and in the extent of randomness in voter loyalties.

An analogy to the context of price competition between two firms helps explain this point. Consider two firms that select price and quality of their respective products, in a context where there is uncertainty about their relative demands. Bertrand competition will then allow firms to price above cost and select suboptimal qualities.

Political corruption with electoral competition in a probabilistic voting setup was considered earlier by Brennan and Buchanan (1980), Polo (1998) and Persson and Tabellini (2000). Our model extends and generalizes their work in a variety of directions, and more importantly allows us to consider, (and compare the effectiveness of), different anti-corruption reform proposals within the same framework. In comparison with Brennan and Buchanan, for instance, we assume that theft is not the only source of rents for elected officials. Power (ego-rents) may be valued for its own sake. Besides, salaries and perquisites of office represent a source of legal rents that represent a policy parameter. This difference in assumptions about the motivation of politicians has important implications for the effects of different kinds of policies on corruption and welfare.

Consider the effects of constitutional constraints on tax rates that Brennan and Buchanan (1980) promote as instruments for reducing corruption. Their argument is based on the assumption of a (Leviathan) government, which faces no competition and for whom theft constitutes the sole source of rents. We investigate the effects of tax constraints in a setting with duopolistic competition and multiple sources of rents. We find that tax constraints are effective in the case where competing candidates are *ex ante* identical, but may be counterproductive when they are not.

The analogy with market competition is again helpful in explaining this. Although they consider an example of elected politicians in the beginning of their book, for all parts the Brennan-Buchanan theory is analogous to a monopolist, who does not face any political competition, who selects a quality (pubic good level) and charges the highest price (tax) that leaves the buyer indifferent between buying the good and not. In such case, the only incentive to produce any quality is that quality increases willingness to pay (taxable income level). The tax rate constraints in that setup is the same as regulating the price charged by the monopolist, as studied in Spence (1975). When tax rates and public good levels are determined by a candidate who has to win, maybe an imperfectly competitive, election rather than a Leviathan, the nature of the problem changes. Now, we have two sellers with possibly different marginal costs (ability), who both decide on the quality and the price of the good that they are selling without knowing their relative demands. What prevents each seller from providing low quality is not a decrease in the overall willingness to pay (because we assume that taxable income level is independent of public good level), but the possibility of rival firm producing a lower price-quality ratio and stealing the demand, electoral competition. The question of whether imposing a price ceiling in such a duopoly results in higher consumer welfare is, however, more complicated than the case of monopoly. In a duopoly, when firms have different marginal costs, they end up with price and quality levels that are different from each other. Unlike Polo (1998) and Persson and Tabellini (2000) who assume quasi–linear preferences, we study voters who have separable but not necessarily quasi–linear preferences. That generality allows us to notice that a constitutional constraints enforcing a "lean" government is not the only solution to political corruption. Another type of constitutional constraint, would be a binding lower bound on the public good level, –a minimum quality regulation using the market analogy, where any leader has to provide at least a given level of public good when he wins the election. As in tax rate limits, we find that when candidates are identical and marginal utility of consumption is decreasing, a constitution that enforces candidates to propose a public good level that is larger than they would propose otherwise. In order to calculate the appropriate constraints in both cases, however, drafters of a constitution will require information that is privately held by the (current and future) candidates, such as how able, popular, and honest they are. If such information is held by the drafters of the constitution, then by employing both type of constraints corruption may be eliminated at no cost. When drafters do not have such information, than constitution is just another level where voters with conflicting interests try to tilt the outcome towards their most preferred policy platform, since the poor prefer a constitution that enforce a minimum public good level, and the rich prefer tax rate limits, even when these constraints decrease aggregate welfare. Hence, a qualified majority agreeing on some type of constitution is difficult to gather. We also find that when the candidates are not identical, either of the constraints may have the *opposite* effect of raising corruption and lowering welfare.

A commonly proposed reform to reduce the illegal appropriation of public funds is to increase the legal compensations of politicians, e.g., as suggested by Becker and Stigler (1974). In the market analogy, this corresponds to a prize (financed by consumers) given to the firm with the highest sales. In that case, a firm has incentives to increase its sales, which can be accomplished by proposing a better price-quality ratio, i.e., lowering the level of corruption. Increasing the wage is, however, costly, since customers eventually finance the wage bill. We find that when candidates are identical and there are no legal incentives for corruption the benefit of wage increase (lower corruption) justifies the cost. But in the presence of legal penalties, this is not always so. The distributional impact of wage increases is also different from those of tax rate limits, i.e., most of the burden of the former is borne primarily by the rich, the latter by the poor. For the wage reform that would implement a second–best optimum we need private information about the cost of stealing for a candidate.

When legal incentives are very strong (a high probability of getting caught and resultant harsh penalties), a candidate will remain honest no matter what the electoral incentives. When the legal incentives are weak, the political competition game may have multiple (two) equilibria: either both candidates stay honest or at least one steals. In terms of anti-corruption effects, one has to be careful. Since the legal incentives reduce the expected rents from the office, a small increase in legal penalties can raise corruption and lower welfare.

Finally, we consider the incentives of candidates to propose an anti-corruption reform. When both candidates are corrupt, it is not surprising that they would have no interest in proposing a reform that would eliminate some of their rents. We argue that even an honest candidate may not want to support such a reform if his opponent is corrupt, since it removes an important source of his competitive advantage.

In summary, modelling it as an agency problem, our model contributes to an understanding of persistence of corruption in democracies in a variety of ways. Political corruption may stem from factors that are beyond the control of constitution designers, such as voter loyalty and candidate heterogeneity. There is no such thing as the best anti-corruption reform. Many reforms commonly suggested may increase corruption under certain conditions. It is especially difficult to design an effective reform when one candidate is honest. Even when a reform could improve voter welfare, implementation requires information that reformers may not have. Different reforms have different distributional effects and hence the voters may not unanimously support a welfare improving reform. And even when there exists a welfare improving reform that is supported by electorate, it may not be proposed by any of the politicians, corrupt or honest, competing for public office.

Section 2 presents the model without law enforcement. In section 3, we prove existence and uniqueness of Nash Equilibrium. In section 3 we also present comparative statics, an example using quasi–linear utility function, and a discussion and generalization of results from the literature. In section 4, we discuss constitutional constraints on tax rates. In section 5, we introduce law enforcement, and then discuss the effect of higher wages and higher legal penalties. In Section 6, we present other approaches to model the agency problem in politics. We discuss that the approach we follow is better in evaluating different reforms, since it models strategic interaction between candidates. In Section 7, we discuss the extensions of the model and conclude. Most of the proofs are presented in the Appendix.

2 The Model

Let us imagine a society where each voter i has income Y_i, out of which he pays an income tax at flat rate τ and consumes the rest. The income in society is distributed over $[Y_{\min}, Y_{\max}]$ with measure $\mu(Y_i)$. The size of the population, N, and the average income $y = \frac{1}{N}\int Y_i d\mu(Y_i)$ are both normalized to one. There are two political agents (candidates) who compete for votes. Candidate $j \in \{1, 2\}$ chooses a policy platform, i.e., promises a tax rate, τ_j, and a per capita public good level, G_j. He implements the promised policy platform when he wins the election.

Each voter i has preferences over his consumption of the private good, $c_i = (1-\tau)Y_i$, and the public good, G. Preferences over consumption are represented by a separable utility function

$$U(c_i, G) = I(c_i) + H(G),$$

where $I()$ and $H()$ are two strictly increasing, C^2, and concave functions from R_+ to R with at least one of them being strictly concave. Strict concavity ensures the single–peaked preferences over tax rates. Unless we use quasi–linear form to simplify calculations[2], to ensure interior outcomes we

[2] When we use quasi–linear form we can explicitly calculate the condition for interior solutions, see footnote 9 on page 15.

assume

Assumption (No Extreme Platforms): The marginal utility of consumption converges to infinity as the good consumed goes to zero, i.e., $\lim_{c\downarrow 0} I'(c) = \lim_{G\downarrow 0} H'(G) = \infty$.

The voters have preferences over the characteristics of political agents as well. The utility of voter i from agent j is

$$U_i^j = U(c_i^j, G_j) + (j-1)\xi_{i2}. \tag{1}$$

where c_i^j denotes consumer i's consumption when the policy platform of candidate j has been implemented. Following the probabilistic voting literature, we assume that ξ_{i2} can be written as $b + b_2 + b_{i2}$, where b is the electorate's average bias in favor of candidate 2 which is known *ex ante*. A positive (negative) b means candidate 2 is more (less) popular. From the candidates' point of view, the other terms in voter preferences, b_2 and b_{i2}, are random variables uniformly distributed on (respectively) $[\frac{-1}{2g}, \frac{1}{2g}]$ and $[\frac{-1}{2f}, \frac{1}{2f}]$. The first term, b_2, reflects uncertainty about a correlated preference shock, while the second term, b_{i2}, reflects an idiosyncratic shock on individual i's preferences. We assume that these preference shocks are statistically independent of each other and of b, i.e., $E[b_2 \mid b, b_{i2}] = 0$ and $E[b_{i2} \mid b, b_2] = 0$.

We assume sincere voting: Voter i votes for candidate j when $U_i^j > U_i^k$. If $U_i^j = U_i^k$, then each candidate gets the vote with equal chance. Both candidates run for the same position, which we call the position of leader. The leader produces the public good from the available public funds using a linear technology, that depends on his ability. The (non-verifiable) ability levels of each candidate, a_j, can be different. The higher is the ability of the leader, the lower is the cost of producing any level of public good. The available public funds that can be used by the leader in the production of public good is equal to collected tax revenues minus the salary of the leader, (denoted by w), and an amount that he chooses to steal. Let S_j denote the public funds stolen. The per capita public good delivered when candidate j is the leader is

$$G_j = a_j(\tau_j - w - S_j). \tag{2}$$

We assume that a politician has to offer a non-negative public good level. The set of feasible policy platforms for a candidate is any tax rate from the interval $[w, 1]$ and any level of stealing that provides at least a zero public good level. Then the strategy space of candidate j is $\Sigma_j = \{(\tau_j, S_j) : \tau_j \in [w, 1] \text{ and } S_j \in [0, \tau_j - w]\}$.

When a candidate wins the election, he is going to get legal rents and will have access illegal rents. In addition to salary, legal rents include ego rents, E.[3] Following the corruption literature, we assume that there are deadweight losses from illegal rents: when the leader diverts a dollar from the public budget, a fraction $1 - L_j$ will be wasted, so the leader will appropriate only $L_j < 1$. This assumption, known as "leakage" or "deadweight loss of corruption" in the literature, reflects the possibility that the leader should share the illegal rents with some of his political supporters or with corrupt bureaucrats, or that there is a moral cost of stealing. When the leader is what Rose-Ackerman (2001) calls "pathologically honest," we have $L_j = 0$.

We assume that candidates are expected rent maximizers. The rents that candidate j receives conditional on being elected are

$$R_j(S_j) = w + E + L_j S_j. \tag{3}$$

[3] We consider changes in the wage as a possible way to reduce the politician's incentives to steal; hence, we want to seperate rents into ego rents, rents that can not be (at least easily) designed and wages, rents that can be perfectly controlled, at the cost of higher taxes.

