Mackerels in the Moonlight: A Model of Corrupt Politicians

Haldun Evrenk[1]
Economics, Boston University, 264 Bay State Road, Boston, MA 02215
email:haldun@bu.edu

January 26, 2004

[1]Gregory Besharov, Martino DeStefano, Hsueh-Ling Huynh, Bart Lipman, Michael Manove, Zvika Neeman, Rasim Ozcan, Bedri Kamil Onur Tas, and Jorgen Weibull provided helpful discussions and comments. I especially want to thank Dilip Mookherjee for all the advice and encouragement. Errors are mine.

Abstract

This paper examines causes of the persistence of corruption among elected politicians in democracies. 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 public good cost and the amount stolen by the elected politician. We solve for the tax rates chosen by the candidates and how much each candidate chooses to steal depending on his ability and popularity. We, then, analyze the effects of various commonly discussed reforms as potential ways of deterring political corruption. We identify conditions under which (i) imposing tax rate limits, (ii) increasing compensation of elected politicians, and (iii) raising legal penalties for corruption, will increase corruption and/or reduce the social welfare. Under certain conditions, the reforms that will reduce corruption will not be supported by either corrupt or 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 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

[1] Quoted in Ehrenhalt [2002].

in equilibrium is increasing in heterogeneity among candidates with respect to their popularity, 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.

Models of corruption based on competition with probabilistic voting were considered earlier by Brennan and Buchanan [1980], Polo [1998] and Persson and Tabellini [2000]. Our model extends and generalizes these models in a variety of directions. In comparison with Brennan and Buchanan, for instance, 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. The Brennan-Buchanan theory is analogous to a monopolist who selects minimum quality and charges the highest price that leaves the buyer indifferent between buying the good and not. In such case, imposing a price ceiling raises consumer welfare. Whether imposing a price ceiling in a duopoly will result in higher consumer welfare is, however, more complicated. In a duopoly, the quality provided by a firm is not necessarily at the minimum level. Forcing firms to lower their price may result in a proportional reduction in quality, which is not necessarily welfare-increasing. We find that when both firms (resp. candidates) are identical and maximize profits (resp. are corrupt), a price ceiling (tax rate constraint) slightly lower than the equilibrium is welfare-increasing if and only if the utility from quality (resp. public good) is strictly concave. In order to calculate the appropriate constraints, however, drafters of a constitution will require information that is privately held by the (current and future) candidates, such as how able and honest they are. And when the candidates are not identical, the constraints may have the opposite effect of raising corruption and lowering welfare. In equilibrium, a less popular candidate may differentiate himself by providing higher public good, financed by higher tax rates and less corruption. A tax rate constraint can affect the policy of this candidate, resulting in a higher competitive advantage for the more popular and corrupt candidate. This will encourage the latter to become more corrupt.

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 constitutional tax constraints, i.e., most of the burden of the former is borne primarily by the rich, the latter by the poor.

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 legal incentives are weaker, the political competition game has multiple (two) equilibria: either both candidates stay honest or both steal. 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 demonstrate 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, 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. Many reforms commonly suggested (such as constitutional tax constraints, and legal and salary reforms) may increase corruption. And even when there exists a welfare improving reform that is supported by electorate, it may not be proposed by any of the politicians 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 quasilinear 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 two reforms: higher wages and higher legal penalties. At the end of section 5 we compare the two reforms (constitutional constraints on tax rates and higher wages) from a distributional point. 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.

Voters.

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. In order to ensure interior outcomes we assume

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

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

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

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.

Candidates.

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 $E[b_{i2} \mid b, b_2] = 0$.

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 technology, that depends on his ability. The 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

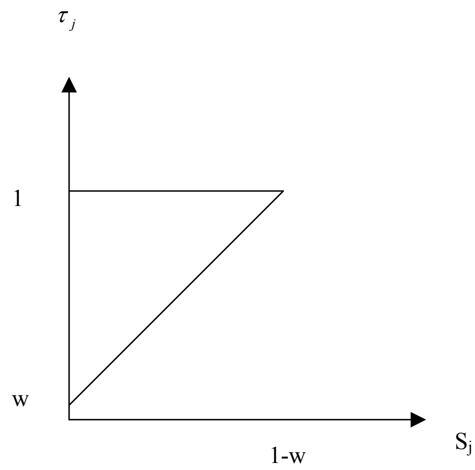

Figure 1:

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

$\Sigma_j = \{(\tau_j, S_j) : \tau_j \in [w, 1] \text{ and } S_j \in [0, \tau_j - w]\}$, as shown in Figure 1.

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.[2] 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

[2]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.

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

The probability that j wins the elections when he competes with k is[3]

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

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 normalize the outside option for candidates to zero. Then candidate j selects a policy platform[4] 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:

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

The voters' expected (utilitarian) welfare, $\mathbf{E}[\mathbf{W}]$, as a function of policy platforms and popularity of each candidate is[5]

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

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 $\mathbf{E}[U_i((1-\tau_j)Y_i, G_j(\tau_j, 0))]$, the average utility of the electorate.[6] The optimality of zero corruption/shirking is intuitive: Given the tax rate, less stealing means higher public goods delivered.

[3] See Appendix.

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

[5] See Appendix.

[6] Thus, the first-best tax rate is $\tau_j^0 = \arg\max_{\tau_j \in [w,1]} \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - w))]$ and the first-best public good level is $G_j^0 = a_j(\tau_j^0 - w)$.

3 Nash Equilibrium.

First Order Conditions.