The probability that j wins the elections when he competes with k is[4]

$$\rho_j = \frac{1}{2} + g[\mathbf{E}[U(c_i^j, G_j) - U(c_i^k, G_k)] + \mathbf{P}_j], \tag{4}$$

where $\mathbf{P}_j = 2(j - \frac{3}{2})b$ is the effect of *ex-ante* popularity advantage of candidate j and the expectation is taken with respect to μ. Note that ρ_j can also be written as a function of $(\tau_j, S_j, \tau_k, S_k)$, i.e.,

$\rho_j = \frac{1}{2} + g[\mathbf{E}[U((1-\tau_j)Y_i, a_j(\tau_j - w - S_j)) - U((1-\tau_k)Y_i, a_j(\tau_k - w - S_k)] + \mathbf{P}_j]$.

2.1 Agency Problem

Let us set the outside option for candidates at zero. Then candidate j selects a policy platform[5] to maximize his expected rents:

$$\max_{(\tau_j, S_j) \in \Sigma_j} \ \rho_j(\tau_j, S_j, \tau_k, S_k) R_j(S_j). \tag{5}$$

Let (τ_j^*, S_j^*), $j = 1, 2$ denote a Nash equilibrium:

$$(\tau_j^*, S_j^*) \in \arg\max \ \rho_j(\tau_j, S_j, \tau_k^*, S_k^*) R_j(S_j). \tag{6}$$

To evaluate an anti–corruption reform, we look at its effect on the equilibrium outcome, and check whether voters are better off or not in the new equilibrium. Since there are a continuum of voters with differing preferences on policy, we use an aggregate measure of voters' welfare. The voters' expected (purely utilitarian) welfare, $\mathbf{E}[\mathbf{W}]$, as a function of policy platforms, popularity and probability of winning the election of each candidate can be written as[6]

[4]See Appendix.
[5]From the candidate's point of view (τ_j, G_j) and (τ_j, S_j) are interchangeable.
[6]See Appendix.

$$\mathbf{E}[\mathbf{W}] = \mathbf{E}[U_i((1-\tau_2)Y_i, G_2(\tau_2, S_2))] + b + \frac{1}{2g}(\rho_1)^2. \tag{7}$$

The policy platform, (τ_j^0, S_j^0), which maximizes $\mathbf{E}[\mathbf{W}]$ when adopted by candidate j will be referred as the first–best policy platform for candidate j. It is easy to check that the first best policy platform for candidate $j \in \{1,2\}$ involves zero corruption and a tax rate which maximizes $\mathbf{E}[U_i((1-\tau_j)Y_i, G_j(\tau_j))]$, the average utility of the electorate. The optimality of zero corruption is intuitive: Given the tax rate, less stealing means higher public goods delivered. On the other hand the optimal tax rate depends on our choice of aggregate welfare function.

2.2 Nash Equilibrium

Conditional on S_j, τ_k, S_k, candidate j selects τ_j to maximize ρ_j. This implies,[7] given (5), that he selects τ_j to maximize average voter utility conditional on S_j. So, in our model the agency problem exists, if at all, in only one dimension, i.e., stealing. This is due to the assumptions that candidates are rent-maximizing, that voters are well informed, and that there are no special interest lobbies. This observation also simplifies the analysis, since the strategy space reduce to the level of stealing alone.

To see when we have an agency problem, we need to consider the first order condition with respect to stealing. The marginal expected utility of stealing for candidate j,

$$gR_j \frac{\partial \mathbf{E}[U(c_i^j, G_j)]}{\partial S_j} + L_j \rho_j, \tag{8}$$

should be less than or equal to zero in equilibrium. It is equal to a weighted average of two marginal gains: (i) the average marginal disutility of voters from corruption weighted by gR_j, and (ii) the

[7] See Appendix.

marginal utility from a stolen dollar conditional on being elected, weighted by the probability of winning election, ρ_j. If (8) is always negative, reducing S_j makes the candidate better off. Then candidate sets $S_j = 0$, and there is no agency problem. When (8) is positive at $S_j = 0$, then candidate j keeps stealing until (8) becomes zero.[8] Let $s_j^0(S_k)$ denote the best response of candidate j to a rival stealing S_k. We show in the Appendix that the corruption levels of candidates are strategic complements, i.e., $\frac{\partial s_j^0(S_k)}{\partial S_k} \geq 0$, and that the best response functions are continuous and they intersect only once. We therefore obtain

Theorem 1 *There exists a unique pure strategy Nash equilibrium for the political competition game.*

Depending on the parameters the outcome is (i) overall corruption (both candidates steal), (ii) partial corruption (only one candidate steals), or (iii) no corruption (both candidates offer policies that maximize voters' welfare). Figure 1 describes four different subsets of parameters that give rise to these different outcomes. The thick curve is $s_1^0(S_2)$. In graphs (a) and (c) both candidates steal. Only Candidate 1 steals in (b). In (d) none of them steals. Note that to determine the outcome of the game, we need to know (i) whether $s_j^0(0) > 0$ or not, and (ii) if $s_j^0(0) = 0$ for at least one candidate, then whether $\underline{S}_j < s_j^0(0)$ or not, where $\underline{S}_j = \inf\{S_j \mid s_k^0(S_j) > 0\}$.

Which subset of parameters gives rise to which of the graphs in Figure 1?. We do not have closed form solutions for those sets. Incorporated into our model $I(.)$ and $H(.)$ are also parts of the parameter space, which makes the conditions particularly messy, (see the Appendix). To be able to convey the intuition about which parameters increase/decrease incentives to be corrupt, one can either (i) choose a "nice" functional form for U, where these conditions become more tractable, or (ii) look at the comparative statics. We do both.

[8] We show in the appendix that (8) is strictly decreasing in S_j.

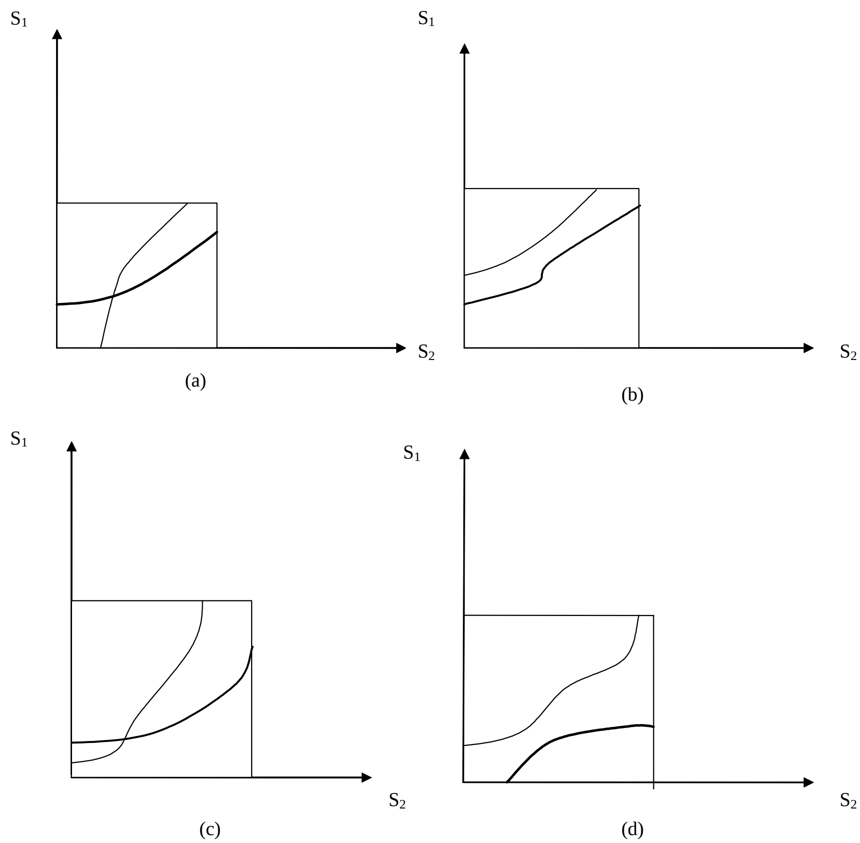

Figure 1: Nash Equilibria

2.3 An Example: Quasi–linear Utility

Assume[9] that $U = c + 2\theta\sqrt{G}$, then candidate j's best response to candidate k is

$$s_j^0(S_k) = \max\{0, \frac{1}{4g} + \frac{S_k}{2} + \frac{A_j}{2} - \frac{K_j^0}{2}\},$$

where $A_j = \theta^2(a_j - a_k) + \mathbf{P}_j$ is the relative advantage that the candidate j has and K_j^0 is equal to the illegal rents that are payoff equivalent to legal rents, $\frac{w+E}{L_j}$. The unique Nash Equilibrium of

[9] Note that quasilinear form does not satisfy our assumption on infinite marginal utility at zero zorruption. Since it simplifies calculations considerably and earlier studies, both Polo (1998) and Persson and Tabellini (2000) use quasilienar form, we provide that example. On the other hand the policy platforms proposed by candidates in equilibrium may involve 100% taxes when utility is quasilinear. To prevent that we need to assume that neither the overall uncertainty nor one candidate's relative advantage is too high, i.e., $\frac{3}{4gy} + \theta^2(3a_j - a_k) + 2\mathbf{P}_j - W + E(\frac{1}{L_j} + \frac{1}{2L_k}) < \frac{3}{2}$ for $j \in \{1, 2\}$.

the political competition game is

(i) $S_j^* = \frac{1}{2g} + \frac{A_j}{3} - \frac{w+E}{3}(\frac{2}{L_j} + \frac{1}{L_k})$ for all $j \in \{1,2\}$ if and only if we have either $\frac{1}{4g} + A_j - K_j^0 > 0$ for all $j \in \{1,2\}$, or $\frac{1}{4g} + A_j + K_j^0 > 0$ for only only one $j \in \{1,2\}$ but we have $(w+E)(\frac{1}{2L_j} + \frac{1}{L_k}) < \frac{3}{4g}\frac{\theta^2(a_k - a_j)}{2}$ for $k \neq j$.

(ii) $S_j^* = \frac{1}{4g} + \frac{A_j}{2} - \frac{K_j^0}{2} > 0$, $S_k^* = 0$ if and only if $\frac{1}{4g} + A_j + K_j^0 > 0$ for only $j \in \{1,2\}$ and we have $(w+E)(\frac{1}{2L_{j'}} + \frac{1}{L_k}) > \frac{3}{4g}\frac{\theta^2(a_k - a_j)}{2}$ for $k \neq j$.

(iii) $S_1^* = S_2^* = 0$ iff $\frac{1}{4g} + A_j - K_j^0 \leq 0$ for all $j \in \{1,2\}$.

As both Polo(1998) and Persson and Tabellini (2000) note, the quasi–linear utility function implies that the effects of higher corruption will be only higher tax rates, while public good levels are always first-best. Also, the slope of the reaction function that we find above is independent of the parameters of the model. Both of those results are driven by the quasi–linearity. In Appendix, we show how the effect of corruption on tax rates and public good levels differ for separable but not necessarily quasi–linear utility functions. For such utility functions, the slope of reaction function is not necessarily independent of parameters of the model either. However even when we consider that larger set of utility functions, the *direction* of comparative statics does not change. In the next section we present comparative statics again for general U.

2.4 Comparative Statics and Relation to Previous Literature

Let us calculate the effect of a small change in one of the parameters, g, b, E, a on the reaction functions. Then, we show that the results of previous studies can be considered as applications of those comparative statics in special environments.