Conditional on S_j, τ_k, S_k, candidate j selects τ_j to maximize ρ_j. This implies (given (4)) 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. The marginal utility of S for candidate j 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 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.[7] Let $s_j^0(S_k)$ denote the best response of candidate j to a rival stealing S_k. The corruption levels of candidates are strategic complements, i.e., $\frac{\partial s_j^0(S_k)}{\partial S_k} \geq 0$, The best response functions 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 steal), or (iii) no corruption (both candidates offer policies that maximize voters' welfare). Figure 2 describes four different subsets of parameters that give rise to these different outcomes. In graphs (a) and (c) both candidates steal. Only Candidate 1 steals in (b). In (d) none of them steals. The

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

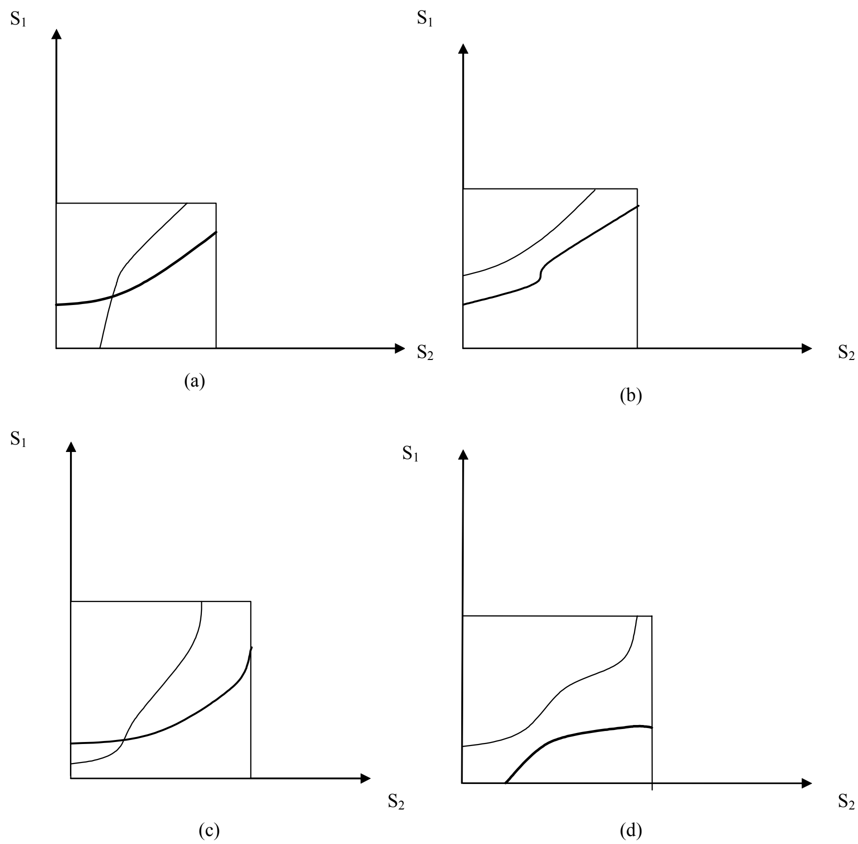

Figure 2:

thick curve is $s_1^0(S_2)$. 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 $\underline{S}_j = \inf\{S_j \mid s_k^0(S_j) > 0\}$.

A natural question to ask is which subset of parameters gives rise to which of the graphs in Figure 2. 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.

3.1 An Example: Quasilinear Utility.

Assume[8] 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 comparative advantage that the candidate j has and $K_j^0 = \frac{W+E}{L_j}$. The last term is equal to the illegal rents that are payoff equivalent to legal rents, i.e., the amount of corrupt rents when stolen that would yield the same income as legal rents.[9] To rule out policy platforms that involve zero consumption when $U = c + H(G)$, the overall uncertainty and the relative advantage of a candidate should not be too high and/or the legal rents should not be too low.[10]

The unique Nash Equilibrium of the political competition game when $U = c+2\theta\sqrt{G}$ is

(i) $S_j^* = \frac{1}{2g} + \frac{\theta^2(a_j-a_k)+2(\frac{1}{2}-j)b}{3} - \frac{W+E}{3}(\frac{2}{L_j} + \frac{1}{L_k})$ for all $j \in \{1,2\}$ iff

-either $\frac{1}{4g} + \theta^2(a_j - a_k) + 2(\frac{1}{2} - j)b - \frac{W+E}{L_j} > 0$ for all $j \in \{1,2\}$

-or $\frac{1}{4g} + \theta^2(a_{j'} - a_k) + 2(\frac{1}{2} - j')b + \frac{W+E}{L_{j'}} > 0$ for only $j' \in \{1,2\}$ 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) $S_{j'}^* = \frac{1}{4g} + \frac{\theta^2(a_j-a_k)+2(\frac{1}{2}-j)b}{2} - \frac{W+E}{2L_j} > 0$, $S_k = 0$ iff $\frac{1}{4g} + \theta^2(a_{j'} - a_k) + 2(\frac{1}{2} - j')b + \frac{W+E}{L_{j'}} > 0$ for only $j' \in \{1,2\}$ 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 $k \neq j'$.

(iii) $S_1^* = S_2^* = 0$ iff $\frac{1}{4g} + \theta^2(a_j - a_k) + 2(\frac{1}{2} - j)b - \frac{W+E}{L_j} \leq 0$ for all $j \in \{1,2\}$.

As both Polo[1998] and Persson and Tabellini [2000] note, the quasilinear utility

[8] 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 may involve 100% taxes when utility is quasilinear (see Appendix for a detailed discussion of why). In footnote 10, we provide conditions that rules out 100% taxes for quasilinear utility function.

[9] So when $S_j^* > K_j^0$, then the proportion of illegal rents in candidate j's income is larger than the proportion of legal rents.

[10] It is easy to calculate that a 100% tax rate is not part of candidate $j's$ best response iff

$\frac{1}{4gy} + \frac{\theta^2(3a_j-a_k)+\mathbf{P}_j}{2} + W - \frac{W+E}{2L_j} < \frac{1}{2}$.

We need a weaker condition for zero private corruption not to be an equilibrium, although under this condition, it may still be a best response to some policy platform. For positive private good consumption in equilibrium, the parameters should satisfy,

$\frac{3}{4gy} + \theta^2(3a_j - a_k) + 2\mathbf{P}_j - W + E(\frac{1}{L_j} + \frac{1}{2L_k}) < \frac{3}{2}$ for both $j \in \{1,2\}$.

function implies that the effects of higher corruption will be 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 special functional form. In Appendix, we show how the effect of corruption on tax rates and public good levels differ for different utility functions. For different utility functions, the slope of reaction function is not necessarily independent of parameters of the model either. However even when we consider different utility functions, the *direction* of comparative statics does not change. In the next section we present comparative statics again for general U.

3.2 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,* $2(\frac{3}{2} - j)b$,
- *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. ■

Note that a shift in the reaction function does not always imply a change in the equilibrium. For instance, for Graph (d) in Figure 2, a small change in any of the parameters has no effect on the outcome, i.e., the candidates who stay honest will not start to steal after the uncertainty increases a little bit. On the other hand if a candidate was stealing in the equilibrium, higher uncertainty will make him steal more. The compaartive statics are monotone: after a sufficiently large increase in uncertainty, a candidate who was honest (but not pathologically) will start stealing.