Lemma 1 *Consider S_k such that candidate j's best response is to steal, ($s_j^0(S_k) > 0$). Any of the following would cause j to steal more, (shift $s_j^0(S_k)$ to the right):*

- *an increase in the uncertainty about popularity, $\frac{1}{g}$,*

- *an increase in the popularity of candidate,* $\mathbf{P}_j$,

- *a decrease in the ability of the rival candidate,* a_k, *and*

- *a decrease in ego rents,* E.

Proof. When $s_j^0(S_k) > 0$, we have (8)=0. Then applying the implicit function theorem, the above results are obtained. ■

Let us now examine how the comparative statics in Lemma 1 relates to previous literature on agency problem in politics. Brennan and Buchanan (1980), in their pioneering study of political economy of taxation, consider the state, for most part, as a dictator who uses his powers to further his own private interest *and* does not face any political competition. To justify that assumption, they begin with an election example: two competing politicians offer policies on how to distribute \$300 among three voters. When there is uncertainty on vote shares, they claim that "each party would rationally appropriate some of the \$300, even where the other party did not"(Brennan and Buchanan (1980), p 22). After noting that when the aggregate vote shares are stochastic, "the multi-party competition and more importantly the simultaneous announcement of policies is not fully constraining as Downs claims," Brennan and Buchanan build their theory of "the foundations of a fiscal constitution." However, their conclusion that candidates necessarily steal, is an outcome of specific assumption that there are no legal rents.

Theorem 2 ***(Brennan and Buchanan (1980), Polo (1998))*** *Suppose that candidates are identical,* $(a_1 = a_2,\ b = 0)$, *are not pathologically honest,* $(L > 0)$, *and there are no legal rents,* $(w = E = 0)$. *If there is overall uncertainty,* $(\frac{1}{g} > 0)$, *then* $S_j > 0$ *in equilibrium.*

Proof. Under the above conditions, $R_j = LS_j$. Then (8) can be written as

$$gL_jS_j\frac{\partial \mathbf{E}[U(c_i^j, G_j)]}{\partial S_j} + L\rho_j.$$

When there are no legal rents, the only source of rents is corruption. Hence there is no point of winning the election if a candidate cannot acquire any illegal rents, i.e., the weight on voters' disutility on corruption is zero when S_j is zero. Then at $S_j = 0$, the marginal utility of corruption for candidate j is $L\rho_j$, which is strictly positive when we have $L > 0$. Thus unless the candidate is "pathologically honest", we always have $s_j^0(0) > 0$. Hence the unique equilibrium outcome is corruption by both candidates. ■

As Theorems 3 and 4 reveal, uncertainty about the outcome of elections is neither necessary nor sufficient for corruption to occur. Actually as we discuss after Theorem 4, uncertainty is necessary for corruption not to occur when there are legal rents and candidates are not identical. The effect of uncertainty on electoral incentives of a candidate can be seen from (8): The larger the uncertainty, the smaller g, and the less important the policy issues for winning the elections, hence less weight on voters' welfare. Theorem 2 does not require a specific utility function. Also, as far as there are no legal rents, Theorem 2 would hold even if candidates were not identical.

Polo (1998) does not mention the work by Brennan and Buchanan (1980), but his model does provide a well specified environment for the phenomenon first discussed by them. In Polo, the process that leads to uncertainty in vote shares, probabilistic voting, is explicitly modelled. The policy is two dimensional, preferences are quasi–linear, $U(c_i, G) = c_i + H(G)$ where $H(.)$ is strictly concave. As Brennan and Buchanan, Polo also assumes expected rent maximizing candidates and no legal rents. In Polo's model, popularity differences among candidates are allowed. He finds that such differences are important for candidate's incentives to steal.

Theorem 3 ***(Polo (1998))*** *Suppose that there is no ability difference between the candidates,* $(a_1 = a_2)$, *and no overall uncertainty about candidate preferences,* $(\frac{1}{g} = 0)$. *If one candidate is more popular than the other,* $(b \neq 0)$, *then (only the) popular candidate will steal.*

Proof. When $\frac{1}{g} = 0$, there is no uncertainty about the winner of an election. The candidate who proposes a policy platform that provides higher utility to the median voter wins the election for certain. Suppose that both candidates adopt the (identical) policy platform that is most preferred by median voter. Then the more popular candidate, say k, will certainly win. But unless he is "pathologically honest", he could afford to steal a little and increase R_k without risking his victory in elections, i.e. without lowering ρ_k. Since that would increase his expected rents, he will steal in the equilibrium.[10] ∎

When, in addition to popularity advantage there is uncertainty about voter loyalty swings, the incentives to steal increase even further. The intuition for the effect of greater popularity is that it permits that candidate to steal more without making himself inferior to another candidate. This helps explain the paradox that pointed out by Kurer (2001) as well as by many others, i.e., some corrupt politicians are also quite popular. Our model would explain this by reversing the causality implicit in the expression. Politicians are not popular because they are corrupt, but rather that popular politicians can afford to be corrupt. The logic of Theorem 3 will apply when the politicians differ not in terms of their popularity but in their ability. The one with the higher ability is able to get some "Ricardian rents" in the equilibrium even when there is no uncertainty.

Persson and Tabellini (2000) discuss the agency problem in politics employing a probabilistic voting model and a quasi–linear utility function as Polo (1998) but they consider ego rents as well.

Theorem 4 ***(Persson and Tabellini (2000))*** *Suppose that* $U = c + H(G)$*, candidates are identical,* ($a_1 = a_2$ *and* $b = 0$)*, there is no wage, but there are ego rents coming from the office,* $(E > 0)$*. Then, there is political corruption iff* $E < \frac{L}{2g}$.

[10] Note that we need some discreteness in the strategy space, otherwise the optimum best response, and the equilibrium do not exist.

Proof. When both candidates are identical, the equilibrium (which is unique by Theorem 1) is symmetric, so $\rho_j = \frac{1}{2}$. Then (8) can be written as

$g(E + LS_j)\frac{\partial \mathbf{E}[U_i(c_i^j, G_j]}{\partial S_j} + L\frac{1}{2}$.

Note that if $E > \frac{-L}{2g\frac{\partial \mathbf{E}[U_i(\tau_j^o, G_j^o)]}{\partial S}}$, then (8) is negative at $S_j = 0$, i.e., $s_j^0(0) < 0$ for both candidates. Then $S_1^* = S_2^* = 0$ is (the unique) equilibrium. For the special case of $U = c + H(G)$, we have $\frac{\partial \mathbf{E}[U_i(\tau_j^o, G_j^o)]}{\partial S} = -1$. ■

The result that when ego rents are high enough, there exists an equilibrium without corruption applies for any utility function. That result can be extended to heterogeneous candidates: Whenever the ego rents are sufficiently high *and* there is uncertainty about voter loyalty, $\frac{1}{g} > 0$, both candidates choose not to steal, despite any advantage that one may have over the other.

We have shown which factors lead to political corruption. Now we will address what can be done about it.

3 Constitutional Constraints as Anticorruption Reform

Brennan and Buchanan (1980) discuss how an individual member of society who decides behind a "veil of ignorance" would like to impose constraints on the political decision-making process or on the domain of the political outcomes to maximize the expected utility of his future selves. As a way to reduce political corruption, we consider first constitutional constraints on tax rates[11] as discussed in chapter 10 of Brennan and Buchanan (1980), and then another type of constitutional constraints, lower bounds on public good levels, which, as far as we know, have not been discussed before.

In previous section we find that aggregate uncertainty does not necessarily lead to political corruption. Our point in this section is that even when it does lead to corruption in democra-

[11] An example is the Proposition 13, which was approved by voters in California in 1978. It restricts the tax on real property to 1 percent of market value.

cies, proposed remedies (constitutional constraints) should be discussed in a model of political competition, not using a model of Leviathan. The following is an attempt in that direction.

Let us first assume that the parameters of the model are such that in equilibrium at least one politician steals, so electoral incentives are not enough to deter political corruption. Now we can study how the constitutional constraints interact with electoral incentives. Let us first consider the constitutional constraints on tax rates.

Proposition 1 *It is impossible to implement the first best policy platform, (τ_j^0, G_j^0) through imposing a tax rate constraint on candidate j.*

Proof. The first order condition with respect to taxes in a Nash equilibrium is

$$gR_j \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - w - S_j))]}{\partial \tau_j} - \lambda_j = 0,$$

where λ_j is a Kuhn-Tucker multiplier satisfying $\lambda_j(\tau_j - \overline{\tau}) = 0$. Suppose there exists a $\overline{\tau}$ that implements the first best. Then $\lambda_j > 0$. But in a first-best the expected marginal utility of electorate with respect to tax rate should equal zero. Contradiction. ■

The fact that tax rate constraints cannot implement the first best does not mean that they are useless. It simply means that these constraints may provide a benefit, yet they have a cost as well. Our second question is about the second–best: When does a tax rate constraint increase voters welfare in a society with political corruption?

3.1 Constitutional constraints when candidates are identical.

Let us consider two (ex-ante) identical candidates competing with each other in a country where the tax rate that a politician can propose is constrained to be less than or equal to $\overline{\tau}$. Using (8), we have the equilibrium level of corruption by each candidate, $S(\overline{\tau})$, given by

$$-g(S(\overline{\tau})+w+E)[aH'(\overline{\tau}-w-S(\overline{\tau}))]+\frac{1}{2}L\leq 0.$$

The effect of a tax rate limit on corruption level can be calculated as,

$$S'(\overline{\tau})=\frac{Ra^2H''(G)}{Ra^2H''(G)]-LaH'(G)} \text{ for } S>0,$$

where $R=\frac{L}{2agH'(G)}$, and $G=a(\overline{\tau}-w-S(\overline{\tau}))$. Note that whenever $H(.)$ is a strictly concave function, the derivative, $S'(\overline{\tau})$, is strictly positive, hence reducing the tax limit would reduce stealing. But also note that, although positive, the derivative is less than one: The decrease in corruption comes with a cost, a reduction in public good level. So, the net effect of tax rate constraints on voters' welfare is not clear and needs to be calculated. Using (7), the effect of an incremental change in $\overline{\tau}$ on voters' welfare can be written as,

$$\frac{\partial\mathbf{E}[\mathbf{W}]}{\partial\overline{\tau}}=aH'(G)\frac{1}{1-[\frac{H''(G)}{2g(H'(G))^2}]}-\mathbf{E}[Y_iI'(Y_i(1-\overline{\tau})].$$

When $H(.)$ is strictly concave $\frac{\partial\mathbf{E}[\mathbf{W}]}{\partial\overline{\tau}}$ is always negative at unconstrained political equilibrium with identical corrupt candidates, i.e., at (τ^*,G^*), we have $aH'(G^*)=\mathbf{E}[Y_iI'(Y_i(1-\tau^*)]$. Thus we have the following result.