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 [1983], Polo [1998])*** *Suppose that candidates are identical,* $(a_1 = a_2,\ b = 0)$, *are not pathalogically 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. The marginal utility of corruption for candidate j is $L\rho_j$, which is strictly positive when $L > 0$. Thus we always have $s_j^0(0) > 0$ then, 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. 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 1 does not require a specific utility function. Also, as far as there are no legal rents, Theorem 1 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, $U(c_i G) = c_i + H(G)$ 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 win. But 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.[11] ■

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.

Persson and Tabellini [2000] discuss the agency problem in politics employing a probabilistic voting model and a quasilinear 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 $E > \frac{L}{2g}$.*

[11] 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 as far as marginal utility from public good is strictly positive. 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.

4 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 constitutional constraints on tax rates as discussed in chapter 10 of Brennan and Buchanan [1980].[12]

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 democracies, 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.

[12] 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.

Proposition 5 *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

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

$$gR_j \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - W - S_j))]}{\partial S_j} + (L - pv)\rho_j \leq 0,$$

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 $\lambda_j > 0$, the shadow value of constraint is positive and is equal to the expected marginal utility of electorate with respect to tax rate. But in a first-best this 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?

Let us consider a tax rate constraint that is marginally less than the equilibrium tax rates without the constraint. The effect of a tax rate constraint, $\overline{\tau}$, that is infinitesimally smaller than τ_j^* on voters utility from candidate j can be approximated as the sum of a direct effect and indirect effect (via the amount stolen)
$\frac{\partial \mathbf{E}[U_i^j]}{\partial \overline{\tau}} = \frac{\partial \mathbf{E}[U_i^j]}{\partial \tau_j^*}(d\tau_j) + \frac{\partial \mathbf{E}[U_i^j]}{\partial S_j}\frac{\partial S_j(\overline{\tau})}{\partial \overline{\tau}}$.

At an interior Nash equilibrium, the direct effect is zero, so only the indirect effect operates. Using the implicit function theorem, it is easy to calculate that when both candidates are identical, $\frac{\partial S_j(\overline{\tau})}{\partial \overline{\tau}} = \frac{gRa^2H''(G)}{gRa^2H''(G) - LaH'(G)}$,hence the indirect effect is equal to

$$-aH'(G^*)\frac{gRa^2H''(G)}{gRa^2H''(G) - LaH'(G)}. \tag{9}$$

When $H(.)$ is strictly concave (9) is larger than zero. This implies:

Proposition 6 *Whenever* $H(G)$ *is strictly concave in a neighborhood of* G^**, and both candidates are identical and corrupt, a constitutional constraint that enforces both candidates to offer a tax rate that is slightly lower than* τ^* *is corruption reducing and welfare-improving.*

The intuition is that tax rate constraints lower G, raise marginal utility of public good. This increases the voters' disutility from corruption. Hence the marginal utility of stealing for a candidate is lower. Contrast to the Laffer curve argument for tax limits in Brennan and Buchanan [1980]. The Leviathan taxes its subjects up to a point such that increasing tax rates does not increase tax revenues anymore. It follows from the assumption of monopoly power of the politician. Our argument incorporates effect of political competition.

So far we have discussed identical candidates. But what if they are not? 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. This is effective if the corrupt candidate selects a higher tax rate. But equilibrium may involve the opposite. Consider the quasilinear 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 that there are no legal rents[13],$w = E = 0$. 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. Any tax rate constraint higher than 32 percent 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 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 stole without the constraint, when the tax rate constraint is about 15 percent,[14] $\overline{\tau} = 0.15$. Since Candidate 1, who, in the first-best, should produce public good with 36 percent of total income, is forced to use 15 percent of total income, the welfare loss due to that is much more than the welfare loss due to Candidate 2's theft. Intuitively, candidate 2 steals because of his popularity advantage. The other candidate is more able and thus

[13] 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.

[14] 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 $\tau = 0.15203$.

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 are useful only when the candidate who proposes larger tax rates is corrupt.

In our model, we can calculate the cost and benefit of constraints and the optimal constraint, as well as the necessary information to set the optimal constraint. Below we calculate the optimal tax rate constraints for $U = c + 2\theta\sqrt{G}$ when both candidates are identical and corrupt.

Lemma 2 *When $U = c + 2\theta\sqrt{G}$ and candidates are identical (and corrupt), the tax rate constraint that maximizes voters' expected welfare is $\overline{\tau} = \tau^o + S^* - \frac{1}{8\theta^2 g^2 a}$.*

The above lemma demonstrates that when both candidates are identical and corrupt, drafters of constitution can set an optimal tax limit. It may not eliminate corruption totally. For example if $S^* = \frac{1}{2g} + \frac{W+E}{L} > \frac{1}{8\theta^2 g^2 a}$, then even under the best tax rate constraint the candidates keep stealing, and the tax rate is higher than the first best tax rate.

Consider the necessary information required to set the optimal tax rate constraint. Suppose that the writers of constitution know both U 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] (also common in works by scholars from Virginia school of public choice) –

"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, Hume (1985)" –

makes one 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.[15] This idea would be correct only

[15]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-

when such restrictions are costless. However 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. [16]

5 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.[17] 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 the country, in general it involves some monetary penalty and imprisonment.[18] 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

$$R_j^p = W + E + \mathbf{1}_{\{S_j>0\}}[L_j S_j - p(vS_j + C + W + E)]. \tag{10}$$

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

interested motivation is defended in the constitutional context.

[16]It is interesting to note the similarities between constitutional constraints projects by Virginia school of public Choice and the regulation of a market. The previous analysis could be done with politicians replaced with firms and drafters of constitution replaced with regulatory agencies. Yet, regulating a duopoly is less difficult, because it can be done through a "law" rather than a "constitution" and there is a larger consensus on the motives of the firms .

[17]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. One can find a functional form where $p = p(S_j)$ is an increasing/decreasing function, without changing our results qualitatively.