Proposition 2 *Whenever marginal utility from public good is decreasing, ($H(G)$ is strictly concave) and both candidates are identical and corrupt, there always exists a constitutional constraint that enforces both candidates to offer a tax rate that is lower than τ^* and that constraint is both corruption reducing and welfare-improving.*

The intuition is that the tax rate constraints lower G, for a given level of corruption, raising marginal utility of public good. This increases the voters' disutility from corruption. Hence the marginal utility of stolen funds for a candidate becomes negative at S^*. A candidate reduces the level he steals because now the cost of stealing in terms of votes foregone is higher. This is, as far as we know, the first formal analysis of constitutional tax rate limits in an electoral setup. The difference between Brennan and Buchanan's analysis is not simply that we have two candidates where they have one Leviathan. What derives their result is the assumption that higher public good levels increase the taxable income, yet higher taxes reduce taxable income, (Laffer curve argument) and the assumption of monopoly power of politician. Our argument incorporates the effect of political competition. In our model, the elasticity of taxable income with respect to either public goods or tax rates is zero, so there is no Laffer curve, yet limits increase voters' welfare. Although imperfect, the electoral competition is what derives our result and make limits work.

As we mention earlier, the market analogy is helpful in thinking about the political corruption. Using that analogy, constitutional constraints are just tools for regulating a market. What we have shown that if we have a duopoly with a special demand in that market, a price cap would increase consumer welfare. Then one wonders, since in our model each firm chooses both its price and quality, what would be the effect of a minimum quality regulation on voters' welfare? That is to ask, how voters' welfare would change if we have a constitutional constraint that requires each candidate to provide at least a minimum level of public good, $\underline{G}$? As in tax limits, let us first note that the first best can not be implemented using a minimum public good level constraint. For a second-best, note that in symmetric equilibrium we have

$$-g(S(\underline{G})+w+E)\mathbf{E}[Y_i I'(1-\frac{1}{a}\underline{G}-w-S(\underline{G}))]+\frac{1}{2}L \leq 0.$$

The effect of public good constraints on stealing is given by

$$S'(\underline{G}) = \frac{R\mathbf{E}[Y_i^2 I''(1-\tau)]}{-R\mathbf{E}[Y_i^2 I''(1-\tau)] + L\mathbf{E}[Y_i I'(1-\tau)]} \text{ for } S > 0.$$

where $R = \frac{L}{2g\mathbf{E}[Y_i I'(1-\tau)]}$ Now, whenever the marginal utility from consumption is decreasing, the derivative, $S'(\underline{G})$, is negative, hence to reduce stealing we need higher levels of public goods, i.e., another solution to political corruption may be a constitution that enforces a large government! When we calculate the total effect of public good limits on voters' welfare, we find that

$$\frac{\partial \mathbf{E}[\mathbf{W}]}{\partial \underline{G}} = aH'(\underline{G}) - \mathbf{E}[Y_i I'(Y_i(1-\tau)] \frac{1}{1 - \frac{\mathbf{E}[Y_i^2 I''(1-\tau)]}{2g(\mathbf{E}[Y_i I'(1-\tau)])^2}}.$$

Note that, whenever $I(.)$ is a strictly concave function, $\frac{\partial \mathbf{E}[\mathbf{W}]}{\partial \underline{G}}$ is larger than zero at unregulated political equilibrium, (τ^*, G^*). So, a constitution that enforces any candidate to propose a public good level that is larger than he would propose without any such constraint would increase the voters' welfare. Or more formally,

Proposition 3 *Whenever marginal utility from private good is decreasing, ($I(c)$ is strictly concave), and both candidates are identical and corrupt, there always exists a constitutional constraint that enforces both candidates to offer a public good level that is higher than G^* and that constraint is both corruption reducing and welfare-improving.*

The intuition for why lower bounds for public good levels work is similar to the intuition for tax limits. The public good limit increases the taxes, reducing after tax income, and increasing the marginal utility of private good consumption for a voter, $I'(.)$. That, in return increases a voter's disutility from corruption, and makes it more costly for a candidate.

We find that when candidates are identical voters may regulate the political market through a price cap or a minimum quality requirement. Which type of regulation is better for voters? In general using *both* constraints would be much better than using only one. As we discuss below, in that case one could implement the first–best. More importantly there are two issues that one has to consider when discussing these constitutional constraints. These are (i) redistributive effects of these constraints, i.e., the burden each constraint puts on different income groups in society, and (ii) (lack of) access to private information, as in any case of regulation.

Suppose that we can achieve the same welfare level whether we use the tax limits or the minimum public good levels, but want to use only one of them. Furthermore suppose that the most preferred tax rate of a voter decreases with his income. Then, despite the fact that overall welfare is the same under both reforms, tax limits will be strictly preferred by rich and public good limits will be strictly preferred by poor. To see that, note without any limits low (high) income voters actually would prefer higher (lower) taxes and higher (lower) public good levels given the level of corruption in the status quo. If they can move towards lower corruption in one of these two ways; (i) higher τ and G, and, (ii) lower τ and G, low income voters will choose (i), where high income voters choose (ii). Because of the distributional effects of these two reforms the identity of designers of the constitution matters. To see that note that, when the rich (poor) design the constitution, they have the incentives to choose the tax rate constraints (minimum public good level constraints) even when such constraints are welfare–reducing. That makes our problem different from other kinds of principal–agent problems, because there is a conflict of interest among principals, which in turn makes an agreement by a qualified majority difficult if not impossible. The fact that elections do not give enough incentives to agents to maximize aggregate welfare does not necessarily mean if some members of society write a constitution it will always be a welfare improvement.

If the drafters of the constitution have a self interest to pursue as candidates do, then why not to allow everyone to participate in writing the constitution, –in which case an agreement that would maximize the median voter's welfare would emerge? One practical reason is the informational requirements, i.e., drafters should have information on issues such as how easy to steal and how able the future politicians are. Such information may not be held by every voter. In our model, if we know the parameters, we can calculate the cost and benefit of constraints and the optimal constraint, as well as the necessary information to set the optimal constraints.[12] Consider the necessary information required to set the optimal tax rate constraint. Suppose that the writers of the constitution know both U (to know preferences is not enough, one needs to know *the* utility function) and that all future candidates are going to be identical and corrupt. Are they able to set the correct constraints with this information? The answer is no. The optimal tax rate constraints in a democratic society depends on the ethics and ability *levels* of all future candidates as well. A quote from Hume in Brennan and Buchanan (1980) –"in contriving any system of government, and fixing the several checks and controls of constitution, every man ought to be suppose a knave, and to have no other end, in his all actions, than private interest" – makes us think that the optimal rules should be designed under the assumption that all politicians are totally corrupt, not because they will be, but if we are protected from the worst then we are protected from all.[13] This idea would be correct only when such restrictions are costless. However as Proposition 1 shows, tax rate constraints are costly in terms of lowering public good level. Whenever candidates are not as corrupt as the designers of the constitution assume, then tax rates prescribed by drafters will be set too low. We expect the problem of information to be severe. Otherwise, both for our analysis and for Brennan and Buchanan's model, when designers have the information about all, current and future, (identical) politicians, and are maximizing the aggregate voters' welfare, why to settle with the second–best? In that case, the designers may simply specify *both* the tax rate and public good limits at the optimal levels and implement the first best, (τ^o, G^o).

[12] Let us assume that $U = c + 2\theta\sqrt{G}$ and that designers maximize the expected voters welfare, and candidates are identical (and corrupt). Then it is easy to show that the tax rate constraint that maximizes voters' expected welfare is $\overline{\tau} = \tau^o + S^* - \frac{1}{8\theta^2 g^2 a}$. (Note that this constraint does not necessarily reduce the corruption to zero.)

[13] One of the authors, Geoffrey Brennan in a recent book, Brennan and Hamlin (2000), notes the problems with that assumption and notes the importance of "economising on virtue" where he describes his new position as "this marks a sharp departure from earlier writing... where the assumption of self-interested motivation is defended in the constitutional context."

3.2 Constitutional constraints when candidates are not identical

So far we have discussed identical candidates, and have shown that when the designers maximize expected voters' welfare rather than the welfare of the income group that they belong to, and have information, they can design reforms that are welfare increasing. But what if the candidates are not identical? Then it may be the case that none of the reforms above will work. Let us provide an example for the tax rate limits, as the case for minimum public good levels is the same. Whenever two candidates propose different tax rates in the equilibrium, the one who proposes the higher tax rate can be targeted by a constitutional limit on tax rates. This is effective if the corrupt candidate selects a higher tax rate. But equilibrium may involve the opposite. For computational simplicity, let us consider again the quasi–linear utility function

$$U = c_i + 2\sqrt{G}$$

with $a_1 = 0.36$, $a_2 = 0.30$, $b = 0.08$, $g = 25$, $L_1 = L_2 = 0.8$ and assume that there are no legal rents,[14] $w = E = 0$. In the equilibrium the first candidate proposes a tax rate of 36% and the second candidate proposes 32 percent taxes. The public good levels that they propose are $G_1 = (0.36)^2$, and $G_2 = (0.30)^2$. Only the second candidate steals. First let us note that any tax rate constraint higher than 32 percent, that is not binding for the candidate who steals in the equilibrium, makes voters (and honest candidate) worse off. It will induce candidate 1 to propose a platform that provides less utility to voters which increases Candidate 2's incentives to steal even further. A tax rate constraint that is less than 32 percent does not work either. For, when the tax rate constraint is equal to 32 percent, Candidate 2 is stealing more than what he stole when there was no constraint. Candidate 2 will reduce the amount he steals back to 2 percent, what he

[14] To assume that the legal rents are small would do it as well, here we follow Brennan and Buchanan (1980) by assuming no legal rents.

stole without the constraint, when the tax rate constraint is about 15 percent,[15] $\overline{\tau} = 0.15$. Since Candidate 1, who, in the first-best, should produce public good with 36 percent of total income, is forced to use only 15 percent of total income, the welfare loss due to that is much more than the welfare loss due to Candidate 2's 2 percent theft. Intuitively, candidate 2 steals because of his popularity advantage. The other candidate is more able and thus attempts to deliver higher public good, financed by higher taxes. Imposing tax rate constraints that bind for the honest candidate, makes popularity advantage even more important, allowing the corrupt candidate to steal even more.

Accordingly when candidates are not identical, tax rate constraints works for sure only when the candidate who proposes larger tax rates is corrupt. Similarly the minimum public good level requirements will work for sure when the candidate who proposes a smaller government is corrupt. But, even when both candidates are corrupt, but not identical, and the reform is welfare increasing, each of the constraints will give a relative advantage to one of the candidates.

4 Legal Incentives

When S_j stands for stealing, as it does in most parts of this paper, one anticipates the possibility of legal punishment. Let us assume that a corrupt candidate believes that with a small probability, p, he will get caught and even punished.[16] When the leader is caught in corruption, he will be deprived of his position and hence will lose the legal rents, both w and E. Let us further assume that there is a legal penalty as well. Although the details of the penalty depend on the laws of

[15] The solution to $-\frac{(0.02)(25)\sqrt{0.3}}{\sqrt{\tau - 0.02}} + 0.5 + 25[0.08 + 2(\sqrt{0.3(\tau - 0.02)} - \sqrt{(0.36)\tau})] = 0$ is $\tau = 0.15203$.

[16] Note that we assume that the probability is independent of the amount the leader steals. It is possible to imagine situations where stealing a great deal will increase (because of more attention) or decrease (because the politician becomes very strong and can threaten or bribe) the probability of punishment. We think that, one can find a functional form where $p = p(S_j)$ is an increasing/decreasing function, without changing our results qualitatively.

the country, in general it involves some monetary penalty and imprisonment.[17] The legal penalty for corruption, we assume, is linear in the amount at rate stolen. There is also a fixed component of the penalty with monetary equivalent of C. Thus, the expected rents that candidate j receives when he is the leader is

$$R_j^p(S_j) = w + E + \mathbf{1}_{\{S_j>0\}}[L_j S_j - p(vS_j + C + w + E)]. \tag{9}$$

It is clear that with a sufficiently strong legal enforcement, the problem of corruption can be eradicated. For example whenever $pv > 1$, the expected gain from corruption is definitely negative since in that case, $L_j - pv < 0$ for any L_j. Thus when the legal incentives are high enough, no one will steal no matter what the electoral incentives are.