[18]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])

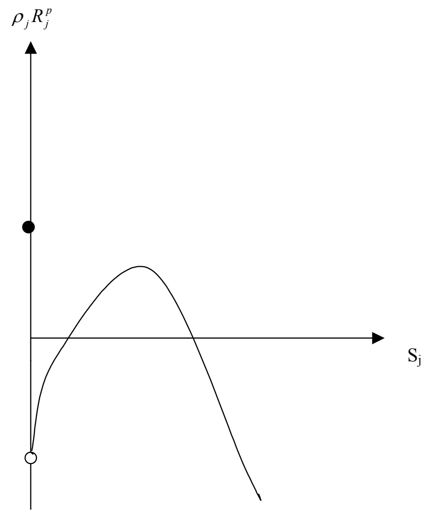

Figure 3:

are.

We assume that such strong legal incentives are not feasible due to administrative and legal constraints.[19]

5.1 Equilibrium Under Law Enforcement.

Now the analysis of equilibria is more complicated owing to a discontinuity in the objective function at $S_j = 0$, (see Figure 3). 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 looks like either Figure 4, or Figure 5.

[19] 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.

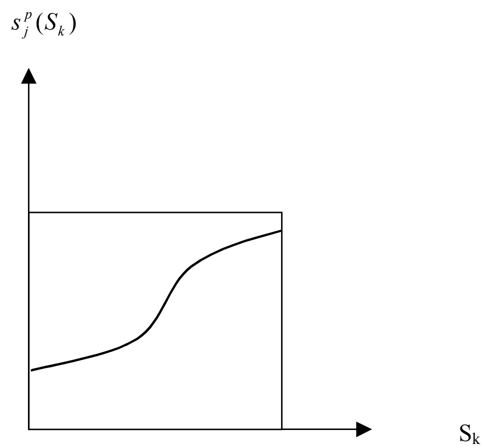

Figure 4:

Owing to this discontinuity there can be multiple (two) equilibria. The conditions for the existence of multiple equilibria for a general utility function are quite messy. Here we provide these conditions only for our quasilinear example. When $U = c + 2\theta\sqrt{G}$, the reaction function is

$$s_j^p(S_k) = \begin{cases} 0 \text{ if } S_k \leq \widetilde{S}_k \\ \frac{1}{4g} + \frac{S_k}{2} + \frac{A_j}{2} - \frac{K_j^p}{2} \text{ otherwise} \end{cases},$$

where $\widetilde{S}_k = (\frac{K_j^p(1+p)}{(1-p)} + \frac{2pC}{(1-p)(L_j-pv)}) + \sqrt{(\frac{K_j^p(1+p)}{(1-p)} + \frac{2pC}{(1-p)(L_j-pv)})^2 - 2(\frac{K_j^p}{2})^2} - \frac{1}{2g} - A_j$,

and $K_j^p = \frac{(W+E)(1-p)-pC}{(L_j-pv)}$.

The effect of law enforcement on the point of discontinuity, $\widetilde{S}_k$, is clear: the higher the law enforcement, K_j^p, the higher is $\widetilde{S}_k$. The effect of uncertainty and relative advantage is as before: the higher $\frac{1}{g}$ or $\mathbf{P}_j$ the incentives for candidate j to steal is higher, hence $\widetilde{S}_k$ is lower. When $\widetilde{S}_k$, as calculated above, is negative for both candidates, then the unique equilibrium always involves no corruption. When $\widetilde{S}_k > 1$ for both candidates, the unique equilibrium involves corruption by both candidates. The necessary and sufficient condition for multiple equilibria is $\widetilde{S}_k < S_k^p(\widetilde{S}_j)$ for at least one $j \in \{1, 2\}$ and $k \in \{1, 2\} \backslash \{j\}$. For the quasilinear example this condition is equivalent to $\widetilde{S}_k - \frac{\widetilde{S}_j}{2} < \frac{1}{4g} + A_j$. If this condition holds, the game has two equilibria (stay clean, stay clean) and (steal, steal), where the second one Pareto dominates the first one from the players' point of view.

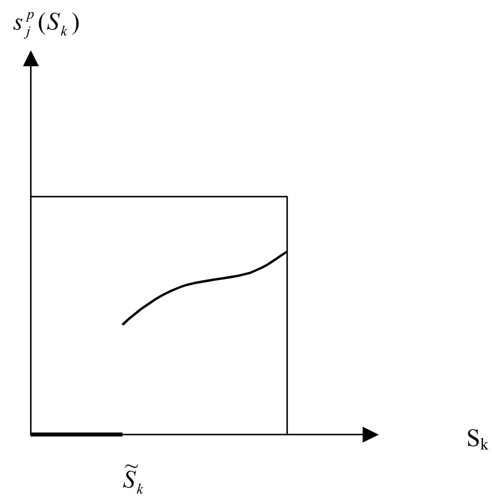

Figure 5:

The comparative statics with respect to parameters in Lemma 1 are similar, i.e., the existence of legal incentives do not change the *direction* of electoral incentives. In Section 5.3, we present comparative statics w.r.to penalties.

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.

5.2 Wage reform.

As Persson and Tabellini [2000] observed higher ego rents imply lower political corruption.[20] 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 solution to bureaucratic corruption. Wittman (1995) mentions contractual solutions among the ways to solve the agency problem in democracies. Here, we discuss the effect of an increase in wages on S_j and on voters' expected welfare.

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

[20] See Theorem 4.

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 on (expected) voter welfare is[21]

$$\frac{d\mathbf{E}[\mathbf{W}]}{dw} = \sum_{j\in\{1,2\}} \rho_j a_j \frac{d\mathbf{E}[U_i(.)]}{dG}(1 + \frac{dS_j(W)}{dw}). \tag{11}$$

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). The net benefit from one candidate affects voters' welfare proportional to the likelihood of that candidate winning the election. One implication of (11) 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.[22] 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 $\frac{dS_j}{dW} < -1$.

Proposition 7 *(i) When both candidates are identical, a small increase in wages increases voter 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

[21] See Appendix for the derivation.

[22] 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 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 $b = 0$.[23]

5.2.1 Comparing the reforms: ***Chicago versus Virginia.***

When we have an increase in social welfare, the distribution of benefits/costs of that increase is also of interest. Let us compare the two reforms, higher wages and constitutional constraints on tax rates, in terms of the burden they put on different income groups in society.

The two reforms will have different effects on the welfare of single individuals even when the effects on aggregate voter welfare is the same. The relative burden with tax rate constraints is on the poor, since they pay a smaller share of the taxes compared with the rich. The benefit of the reform, i.e., relatively higher per capita public good, is distributed equally among people.

In contrast, when wages increase, 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 constitutional constraints and low income voters would prefer the wage increases.

5.3 Small changes in penalties.

There is always pressure on politicians from the public and nowadays from multinational organizations for harsher penalties on corruption. If in reaction to these pressures some small steps are taken, how would the outcome be changed? The following proposition considers the effects of a small increase in either constant or variable components of corruption penalties.

Proposition 8 *A small increase in*

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

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

(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 usual, the result depends on the change in the relative weights discussed in (8). If the decrease in the weight on voter welfare 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.

5.4 Political support for anti-corruption reform

We have seen that a sufficiently large improvement in legal incentives will stop corruption. But such a reform needs to be proposed and implemented by politicians. 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.[24]

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

[24] Evrenk [2003b] offers an analysis of this issue in a three-candidate setting.

political leaders will continue.[25] 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 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 be 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.[26]

5.5 Other approaches to agency problem in politics.

Adsera et. al. [2001] 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 5.1, 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.

[25] 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 it is politicians who are going to control the legal system and the law enforcement, so without their support such reforms may not be long lasting.

[26] When there are more than 2 candidates, there are even additional factors that determine the location in the politcal spectrum and honest candidates' support for the reform. Evrenk [2003b] provides an analysis of this issue.

Caselli and Morelli [2001] 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, actually 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.

In his informal, but comprehensive paper, Kurer [2001] asks, "Why do voters support corrupt politicians?" He answers that it is either because the voters desire corruption or because there is no one else to support. The second case, he asserts, can be the result of barriers to entry or factionalism or both.

6 Conclusion.

This paper has discussed possible reasons for the persistence of corruption in democracies. We analyzed some commonly proposed reforms and show when, how and why they may be useful. We also argued that politicians themselves 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 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 Principle-Agent theory. The Agent (candidate) has powers on the word of the contract as well as its enforcement that is unimaginable in standard Principle-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.

References

[1] Adsera, A., Carles Boix, and Mark Payne [2001], "Are you being served? Political accountability and quality of government." Mimeograph,

[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." Mimeograph.

[6] Di Tella, R. "The new political economy of corruption?" Mimeograph.

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

[8] Evrenk, H. [2002a], "Are honest citizens to blame for corruption? An exercise in political economy of tax evasion" Mimeograph.

[9] Evrenk, H. [2002b], "Multi-party competition and political support for anti-corruption reforms" Mimeograph.

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

[11] Kunicova, J. and Susan Rose-Ackerman [2002], "Electoral rules as constraints on corruption." Mimeograph.

[12] Myerson R, B. [1993], "Effectiveness of Electoral Systems for Reducing Government Corruption: A Game Theoretic Analysis" Games and Economic Behavior. 5, 118-132.

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

[14] Persson, T., Guido Tabellini and Francesco Trebbi [2002], "Electoral rules and corruption." Mimeograph.

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

[16] Rose-Ackerman, S. [1999], Corruption and Government: Causes, Consequences and Reform. Cambridge University Press, New York.

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

[18] Sturzenegger, F. and M. Tommasi, (eds.) [1998], The Political Economy of Reform. The MIT Press. Cambridge Mass.

[19] Tirole, J. [1996], "A theory of collective reputations (with application to the persistence of corruption and the firm quality)" Review of Economic Studies, 63(1), pp. 1-22.

7 Appendix

Lemma 3 *When $U(c_i, G) = I(c_i) + H(G)$ with both $I()$ and $H()$ are strictly increasing, C^2 and concave functions from R_+ to R with at least one of them being strictly concave, the preferences of each voter is always single peaked in tax rates.*

Proof. Note that under the above conditions, we have $\frac{\partial^2 U}{(\partial\tau)^2} = Y_i I''() + (a_j)^2 H''(.) <$ 0 which implies that the utility function is strictly concave in tax rate for any given level of S. Then the local maximum is also the unique global maximum. ■

7.1 Vote shares.

Without knowing the personal preferences of each voter, a political 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

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

Since b_2 is a random variable, ϕ is a random variable too. Candidate 1 is going to win the elections and become the leader whenever $\phi > \frac{1}{2}$ or equivalently $b_2 <$

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

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

7.2 Voters' Welfare

The voter i's expected welfare is

$\rho E[U_i^1$ |candidate 1 won the election] + $(1-\rho)E[U_i^2$ |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 voters' welfare can be written as

$\rho_1 U_i((1-\tau_1)Y_i, G_1) + (1-\rho)U_i((1-\tau_2)Y_i, G_2) + (1-\rho_1)b$
$+(1-\rho)\mathbf{E}_{b_2}[b_2 \mid b_2 < \mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b]$.

Note that the second part,

$(1-\rho)E_{b_2}[b_2 \mid b_2 < \mathbf{E}[U(c_i^1, G_1) - U(c_i^2, G_2)] - b]$, is equal to $(1-\rho)\frac{\int_{U(c_i^1,G_1)-U(c_i^2,G_2)]-b}^{\frac{1}{2g}} xgdx}{1-\rho} = \frac{1}{2g}[\frac{1}{4} - (\rho_1 - \frac{1}{2})^2]$. Thus we can write the welfare of voter i as,

$$\rho_1 U((1-\tau_1)Y_i, G_1) + (1-\rho_1)U((1-\tau_2)Y_i, G_2) + (1-\rho_1)b + \frac{1}{2g}[\frac{1}{4} - (\rho_1 - \frac{1}{2})^2]. \quad (12)$$

Summing (12) over i and using (4), we have the desired result,

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

7.3 First Best Policy Platforms

Lemma 4 *The first best policy platform for candidate $j \in \{1, 2\}$ is a platform that maximizes the average utility of the electorate with zero corruption/shirking, i.e., (τ_j^o, G_j^o) is such that τ_j^o satisfies*

$\frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, G_j^0)]}{\partial \tau_j} \leq 0$ *(with equality when $\tau_j < 1$) and $G_j^o = a_j(\tau_j^o - W)$.*

Proof. Note that given the optimal S_j we are able to pin down the optimal tax rate and the public good levels. The derivative of voters' welfare with respect to S_j is $\frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, G_j^0)]}{\partial S_j}$. Since

$\frac{\partial U_i((1-\tau_j)Y_i, G_j^0)}{\partial S_j} = -a_j \frac{U_i((1-\tau_j)Y_i, G_j^0)}{\partial G_j}$ and $\frac{\partial U}{\partial G} > 0$, we have $\frac{\partial \mathbf{E}[\mathbf{W}]}{\partial S_j} < 0$, i.e., the voters' welfare is maximum when S_j is minimum $(= 0)$. The f.o.c. with respect to tax rate from the maximization of $\mathbf{E}[\mathbf{W}]$ implies $\rho_j \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, G_j^0)]}{\partial \tau_j} \leq 0$ (with equality when $\tau_j < 1$). ■

7.4 First order condition w.r.to Tax rate

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

$$gR_j \frac{\partial \mathbf{E}[U_i((1-\tau_j)Y_i, a_j(\tau_j - W - S_j))]}{\partial \tau_j} \tag{13}$$

is zero at τ_j^*.[27] 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/shirk, the policy platform he chooses is optimal, $\tau_j^* = \tau_j^0$.