We assume that such strong legal incentives are not feasible due to administrative and legal constraints.[18]

4.1 Equilibrium Under Law Enforcement

Now the analysis of equilibria is more complicated owing to a discontinuity in the objective function at $S_j = 0$. Theorem 1 no longer applies since it made use of the continuity of reaction functions. In the appendix, we show, however, that the reaction function under law enforcement, $s_j^p(S_k)$, can have at most one point of discontinuity. Accordingly, the reaction function has the form

[17] For instance, in the U.S., a public official who has accepted a bribe shall be "fined not more than three times the monetary equivalent of the thing of value or imprisoned for not more than fifteen years or both." (18 U.S.C. § 201, quoted in Rose-Ackermann (1999))

[18] Increasing p is not easy, since auditing (or prosecuting) the leader is different than, say, a tax collector. Since auditing even tax collectors is not an easy task, we assume that for the leader there is quite inadequate auditing, i.e., p is not zero, but is small. Given the weak auditing, what can be done? One solution, known as Becker conundrum, is to have a low probability of detection, but a very high punishment when the offender is caught. It makes law enforcement effective, despite the low probability of detection. That quick fix we think is not feasible either. In many countries, the legal system itself is not very accurate and is subject to influence by the executive branch. To allow one politician to be severely punished may deter not only corruption but also opposition. So we assume that the system has a weak auditing mechanism that is very expensive to fix, and that easy solutions such as very high punishments are not feasible.

$$s_j^p(S_k) = \begin{cases} 0 \text{ if } S_k < \widetilde{S}_k, \\ S_j^p(S_k) \text{ otherwise,} \end{cases}$$

where $S_j^p(S_k)$ is the continuous and upward sloping reaction function that one would calculate had the value of the objective function be equal to its limit from the right and thus continuos at $S_j = 0$,[19] and $\widetilde{S}_k$ is defined as the $\sup\{S_k \in [0,1] : \rho_j(S_j^p(S_k), S_k)R_j^p(S_j^p(S_k)) \leq \rho_j(0, S_k)R_j^p(0)\}$.

The point of discontinuity, $\widetilde{S}_k$, increases in the parameters of law enforcement, $p, v.C$.[20] Owing to this discontinuity there can be multiple (two) equilibria. In one of these equilibria, both candidates stay clean, and in the other one at least one of them steals.[21] The conditions for the existence of multiple equilibria for a general utility function are quite messy.[22] The question that we are interested in is when there is an equilibrium with at least one candidate stealing despite the legal incentives how effective are the higher wages in deterring political corruption?

Unlike constitutional constraints, the legal incentives are important here in evaluating the effect of higher wages and the effect of higher penalties on corruption and social welfare. In the following sections we explain why this is so.

4.2 Wage reform

As Persson and Tabellini (2000) observed higher ego rents imply lower political corruption.[23] Although politicians who get higher ego rents from being leaders are good for the voters, it is not clear how to find such people and replace the current (and corrupt) political elite with them. After Becker and Stigler (1974), efficiency wages are proposed by many authors in the literature as a

[19] For more on that see appendix.

[20] Simply applying implicit function theorem gives this result.

[21] The second equilibria Pareto dominates the first one, from *each* candidate's point of view. For more on this, see section 4.4.

[22] The necessary and sufficient condition for existence of an equilibrium where both candidates steal is $\widetilde{S}_k < S_k^p(\widetilde{S}_j)$ for at least one $j \in \{1, 2\}$ and $k \in \{1, 2\} \backslash \{j\}$.

[23] See Theorem 4.

solution to bureaucratic corruption.

Similar to ego rents, higher wages also makes winning the election more attractive, and induce the agents to comply more with voter will. The advantage of increasing wages over increasing ego rents is that it is easier to increase the monetary compensation than rents based on psychological factors. On the other hand, wage increases unlike increases in ego rents, should be financed from the public budget. Since, a clean government may have a high cost in terms of high wages paid to the political agents, one should calculate not only the effect of wages on corruption, but also the net effect, including the effect of wages on taxes and on public good levels. The total effect of an infinitesimal increase in wage, w, on (expected) voter welfare, $\mathbf{E}[\mathbf{W}]$, is[24]

$$\frac{d\mathbf{E}[\mathbf{W}]}{dw} = -\sum_{j\in\{1,2\}} \rho_j a_j \frac{d\mathbf{E}[U_i(.)]}{dG}(1+\frac{dS_j(w)}{dw}). \tag{10}$$

That implies, if we increase the wage candidate j receives, this will increase voter welfare only when the benefit of high wages (a decrease in S_j and hence an increase in G_j) is larger than the cost of high wages (a decrease in public good due to higher wages), i.e., only when $\frac{dS_j(w)}{dw} < -1$. Also note that the net benefit from one candidate affects voters' welfare proportional to the likelihood of that candidate winning the election. One implication of (10) is that whenever both candidates are honest, increasing wages is always bad for voter welfare, since it does not improve the quality of service, but instead, increases the cost of it.[25] So when one of the candidates is honest, increasing wages is not as effective as when both are stealing. Even when $S_j > 0$ for both candidates, the wage increase is good for voter welfare only when a dollar increase in wages reduces stealing more than a dollar. The next proposition characterizes exactly when that happens.

Proposition 4 *(i) When both candidates are identical, a small increase in wages increases voter*

[24] See Appendix for the derivation.

[25] Here we disregard the possibility that higher wages will attract higher ability candidates to politics, see Morelli and Caselli (2001) for a model of endogenously determined candidate characteristics.

welfare if and only if

$$L - pv < 1 - p.$$

(ii) If the candidates are not identical, yet both steal in the equilibrium, then for a small increase in wages to be welfare-increasing, a necessary condition is $\min\{L_1, L_2\} - pv < 1 - p$*, while a sufficient condition is* $\max\{L_1, L_2\} - pv < 1 - p$.

Proof. See the Appendix. ■

The wage increases work in two channels. The "direct" effect is that higher wages increase the rents from the office and hence the weight the candidate puts on voter welfare goes up, inducing lower corruption. The "strategic" effect, on the other hand, works on the last part of (8): a rival candidate also reduces his corruption, ρ_j is now lower, which further reduces the incentives to steal. Obviously the strategic effect occurs only when the rival candidate is also corrupt. An honest candidate cannot lower his level of corruption. Hence, the prize (higher wages) are most efficient inducing higher compliance with voter will when both candidates are identical and corrupt, i.e., $a_1 = a_2$ and $b = 0$.[26]

Similar to the optimal constitutional constraints, we have redistributive and informational issues with the salary reform: Suppose, again, that the most preferred tax rate decreases with income, and there are several reforms providing the same level of aggregate voters' welfare. When salary reform is chosen, taxes go up, but also the public good level. Everyone pays the cost (higher taxes), but the rich pay proportionally higher fraction. While the benefit (higher public good level), is also distributed equally. So for the same effect on (aggregate) voter welfare, high income voters would prefer the tax limits and low income voters would prefer either the wage increases or minimum public good requirements. Similarly, the optimal wage *for candidate j,* a wage that would maximize the expected welfare of the voters, can be calculated, but one needs to know about the

[26] In Appendix, we calculate the effect of wages when one of the candidates is honest.

honesty, efficiency and popularity of the candidate and his rival.

4.3 Small Changes in Penalties

There is always pressure on politicians from the public and nowadays from multinational organizations for harsher penalties on corruption. Although we have not specified the cost of higher legal incentives,[27] we feel that we may still speculate on the question that if in reaction to these pressures some *small* steps are taken, how would the outcome be changed?[28] The following proposition considers the effects of a small increase in either constant or variable components of corruption penalties.

Proposition 5 *A small increase in*

(i) constant penalty, C, leads to an increase in political corruption,

(ii) variable penalty, v, reduces corruption only when the expected constant penalty is less than the expected legal rents for a corrupt candidate, $pC < (1-p)(w+E)$.

Proof. By applying the implicit function theorem on (8)=0. ■

The intuition for (i) is that an increase in C actually reduces the expected rents from office and hence reduces the weight politician puts on voter welfare. Then, the marginal utility of stealing is higher for candidate j, so S_j is higher in the equilibrium. We have the same effect for the variable penalty as well, i.e., lower rents from the office as a result of higher penalties. But for the latter, there is another effect that works in the opposite direction, the higher the v, the lower is $L_j - pv$, i.e., the expected penalty per dollar stolen increases. As the previous ones, that result too depends on the change in the relative weights discussed in (8). If the decrease in the weight on voter welfare

[27] That we beleive requires explicitly modelling the relationship between law enforcers and the political leaders, a relationship that differs from the typical principal agent interaction. Here each side is principal and an agent in different instances.

[28] We have already seen that if the incentives are "strong enough" there will be no corruption.

due to the first effect is lower than the decrease in expected monetary benefit of a dollar stolen, then the second effect dominates and the equilibrium level of S_j will be lower.

The constant penalty is good only if it is high enough to completely deter corruption. Note that the condition for the effectiveness of a variable penalty will be more difficult to hold when the constant penalty is higher. Thus, in our model, the constant penalty can be justified only when it is sufficiently high to completely deter the political corruption.

4.4 Political Support for Anti-corruption Reform

We have seen that under some conditions some reforms, such as a sufficiently large improvement in legal incentives, will stop corruption. But such a reform needs to be proposed and implemented. An interesting question, then, is whether politicians will support the reform. A utility-maximizing politician should compare the benefits and costs of the reform for himself. Adding the reform to policy platform would increase his vote shares in current elections, yet curbing corruption might reduce his current and future payoffs. Since the problem is a dynamic one and our model is static, we discuss this question only informally here.[29]

Successful anti-corruption reforms, will be welcomed by the electorate. Yet, we have corruption to begin with exactly because there is an agency problem: policies that the electorate appreciates are not necessarily being implemented. If all candidates agree not to propose the reform, it will never be implemented and the corruption among the political leaders will continue.[30] When both candidates are corrupt it is not difficult to see that if the illegal rents from corrupt status quo are significantly high, then each of the (corrupt) candidates would rationally choose not to propose the reform.

One may be inclined to think that this corruption trap is possible only when all the politicians are

[29] We discuss this issue in a simpler setting with three candidates in Evrenk(2004).

[30] Of course, the reform can be proposed and be implemented by people other than politicians, as was the case in Italy with clean hands. But, eventually without the support of political leaders such reforms may not be long lasting.

corrupt. Since an honest politician receives no benefit from the corrupt status quo, he will incur no cost by supporting the reform. This reasoning is, however, not always correct. Consider an honest leader, Candidate 1, who is going to compete with a corrupt rival in the next election. An anti-corruption reform that will prevent all future corruption will affect the policy platform of Candidate 2 in future elections. It will induce Candidate 2 to offer a more voter friendly platform. This will reduce the honest candidate's vote share. So, the honest candidate may also not propose the reform. The intuition for this is that political competition is a zero sum game without corruption, but this is not true with corruption. The existence of corruption benefits both candidates, even when one of the candidates is completely honest. When one candidate is corrupt, he is better off, since he can get the illegal rents. The (honest) competitor is better off because by stealing the candidate makes his policy platform less attractive and hence the policy platform of his rival becomes more attractive. When the choice to be corrupt is no longer available, the corrupt candidate is going to lose his rents, but the honest one will lose some of his voters.