7.5 Effect of Corruption on Tax rates and on Public Good levels.

Note that the f.o.c w.r.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.

When the candidate steals, i.e., $S_j^* > 0$, the tax rate he chooses is not necessarily τ_j^0. Using the implicit function theorem, we can calculate the effect of a small change in S_j on tax rate: $\frac{\partial \tau_j^*(S_j)}{\partial S_j} = \frac{E[(a_j)^2 U_{22}]}{E[Y_i^2 U_{11} + (a_j)^2 U_{22}]} \in [0, 1]$. Figure 6 shows how S_j determines τ_j^* for three different utility functions, U.

The quasilinear utility functions determine the borders of the derivative: When $I(.)$ is linear, $U_{11} = 0$, 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, $U_{22} = 0$, we have $\frac{\partial \tau_j^*(S_j)}{\partial S_j} = 0$. In such case the tax rates are always optimal, candidate steals from the public good. When both $I()$ and $H()$ are strictly concave the derivative is between 0 and 1, and thus, the effect of corruption is both lower than optimal public good levels *and* higher than optimal taxes. The kinks in the figure that we encounter in two quasilinear 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.

[27] By no extreme platforms assumption corner solutions have been ruled out.

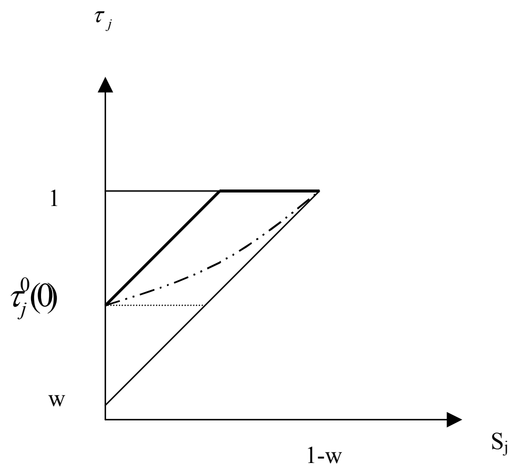

Figure 6:

Thus, a candidate may find it good policy to supply optimal public good yet impose 100 percent taxes. We rule out those "extreme" platforms, i.e., platforms that when implemented voters have zero (public or private good) consumption, by assuming[28] that even in the quasilinear case[29], the utility becomes strictly concave and the marginal utility goes to infinity around an epsilon neighborhood of zero consumption.[30] Hence, the strategy space relevant to our analysis, $(\tau_j^*(S_j), S_j)$ is a curve in Σ_j, and its slope

[28] See page 5.

[29] The quasilinear form used both by Polo [1998] and Persson and Tabellini [2000], $U = c + H(G)$, does not satisfy that restriction. We also used quasilinear form in some of the exmaples, since it makes calculations much easier. Since our model is more general, by this way we also show how their results would change when other factors are included into the model. Both papers implicitly focus on interior equilibria, where both candidates offer lower than 100 percent taxes. We calculate the interior equilibrium and specify the necessary and sufficient conditions on other parameters of the model for interior equilibrium, when $H(G) = 2\theta\sqrt{G}$..

[30] Let us provide an example for the other quasilinear form, $U = I(c) + H(G)$ where $H(G) = G$. When we replace $H(G)$ with

$$H^\varepsilon(G) = \begin{cases} G + \sqrt{\varepsilon} \text{ for } G > \sqrt{\varepsilon} \\ 2\sqrt{G} \text{ for } G \leq \sqrt{\varepsilon} \end{cases}$$

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. This example gives an idea of how to eliminate extreme positions in equilibria.

is between zero and one.

7.6 Existence and Uniqueness of Equilibrium

Lemma 5 *Over* $\left(\tau_j^*(S_j), S_j\right)$, *the Marginal utility of corruption for candidate* j*, (8), is continuous and strictly decreasing in* S_j *and continuous 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. ■

Corollary 9 *The objective function,* $\rho_j R_j$ *is quasi-concave in* S_j *over* $(\tau_j^*(S_j), S_j)$.

Proof. Follows from Lemma 5. ■

Lemma 6 *The corruption levels of candidates are strategic complements,* $\frac{\partial s_j^p(S_k)}{\partial S_k} \geqslant 0$, *with inequality being strict when* $s_j^p(S_k) > 0$.

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

$$\frac{\partial s_j^0(S_k)}{\partial S_k} = \frac{-\frac{\partial^2 \rho_j R_j}{\partial S_j \partial S_k}}{\frac{\partial^2 \rho_j R_j}{(\partial S_j)^2}} = \frac{z_{jk}}{2z_{jj} + a_j R_j \frac{\partial^2 \mathbf{E}[U(c_i^j, G_j)]}{(\partial G_j)^2}\left(\frac{\partial \tau_j^*(S_j)}{\partial S_j} - 1\right)}. \tag{14}$$

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 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. Hence when $s_j^0(S_k) = 0$ we have $\frac{\partial s_j^0(S_k)}{\partial S_k} = 0$. ■

Proposition 10 *Reaction functions* $s_1^0(S_2)$ *and* $s_2^0(S_1)$ *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 7 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.

Note that $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} = \frac{\frac{\partial^2(\rho_j R_j)}{\partial S_j^* \partial S_k^*}}{\frac{\partial^2(\rho_j R_j)}{(\partial S_j^*)^2}}.\frac{\frac{\partial^2(\rho_k R_k)}{\partial S_k^* \partial S_j^*}}{\frac{\partial^2(\rho_k R_k)}{(\partial S_k^*)^2}}.$

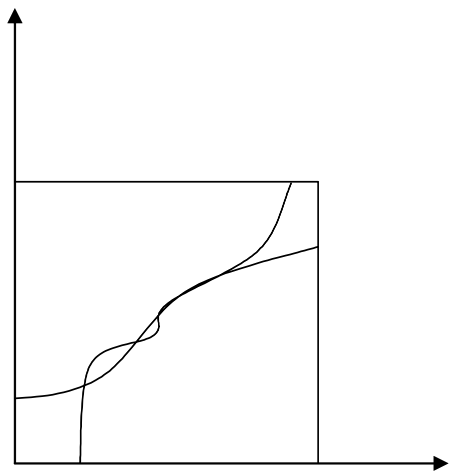

Figure 7

Let $z_1 = a_1 \frac{\partial \mathbf{E}[U(c_i^1, G_1^*)]}{\partial G}$ and $z_2 = a_2 \frac{\partial \mathbf{E}[U(c_i^2, G_2^*)]}{\partial G}$, then using the definition of $\frac{\partial s_j^0(S_k)}{\partial S_k}$ from Lemma 7, we have

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

Using the definition of z_{jk} from Lemma 6, we have

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

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

Corollary 11 *For later use note that the above result can be written as* $S_j^* > 0$ *and* $S_k^* > 0$ *implies that* $\frac{\partial^2(\rho_j R_j)}{(\partial S_j^*)^2}\frac{\partial^2(\rho_k R_k)}{(\partial S_k^*)^2} - \frac{\partial^2(\rho_j R_j)}{\partial S_j^* \partial S_k^*}\frac{\partial^2(\rho_k R_k)}{\partial S_k^* \partial S_j^*} < 0.$

Theorem 12 *The pure strategy Nash equilibrium for the political competition game exists and is unique.*

Proof. Existence.

The objective functions of candidate j is quasi concave in S_j over $\left(\tau_j^*(S_j), S_j\right)$. Then by the Theorem of Maximum the best response correspondence, $s_j^0(S_k)$ is continuous in S_k. No Extreme Platforms assumption implies that $0 \leqslant s_j^0(S_k) < 1$. By standard arguments there exists an equilibrium in pure strategies.

Uniqueness.

-If there exists an interior equilibrum: By Proposition 10, if there exists an interior equilibrium, then 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 steal 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 10 $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} < 1$ 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 where $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} \geq 1$, which is not possible by Proposition 10.

-If there exists a corner equilibrium: 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, since by continuity of reaction functions, we can not have two points of intersection following each other and both satisfying $\frac{\partial s_j(S_k^*)}{\partial S_k}.\frac{\partial s_k(S_j^*)}{\partial S_j} < 1$. ■

7.7 The equilibrium outcome as a function of parameters of the game.

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

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

Let $\Delta_j := \{a_1, a_2, L_1, L_2, w, E, g, I(.), H(.) : COND_j \text{ holds.}\}$. 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. Thus, let $\underline{S}_k$ denote the lowest amount stolen by candidate k that will not induce candidate j to steal, by continuity of reaction function it can also be defined 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

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

Let Υ_j be the set of parameters such that the above condition is satisfied, i.e., $\Upsilon_j := \{a_1, a_2, L_1, L_2, w, E, , g, I(.), H(.) : INEQ_j \text{ holds.}\}$.

Let ω be the set of the parameters of a particular game.

Lemma 7 *The unique Nash equilibrium of the game is:*

(a) $S_1^* = S_2^* = 0$, *iff for all* $j \in \{1, 2\}$, $\omega \notin \Delta_j$.

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

-either for all $j \in \{1, 2\}$, $\omega \in \Delta_j$

-or $\omega \in \Delta_j$, $\omega \notin \Delta_k$ *with* $\omega \in \Upsilon_j$.

(c) $S_j^* = s_j^0(0) > 0$ *and* $S_k^* = 0$ *iff* $\omega \in \Delta_j$, $\omega \notin \Delta_k$ *and* $\omega \notin \Upsilon_j$.

Proof. Note that ω is either in $\Delta_j \cup \Delta_k$ or in $(\Delta_j \cup \Delta_k)^C$. When $\omega \in (\Delta_j \cup \Delta_k)^C$ we have $s_j^0(0) = 0$ for both candidates. Then none of them steals when the rival steals zero. By Proposition 10 and by the continuity of reaction functions an interior equilibrium is not possible either. Hence the unique equilibrium is zero corruption by both candidates. If $\omega \in \Delta_j \cup \Delta_k$ then it is either in $\Delta_j \cap \Delta_k$ or in $\Delta_j \backslash \Delta_k$. When it is in $\Delta_j \cap \Delta_k$ both candidates are going to steal even when the rival does not, then by Proposition 10 and by continuity of reaction functions, there exist a unique equilibrium where both candidates steal positive amounts in equilibrium. If it is in $\Delta_j \backslash \Delta_k$ then it is either in $\Delta_j \backslash \Delta_k \cap \Upsilon_j$ or in $\Delta_j \backslash \Delta_k \cap (\Upsilon_j)^C$. When $\omega \in \Delta_j \backslash \Delta_k \cap \Upsilon_j$ by Proposition 10 and by continuity there only exist a unique interior equilibrium. The last case is $\omega \in \Delta_j \backslash \Delta_k \cap (\Upsilon_j)^C$. Now candidate k does not steal when candidate j steals $s_j^0(0) > 0$. Using Proposition 10 and continuity of reaction functions, we find that in the unique corner equilibrium only candidate j steals. ■

7.8 Analysis of Equilibrium Under Law Enforcement.

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

The function $\rho_j r_j(S_j)$ does not have any discontinuity. What we do is, to derive a "fake" reaction function for candidate j, $\sigma_j(S_k)$, from the optimization of $\rho_j r_j(S_j)$ and then take the relevant part of this reaction function, i.e.,

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

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

The "fake" reaction function, $\sigma_j(S_k)$, 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 10 can be applied to the intersection of $\sigma_j(S_k)'$s. It is this similarity that we use to extend the results from the analysis with no law enforcement. Since we know quite a lot about $\sigma_j(S_k)$, let us try to understand when it is relevant. The following Proposition shows that if it becomes relevant at some level of candidate k's corruption, it is always relevant for any higher level of corruption. By this proposition, $s_j^p(S_k)$ can have at most discontinuity and is strictly increasing in S_k as far as $s_j^p(S_k) > 0$.