5 Other Approaches to Agency Problem in Politics.

Adsera, Boix and Payne (2003) extend the incumbency model by Persson and Tabellini (2000). They examine the incentives of incumbents to steal, given that voters have incomplete information about the state of the world and support the incumbent whenever he achieves a minimal performance standard. In their model, the minimum performance standard is the expected utility from the challenger and is exogenous. As can be seen in section 4.2, the strategic effects, the change in the challenger's performance as a result of, say a change in wages, is absent when the performance of challenger is fixed. Caselli and Morelli (2004) studied what determines the honesty and quality of elected politicians. Unlike us, they allow the quality to be determined endogenously. But in their model corrupt politicians do extract as much rents as they possibly can, i.e., there is no concern for reelection. The difference is mainly due to the fact that we study competition among finitely many, two, politicians whereas they study a continuum of politicians. In their model the large number of players reduces the strategic incentives in rent extraction to zero. So, politicians either steal everything or they do not steal at all. Our analysis differs from both of these studies by modeling the strategic interaction between candidates.

Our model is a static one, politicians compete only once. That framework, in general would give rise to commitment problems, i.e., when the campaign promises are non-enforceable, non-verifiable, or non-observable, the agent has incentives to deviate from the promised behavior. In our analysis, we assume away that problem. The corrupt politicians in our model are honest thieves: they will deliver what they promised, although if political competition happens once a corrupt politician who won the election has incentives to steal all public revenue, not a fraction of it. What we want to show is that even when there is no agency problem in terms of commitment, corruption still occurs. Obviously if a candidate steals even when the campaign promises are enforceable, he will do so when they are not. It is possible to get rid of problem of commitment by assuming future elections and candidates with high discount rates as described in detail by Persson and Tabellini (2000, Ch. 4). Then the fear of losing the future gains induces at least some commitment by the candidates.

6 Conclusion and Future Research

In this paper, we discussed possible reasons for the persistence of corruption in democracies. We analyzed some commonly proposed reforms and show when, how and why they may (not) be useful. We pointed out serious informational problems and the conflict of interest among voters when one wants to regulate the political competition. We also argued that politicians themselves, honest and corrupt, may oppose anti-corruption reforms. For the analysis, we use a static probabilistic voting model with heterogenous candidates. We are planning to extend our analysis in following directions: (i) campaign financing, (ii) candidates with ideological motivations, and (iii) Principal-Agent analysis when agent has some authority over the principle.

In our model, the candidates steal for their own consumption which reduces their vote shares. We also observe that when campaign financing matters, candidates steal (or have alliances with businesspeople who will steal when candidates win the elections) to be able to raise money for campaign financing. To look at the corruption as the source of campaign financing, one would require a different model with voters who have imperfect information.

A candidate can have strong preferences on policy on the one hand and use his opportunities to steal on the other. The interaction of a candidate's policy preferences (on the tax rate and public good) and the amount he steals, as well as which part of the policy platform he steals from, could shed some light on the relationship between economic development and corruption.

The design and implementation of legal incentives for politicians are not simple applications of Principal–Agent theory. The Agent (candidate) has powers on the word of the contract as well as its enforcement that is unimaginable in standard Principal–Agent models. We believe that the analysis of the optimal contract as well as that of optimal auditing structure (in terms of institutions) in that framework is worth attention.

7 Appendix

Lemma 2 *The probability that $j \in \{1, 2\}$ wins the elections when he competes with k is* $\rho_j = \frac{1}{2} + g[\mathbf{E}[U(c_i^j, G_j) - U(c_i^k, G_k)] + \mathbf{P}_j]$, *where* $\mathbf{P}_j = 2(j - \frac{3}{2})b$.

Proof. Without knowing the personal preferences of each voter, a candidate can not know whether a specific voter is going to vote for him or not. What he can know is that voter i will vote for the candidate 1 iff $U_i^1 > U_i^2$ which is equivalent to say,

$b_{i2} < U(c_i^1, G_1) - U(c_i^2, G_2) - b - b_2.$

Then the probability of voter i voting for candidate 1 is

$\frac{1}{2} + f[U(c_i^1, G_1) - U(c_i^2, G_2) - b - b_2]$. If we sum this over Y_i the expected vote share of the candidate 1 is equal to

$\phi = \frac{1}{2} + f[\mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b - b_2].$

Candidate 1 is going to win the elections and become the leader whenever $\phi > \frac{1}{2}$ or equivalently whenever

$b_2 < \mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b.$

Using the distribution of b_2, we find that the probability of candidate 1 winning the elections as a function of the policy platforms and the popularity of candidates is

$\frac{1}{2} + g[\mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b].$ ■

Lemma 3 *The voters' expected (utilitarian) welfare,* $\mathbf{E}[\mathbf{W}]$, *as a function of policy platforms and popularity of each candidate can be written as* $\mathbf{E}[U_i((1-\tau_2)Y_i, G_2(\tau_2, S_2))] + b + \frac{1}{2g}(\rho_1)^2$.

Proof. The voter i's expected welfare is equal to

$\rho\mathbf{E}[U_i^1 \mid \text{candidate 1 won the election}] + (1-\rho)\mathbf{E}[U_i^2 \mid \text{candidate 2 won the election}]$. The expected value of b_{i2} conditional on candidate 2 winning the election is equal to its unconditional expected value, which is zero. So the voter i's expected welfare is equal to

$\rho_1 U_i((1-\tau_1)Y_i, G_1) + (1-\rho)U_i((1-\tau_2)Y_i, G_2) + (1-\rho_1)b$

$+(1-\rho)\mathbf{E}_{b_2}[b_2 \mid b_2 < \mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b].$

Note that $(1-\rho)E_{b_2}[b_2 \mid b_2 < \mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b]$

is equal to $(1-\rho)\frac{\int_{U(c_i^1,G_1)-U(c_i^2,G_2)]-b}^{\frac{1}{2g}} xgdx}{1-\rho}$, which is equal to

$\frac{1}{2g}[\frac{1}{4} - (\rho_1 - \frac{1}{2})^2]$. Thus we can write the welfare of voter i as,

$$\rho_1 U((1-\tau_1)Y_i, G_1) + (1-\rho_1)U((1-\tau_2)Y_i, G_2) + (1-\rho_1)b \ + \frac{1}{2g}[\frac{1}{4} - (\rho_1 - \frac{1}{2})^2]. \tag{11}$$

Summing (11) over i and using (4), we have the desired result. ■

F.O.C. with respect to Tax Rate and the Effect of Corruption on Taxes and Public Good Levels.

To solve (5), candidate j should choose a tax rate such that the marginal utility of tax rate for candidate j,

$$gR_j \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - W - S_j))]}{\partial \tau_j} \tag{12}$$

is zero at τ_j^*. Since gR_j is always positive, the first order condition w.r.to tax rate holds only when $\frac{\partial \mathbf{E}[U_i(\tau_j, S_j)]}{\partial \tau_j} = 0$. Thus, when maximizing his expected payoffs, candidate j chooses a tax rate that maximizes $\mathbf{E}[U_i(\tau_j, S_j)]$, the average welfare of voters, for *given* corruption level, S_j. Then, when the candidate j does not steal, the policy platform he chooses is optimal, $\tau_j^* = \tau_j^0$.

Note that the f.o.c. with respect to tax rate does not *directly* depend on the policy platform of candidate k. The effect of other candidate's platform will be seen, if at all, through S_j. Using the implicit function theorem, we can calculate the effect of a small change in S_j on tax rate: $\frac{\partial \tau_j^*(S_j)}{\partial S_j} = \frac{E[(a_j)^2 H''(G)]}{E[Y_i^2 I''(c_i) + (a_j)^2 H''(G)]} \in [0,1]$. Figure 2 shows how S_j determines τ_j^* for three different utility functions, U.

The quasi–linear utility functions determine the borders of the derivative: When $I(.)$ is linear, $I''(.) = 0$, we have $\frac{\partial \tau_j^*(S_j)}{\partial S_j} = 1$. Then the effect of political corruption is socially optimal public good levels, G_j^0, but higher than optimal taxes. When $H(.)$ is linear, $H''(.) = 0$, we have $\frac{\partial \tau_j^*(S_j)}{\partial S_j} = 0$. In such case the tax rates are always optimal, candidate steals only from the public good. When both $I(.)$ and $H(.)$ are strictly concave the derivative is between 0 and 1, and thus, the effect of

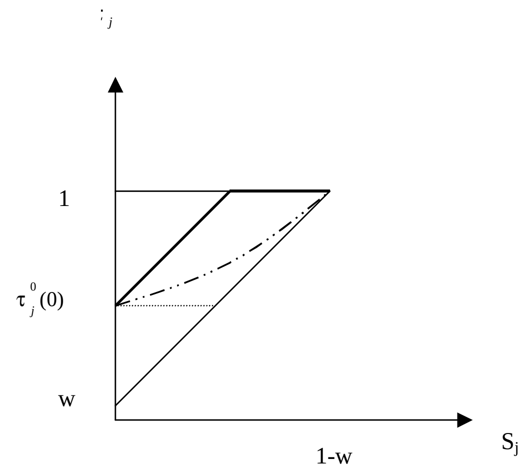

Figure 2: The effect of stealing on the tax rate

corruption is both lower than optimal public good levels *and* higher than optimal taxes. The kinks in the figure that we encounter in two quasi–linear cases are due to the finite marginal utility at zero consumption. In such case, the harm done to voters by stealing the last penny in the public budget or taking the last penny of the taxpayer is not different then stealing a penny from a large budget. Thus, a candidate may find it good policy to supply optimal public good yet impose 100 percent taxes. We can rule out those "extreme" platforms, i.e., platforms that when implemented voters have zero (public or private good) consumption, by assuming that even in the quasi–linear case, the utility becomes strictly concave and the marginal utility goes to infinity around an epsilon neighborhood of zero consumption.[31] Hence, the strategy space relevant to our analysis, $(\tau_j^*(S_j), S_j)$ is a curve in Σ_j, and its slope is between zero and one.

Lemma 4 *Over* $\left(\tau_j^*(S_j), S_j\right)$*, the Marginal utility of corruption for candidate* j*, (8), is continuous and strictly decreasing, -hence the objective function is strictly quasi–concave-, in* S_j*, and continuous*

[31] Say, when $U = I(c) + G$, we replace G with
$$G = \begin{cases} G + \sqrt{\varepsilon} \text{ for } G > \sqrt{\varepsilon} \\ 2\sqrt{G} \text{ for } G \leq \sqrt{\varepsilon}. \end{cases}$$
For ε small enough the distance between $H(G)$ and $H^{\varepsilon}(G)$ is minuscule. Yet, as a result of this change, a candidate never offers zero public good, since offering a little bit of public good increases voters' utility signifcantly. Thus $(\tau_j^*(S_j), S_j)$ is a smooth curve.

and strictly increasing in S_k.

Proof. We need to consider the movements only on $\tau_j^*(S_j)$. Note that $\frac{\partial \mathbf{E}[U(c_i^j, G_j)]}{\partial G_j}$ is continuous in both τ_j and in S_j. Similarly $R_j(.)$ is also continuous in S_j. For the derivative as we increase S_j, $\rho_j \downarrow$ and $R_j \uparrow$. For $\frac{\partial \mathbf{E}[U(c_i^j, G_j)]}{\partial G_j}$ we have two effects but since $\frac{\partial \tau_j^*(S_j)}{\partial S_j} \leq 1$ the net effect is also not a decrease, hence $-a_j g R_j \frac{\partial \mathbf{E}[U(c_i^j, G_j)]}{\partial G_j} \downarrow$. The arguments for S_k is similar, only simpler. ■

Lemma 5 *The corruption levels of candidates are strategic complements, $\frac{\partial s_j^0(S_k)}{\partial S_k} \geqslant 0$, with inequality being strict when $s_j^0(S_k) > 0$.*

Proof. When $s_j^0(S_k) > 0$, we have (8) evaluated at $(s_j^0(S_k), S_k)$ is equal to zero. Then using implicit function theorem it is straightforward to calculate that

$$\frac{\partial s_j^0(S_k)}{\partial S_k} = \frac{-\frac{\partial^2 \rho_j R_j}{\partial S_j \partial S_k}}{\frac{\partial^2 \rho_j R_j}{(\partial S_j)^2}} = \frac{z_{jk}}{2z_{jj} + a_j R_j \frac{\partial^2 \mathbf{E}[U(c_i^j, G_j)]}{(\partial G_j)^2}(\frac{\partial \tau_j^*(S_j)}{\partial S_j} - 1)}. \tag{13}$$

where $z_{jk} = L_j a_k \frac{\partial \mathbf{E}[U(c_i^k, G_k)]}{\partial G}$ and $z_{jj} = L_j a_j \frac{\partial \mathbf{E}[U(c_i^j, G_j)]}{\partial G}$. By concavity of $H(.)$ we have $\frac{\partial^2 \mathbf{E}[U(c_i^j, G_j)]}{(\partial G)^2} \leq 0$ and as we have shown above $\frac{\partial \tau_j^*(S_j)}{\partial S_j} - 1 \leq 0$. Thus both the nominator and denominator is positive. When (8) is negative at $S_j = 0$ then by continuity an infinitesimal increase in S_k is not going to increase the optimal S_j. ■

Proposition 6 *Reaction functions $s_1^0(S_2)$ and $s_2^0(S_1)$, if continuous, do not intersect more than once in the interior, i.e., $S_1^* > 0, S_2^* > 0$ such that $s_j^0(s_k^0(S_j^*)) = S_j^*$ is unique, if it exists.*

Proof. Assume that we have more than one interior equilibria. Then as Figure 3 shows we should have

$\frac{\partial s_j(S_k^*)}{\partial S_k} . \frac{\partial s_k(S_j^*)}{\partial S_j} \geq 1$ in at least one of the equilibria.

Using (13), we have

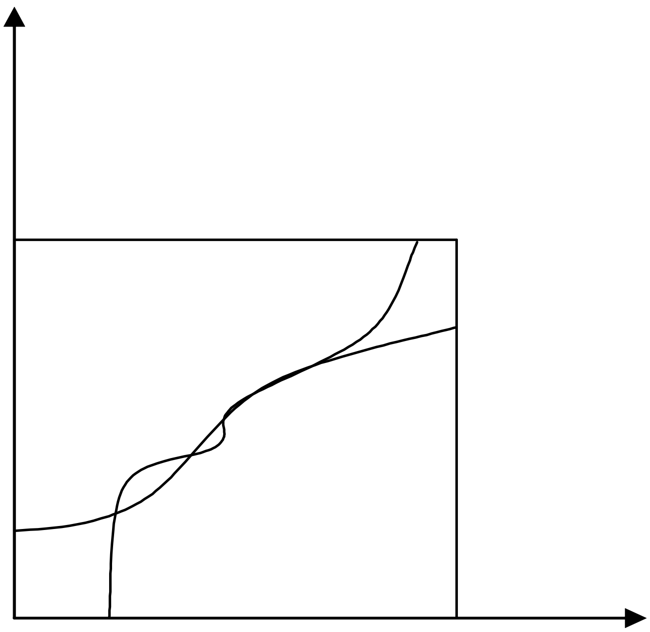

Figure 3: Reaction functions with multiple equilibria

$\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} \geq 1 \Leftrightarrow \frac{z_{jk}z_{kj}}{4z_{jj}z_{kk}+Z} \geq 1$ where $Z \geq 0$.

Note that $z_{jk}z_{kj} = z_{jj}z_{kk}$. Hence

$\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} \geq 1 \Leftrightarrow \frac{z_{jj}z_{kk}}{4z_{jj}z_{kk}+Z} \geq 1$ where $Z \geq 0$. Contradiction. ■

Corollary 1 *For later use note that the above result can be written as $S_j^* > 0$ and $S_k^* > 0$ implies that $D = \frac{\partial^2(\rho_j R_j)}{(\partial S_j^*)^2}\frac{\partial^2(\rho_k R_k)}{(\partial S_k^*)^2} - \frac{\partial^2(\rho_j R_j)}{\partial S_j^* \partial S_k^*}\frac{\partial^2(\rho_k R_k)}{\partial S_k^* \partial S_j^*} < 0$.*

Proof of Theorem 1. The objective functions of candidate j is quasi concave in S_j over $\left(\tau_j^*(S_j), S_j\right)$. By Berge's Maximum Theorem the best response correspondence, $s_j^0(S_k)$ is upper hemi-continuous. In Lemma 5, we find that the objective function is strictly quasi-concave in S_j implying that $s_j^0(S_k)$ is single valued, and hence is continuous. Using this result with Kakutani's Fixed Point Theorem we have the existence of equilibrium in pure strategies. For uniqueness let us first note that by No Extreme Platforms assumption we have $0 \leqslant s_j^0(S_k) < 1$. If there exists an interior equilibrium, then Proposition 6 shows that it is the only interior equilibrium. The continuity of reaction functions with strategic complementarity implies that even a corner equilibrium where one of the candidates steals zero can not exists. To see why, note that in such an equilibrium generically $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} = 0$ (and even when both reaction functions have nonzero slope it is still the case that $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} < 1$). But by Proposition 6 $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} < 1$ holds for the interior equilibrium as well. Since the reaction functions are continuous there should be another point of intersection between the corner equilibrium and the interior equilibrium, but at that interior equilibrium $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} \geq 1$. Contradiction. The same argument can be used to show that when there exist a corner equilibrium $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} < 1$, then we can not have any other corner equilibrium or interior equilibrium. ■

Equilibrium Outcome

The condition that candidate j steals even when his rival does not, $s_j^0(0) > 0$, is equivalent to

$$\mathbf{E}U_i^j(\tau_j^o, G_j^o) - \mathbf{E}U_i^k(\tau_k^o, G_k^o) > \frac{1}{L_j} a_j g(W+E)H'(G_j^o) - \mathbf{P}_j. \qquad (COND_j)$$

When it holds, $s_j^0(0) > 0$ is the point where the reaction function intersects the S_j axis. On the other hand when $s_j^0(0) = 0$, then we can define the point where the reaction function, $s_j^0(S_k)$ intersects S_k axis as $\underline{S}_k = \inf\{S_k : s_j^0(S_k) > 0\}$.

It is straight forward to calculate that $\underline{S}_j < s_j^0(0)$ iff

$$\frac{1}{L_k} a_k g[w+E]H'(G_k^o) < \frac{1}{L_j} a_j g[w+E+L_j s_j^0(0)]H'(G_j(s_j^0(0))). \qquad (INEQ_j)$$

Then, the unique Nash equilibrium of the game is:

(a) $S_1^* = S_2^* = 0$, iff $\mathbf{E}U_i^j(\tau_j^o, G_j^o) - \mathbf{E}U_i^k(\tau_k^o, G_k^o) < \frac{1}{L_j} a_j g(W+E)H'(G_j^o) - \mathbf{P}_j$, $\forall\, j \in \{1,2\}$

(b) a unique pair $S_1^* > 0$, $S_2^* > 0$ iff

-either $\mathbf{E}U_i^j(\tau_j^o, G_j^o) - \mathbf{E}U_i^k(\tau_k^o, G_k^o) > \frac{1}{L_j} a_j g(W+E)H'(G_j^o) - \mathbf{P}_j$, $\forall\, j \in \{1,2\}$

-or $\exists$only one $j \in \{1,2\}$ with $\mathbf{E}U_i^j(\tau_j^o, G_j^o) - \mathbf{E}U_i^k(\tau_k^o, G_k^o) > \frac{1}{L_j} a_j g(W+E)H'(G_j^o) - \mathbf{P}_j$ but

$\frac{1}{L_k}a_k g[w+E]H'(G_k^o) < \frac{1}{L_j}a_j g[w+E+L_j s_j^0(0)]H'(G_j(s_j^0(0))$.

(c) $S_j^* = s_j^0(0) > 0$ and $S_k^* = 0$ iff we have

$\frac{1}{L_k}a_k g[w+E]H'(G_k^o) > \frac{1}{L_j}a_j g[w+E+L_j s_j^0(0)]H'(G_j(s_j^0(0))$

and $\mathbf{E}U_i^j(\tau_j^o, G_j^o) - \mathbf{E}U_i^k(\tau_k^o, G_k^o) > \frac{1}{L_j}a_j g(W+E)H'(G_j^o) - \mathbf{P}_j$

with $\mathbf{E}U_i^k(\tau_k^o, G_k^o) - \mathbf{E}U_i^j(\tau_j^o, G_j^o) > \frac{1}{L_k}a_k g(W+E)H'(G_k^o) - \mathbf{P}_k$

Analysis of Equilibrium Under Law Enforcement

To start with let us define $r_j(S_j) = \begin{cases} R_j^p(S_j) \text{ for } S_j > 0, \\ \lim_{S_j \downarrow 0} R_j^p(S_j) \text{ at } S_j = 0. \end{cases}$

The function $\rho_j(S_j, S_k) r_j(S_j)$ is strictly quasi–concave and continuous in S_j and continuous in S_k. Then we derive a "fake" reaction function for candidate j, $S_j(S_k)$, from the optimization of $\rho_j(S_j, S_k) r_j(S_j)$ and take the relevant part of this reaction function, i.e.,

$s_j^p(S_k) = \begin{cases} S_j(S_k) \text{ if } \rho_j(S_j^p(S_k), S_k) R_j^p(S_j^p(S_k)) \leq \rho_j(0, S_k) R_j^p(0) \text{ and } S_j(S_k) > 0. \\ 0 \text{ otherwise.} \end{cases}$

Now, (8)=0 is necessary but not sufficient for $s_j^p(S_k) > 0$, although it is both necessary and sufficient for $S_j(S_k) > 0$.

The "fake" reaction function, $S_j(S_k)$, is similar to $s_j^0(S_k)$ in the sense that it comes from the maximization of a continuous and strictly quasi-concave objective function over a convex domain, hence it is single valued, increasing and continuous in S_k. Also Proposition 6 can be applied to the intersection of $S_j(S_k)$'s. It is this similarity that we use to extend the results from the analysis with no law enforcement. The following Lemma shows that when $S_j(S_k)$ becomes relevant at some level of candidate k's corruption, it is always relevant for any higher level of corruption. Hence, $s_j^p(S_k)$ can have at most one discontinuity and is strictly increasing in S_k as far as $s_j^p(S_k) > 0$.

Lemma 6 *If* $s_j^p(S_k') = S_j(S_k')$ *for some* S_k' *with* $S_j(S_k') > 0$ *then*

$\rho_j(S_j(S_k), S_k) R_j(S_j(S_k)) > \rho_j(0, S_k) R_j(0)$ *for any* $S_k > S_k'$.

Proof. Take any S_k' such that $\rho_j(S_j(S_k'), S_k') R_j^p(S_j(S_k')) \geq \rho_j(0, S_k') R_j^p(0)$. Let us note that both sides are continuously differentiable in S_k' and consider an infinitesimal increase in S_k'. The increase in l.h.s. is equal to $A = dS_k([\frac{\partial \mathbf{E}[U_i^j]}{\partial S_j}\frac{\partial S_j(S_k')}{\partial S_k} - \frac{\partial \mathbf{E}[U_i^k]}{\partial S_k}]gR_j(S_j(S_k')) + (L_j - pv)\frac{dS_j(S_k')}{dS_k}\rho_j(S_j(S_k'), S_k')) > 0$. The increase in r.h.s. is equal to $B = -dS_k\frac{\partial \mathbf{E}[U_i^k]}{\partial S_k}gR_j(0) > 0$. We need to show that $A - B > 0$. Note that $S_j(S_k') > 0$, implies that (8) holds with equality, which in turn implies that

$\frac{\partial \mathbf{E}[U_i^j]}{\partial S_k}gR_j(S_j(S_k')) + (L_j - pv)\rho_j(S_j(S_k'), S_k') = 0$.

Thus $A - B = dS_k[\frac{\partial \mathbf{E}[U_i^k]}{\partial S_k}gR_j(0) - \frac{\partial \mathbf{E}[U_i^k]}{\partial S_k}gR_j(S_j(S_k'), S_k')]$. We know that candidate j's probability of winning the election is higher if he does not steal, $\rho_j(S_j(S_k'), S_k') < \rho_j(0, S_k')$, so

$\rho_j(S_j(S_k'), S_k')R_j^p(S_j(S_k')) \geq \rho_j(0, S_k')R_j^p(0)$ implies that $R_j(S_j(S_k')) > R_j(0)$. Hence $A - B > 0$.

■

The difference is that now the game can have multiple equilibria, one equilibrium where no candidate steals and another one where at least one does. By an application of Proposition 6, the interior equilibrium is unique, (the intuition is that in the interior equilibrium it is the $S_j(S_k)'$s that intersect each other, and as Proposition 6 shows this can not happen twice in the interior).

Lemma 7 $\frac{d\mathbf{E}[\mathbf{W}]}{dw} = -\sum_{j\in\{1,2\}}\rho_j a_j \frac{d\mathbf{E}[U_i(.)]}{dG}(1 + \frac{dS_j(w)}{dw})$

Proof. The derivative of $\mathbf{E}[\mathbf{W}]$ with respect to w is

$\frac{d\mathbf{E}[U_i((1-\tau_2(w))Y_i, G_2(w))]}{dw} + \rho_1(\frac{d\mathbf{E}[U_i((1-\tau_1(w))Y_i, G_1(w))]}{dw} - \frac{d\mathbf{E}[U_i((1-\tau_2(w))Y_i, G_2(w))]}{dw})$

$= ((1-\rho_1)\frac{d\mathbf{E}[U_i((1-\tau_2(w))Y_i, G_2(w))]}{dw} + \rho_1\frac{d\mathbf{E}[U_i((1-\tau_1(w))Y_i, G_1(w))]}{dw})$.

Note that, $\frac{d\mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{dw}$ is equal to

$\frac{\partial \mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{\partial \tau_j}\frac{d\tau_j(w)}{dw} + \frac{\partial \mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{\partial S_j}\frac{dS_j(w)}{dw} + \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - w - S_j))]}{\partial w}$.

By the f.o.c for the tax rate the first term is zero, so we have

$\frac{d\mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{dw} = \frac{\partial \mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{\partial S_j}\frac{dS_j(w)}{dw} + \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - w - S_j))]}{\partial w}$.

As a last step note that,

$\frac{\partial \mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{\partial S_j} = \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - w - S_j))]}{\partial w} = -a_j\frac{\partial \mathbf{E}[U_i(c_i^j, G_j)]}{\partial G}$.

Thus, $\frac{d\mathbf{E}[U_i((1-\tau_j(w))Y_i,G_j(w))]}{dw} = -a_j \frac{\partial \mathbf{E}[U_i(c_i^j,G_j)]}{\partial G}(1 + \frac{dS_j(w)}{dw})$, and

$\frac{d\mathbf{E}[\mathbf{W}]}{dw} = -\sum_{j\in\{1,2\}} \rho_j a_j \frac{d\mathbf{E}[U_i(.)]}{dG}(1 + \frac{dS_j(w)}{dw})$. ■

Proof of Proposition 4. Taking the derivative of first order conditions and noting that the derivative of $\frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i,a_j(\tau_j-W-S_j))]}{\partial \tau_j}$ with respect S_j is equal to the derivative with respect to wage, w, we have the following equation,
$\begin{bmatrix} \frac{\partial^2(\rho_1 R_1)}{(\partial S_1)^2} & \frac{\partial^2(\rho_1 R_1)}{\partial S_1 \partial S_2} \\ \frac{\partial^2(\rho_2 R_2)}{\partial S_1 \partial S_2} & \frac{\partial^2(\rho_2 R_2)}{(\partial S_2)^2} \end{bmatrix} \begin{bmatrix} \frac{dS_1}{dw} \\ \frac{dS_2}{dw} \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2(\rho_1 R_1)}{\partial S_1 \partial w} \\ -\frac{\partial^2(\rho_2 R_2)}{\partial S_2 \partial w} \end{bmatrix}$. The solution is

$$\left[\frac{dS_j}{dw}\right] = \left[\frac{\frac{\partial^2(\rho_j R_j)}{\partial S_j \partial S_k}\frac{\partial^2(\rho_k R_k)}{\partial S_k \partial w} - \frac{\partial^2(\rho_k R_k)}{(\partial S_k)^2}\frac{\partial^2(\rho_j R_j)}{\partial S_j \partial w}}{\frac{\partial^2(\rho_1 R_1)}{(\partial S_1)^2}\frac{\partial^2(\rho_2 R_2)}{(\partial S_2)^2} - \frac{\partial^2(\rho_1 R_1)}{\partial S_1 \partial S_2}\frac{\partial^2(\rho_2 R_2)}{\partial S_1 \partial S_2}}\right].$$

By Corollary 1,

$$\frac{dS_j}{dw} < -1 \text{ iff } \frac{\partial^2(\rho_j R_j)}{\partial S_j \partial S_k}\frac{\partial^2(\rho_k R_k)}{\partial S_k \partial w} - \frac{\partial^2(\rho_k R_k)}{(\partial S_k)^2}\frac{\partial^2(\rho_j R_j)}{\partial S_j \partial w} < -D. \tag{14}$$

The effects of stealing and higher wages is the same on optimal taxes, $\frac{\partial \tau_j^*}{\partial S_j} = \frac{\partial \tau_j^*}{\partial w}$, thus $\frac{dG_j}{dS_j} = \frac{dG_j}{dw}$. Using this equality we can write

$$\frac{\partial^2(\rho_j R_j)}{\partial S_j \partial w} = \frac{\partial^2(\rho_k R_k)}{(\partial S_k)^2} + \frac{\partial^2(\rho_j R_j)}{\partial S_j \partial S_k} - A_j,$$

where $A_j = a_j g[1 - p - L_j + pv]\frac{\partial \mathbf{E}[U_i(\tau_j^*,G_j^*)]}{\partial G}$. Using this in (14) gives us

$$\frac{dS_j}{dw} < -1 \text{ iff } \frac{\partial^2(\rho_j R_j)}{\partial S_j \partial S_k}A_k - \frac{\partial^2(\rho_k R_k)}{(\partial S_k)^2}A_j > 0.$$

By Lemma 5 $\frac{\partial^2(\rho_j R_j)}{\partial S_j \partial S_k} > 0$ and $\frac{\partial^2(\rho_k R_k)}{(\partial S_k)^2} < 0$. From here it is easy to see that $\min\{A_k,\ A_j\} > 0$ is necessary and $\max\{A_k,\ A_j\} > 0$ is sufficient for higher wages to be welfare increasing when candidates steal. ■

Lemma 8 *If only candidate j steals in the equilibrium, we have*

$\frac{dS_j}{dw} < -1$ *iff* $L_j - pv < \frac{1-p}{[1+(a_k \frac{\partial \mathbf{E}[U_i^k(G_k^o)]}{\partial G} / a_j \frac{\partial \mathbf{E}[U_i^j(G_j^*)]}{\partial G})]}$.

Proof. When only candidate j steals $\frac{dS_j}{dw} = \frac{-\frac{\partial^2(\rho_j R_j)}{\partial S_j \partial w}}{\frac{\partial^2(\rho_j R_j)}{(\partial S_j)^2}}$, which implies

$\frac{dS_j}{dw} < -1$ iff $\frac{a_j}{a_k} \frac{1-p+L_j-pv}{L_j-pv} < \frac{\frac{\partial \mathbf{E}[U_i(\tau_k^o, G_k^o)]}{\partial G}}{\frac{\partial \mathbf{E}[U_i(\tau_j^*, G_j^*)]}{\partial G}}$. ■

References

[1] Adsera, A., Carles Boix, and Mark Payne (2001), "Are you being served? Political accountability and quality of government." Journal of Law, Economics and Organization. 19:2, 445-90.

[2] Becker, G. and George Stigler (1974), "Law Enforcement, Malfeasance and the Compensation of Enforcers" Journal of Legal Studies. 1, 1-19.

[3] Brennan, G. and James M. Buchanan (1980), "The power to tax: Foundations of a fiscal constitution," Cambridge University Press, Massachusetts.

[4] Brennan, G. and George Hamlin (2000), "Devices and Desires" Cambridge University Press.

[5] Caselli, F. and Massimo Morelli (2001), "Bad politicians" Jornal of Public Economics. 88:3-4, 759-82

[6] Ehrenhalt, A. (2002), "The paradox of corrupt yet effective leadership" The New York Times, 20 September 2002.

[7] Evrenk, H. (2004) "When and Why Honest Politicians Oppose Anti–Corruption Reforms: A Multi–Candidate Analysis?" Mimeograph.

[8] Kurer, O. (2001), "Why do voters support corrupt politicians?" in Jain, A. K. (editor) The Political Economy of Corruption, Routledge, New York.

[9] Persson, T. and Guido Tabellini (2000), Political Economics. The MIT Press. Cambridge Mass.

[10] Polo, M. (1998), "Electoral competition and political rents" Working Paper 144, IGIER, Bocconi University.

[11] Rose-Ackerman, S. (2001), "Political corruption and democratic structures" In Jain, A. K. (editor) The Political Economy of Corruption, Routledge, New York.

[12] Spence, M. (1975), "Monopoly, Quality, and Regulation" Bell Journal of Economics. 6, 417-429.