Proposition 13 *If* $\rho_j(\sigma_j(\widehat{S}_k)) r_j(\sigma_j(\widehat{S}_k)) = \rho_j(0, \widehat{S}_k) R_j(0)$ *for some* $\widehat{S}_k$ *with* $\sigma_j(\widehat{S}_k) > 0$ *then*

$\rho_j(\sigma_j(S_k)) r_j(\sigma_j(S_k)) > \rho_j(0, S_k) R_j(0)$ *for any* $S_k > \widehat{S}_k$.

Proof. Take any $\widehat{s}_k$ such that $\rho_j(\sigma_j(\widehat{s}_k)) r_j(\sigma_j(\widehat{s}_k)) \geq \rho_j(0, \widehat{s}_k) R_j(0)$. Let us note that both sides are continuously differentiable in $\widehat{s}_k$ and consider an infinitesimal increase in $\widehat{s}_k$. The derivative of $\rho_j(0, S_k) R_j(0)$ w.r.to S_k evaluated at $\widehat{s}_k$ is equal to

$-\frac{\partial \mathbf{E}[U_i^k()]}{\partial S_k} g R_j(0) > 0$. The derivative of $\rho_j(\sigma_j(S_k)) r_j(\sigma_j(S_k))$ w.r.to S_k evaluated at $\widehat{s}_k$ is

$[\frac{\partial \mathbf{E}[U_i^j]}{\partial S_j} \frac{\partial \sigma_j(\widehat{s}_k)}{\partial S_k} - \frac{\partial \mathbf{E}[U_i^k(\widehat{s}_k)]}{\partial S_k}] g r_j(\sigma_j(\widehat{s}_k)) + (L_j - pv) \frac{\partial \sigma_j(\widehat{s}_k)}{\partial S_k} \rho_j > 0$. We need to show that $\rho_j(\sigma_j(\widehat{s}_k)) r_j(\sigma_j(\widehat{s}_k)) \geq \rho_j(0, \widehat{s}_k) R_j(0)$ implies

$A = \frac{\partial \mathbf{E}[U_i^k(\widehat{s}_k)]}{\partial S_k} g R_j(0) + [\frac{\partial \mathbf{E}[U_i^j]}{\partial S_j} \frac{\partial \sigma_j(\widehat{s}_k)}{\partial S_k} - \frac{\partial \mathbf{E}[U_i^k(\widehat{s}_k)]}{\partial S_k}] g r_j(\sigma_j(\widehat{s}_k)) + (L_j - pv) \frac{\partial \sigma_j(\widehat{s}_k)}{\partial S_k} \rho_j > 0$

Note that (8) = 0, which is necessary for $\sigma_j(S_k) > 0$, implies that $\frac{\partial \mathbf{E}[U_i^j]}{\partial S_k} g r_j(\sigma_j(\widehat{s}_k)) + (L_j - pv)\rho_j = 0$. Thus

$A = \frac{\partial \mathbf{E}[U_i^k(\widehat{s}_k)]}{\partial S_k} g R_j(0) - \frac{\partial \mathbf{E}[U_i^k(\widehat{s}_k)]}{\partial S_k} g r_j(\sigma_j(\widehat{s}_k))$.

Since $\rho_j(\sigma_j(\widehat{s}_k), \widehat{s}_k) < \rho_j(0, \widehat{s}_k)$,

$\rho_j(\sigma_j(\widehat{s}_k)) r_j(\sigma_j(\widehat{s}_k)) \geq \rho_j(0, \widehat{s}_k) R_j(0)$ implies that

$r_j(\sigma_j(\widehat{s}_k)) > R_j(0)$. Hence $A > 0$. ■

If the "fake" reaction function is always relevant for both candidates i.e., if for all $j \in \{1, 2\}$ we have $\sigma_j(0) > 0$ and $\rho_j(\sigma_j(0)) r_j(\sigma_j(0)) > \rho_j(0) R_j(0)$, then the discontinuity in the objective function has no effect on the reaction functions, as shown in Figure 5. Then by the same arguments used in Proof of Theorem 1, the unique equilibrium is $S_1^* > 0$ and $S_2^* > 0$. When the fake reaction function is always irrelevant then the best response is simply staying clean for whatever the rival does, hence the unique equilibrium is no corruption. In those two cases, when $\sigma_j(S_k)$ is always relevant and never relevant, we have unique equilibrium as in the no law enforcement case. On the other hand the law enforcement does make a difference in some cases. There is a third possibility that for both candidates $\sigma_j(S_k)$ is sometimes relevant, i.e., an intermediary case where $\rho_j(\sigma_j(0))R_j(\sigma_j(0)) < \rho_j(0,0)R_j(0)$ yet there exists an $\widetilde{S}_k \in (0, \overline{S}_k)$ such that $\rho_j(\sigma_j(\widetilde{S}_k))r_j(\sigma_j(\widetilde{S}_k)) = \rho_j(0, \widetilde{S}_k)R_j(0)$. In that case the reaction function is discontinuous at $\widetilde{S}_k$. As shown in Figure 6 it is zero until $S_k = \widetilde{S}_k$ and then suddenly it jumps to $\sigma_j(\widetilde{S}_k) > 0$. The difference is that now the game can have multiple equilibria, one equilibrium where no candidate steals and another one where both steal. By an application of Proposition 10, the interior equilibrium is unique, (the intuition is that in the interior equilibrium it is the $\sigma_j(S_k)'s$ that intersect each other, and as Proposition 10 shows this can not happen twice in the interior). Then in the second equilibria no one steals.

7.9 Wage reform

Lemma 8 $\frac{d\mathbf{E}[\mathbf{W}]}{dw} = \sum_{j\in\{1,2\}} \rho_j a_j \frac{d\mathbf{E}[U_i(.)]}{dG}(1 + \frac{dS_j(w)}{dw})$

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

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

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

To simplify that let us note, $\frac{d\mathbf{E}[U_i((1-\tau_j(w))Y_i, G_j(w))]}{dw}$ 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

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

As a last step note that,

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

7.9.1 Calculations for $\frac{dS_j}{dw}$.

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 matrix,

$$\begin{bmatrix} \frac{\partial^2(\rho_1 R_1)}{(\partial S_1)^2} & \frac{\partial^2(\rho_1 R_1)}{\partial S_1 \partial S_2} \\ \frac{\partial^2(\rho_2 R_2)}{\partial S_1 \partial S_2} & \frac{\partial^2(\rho_2 R_2)}{(\partial S_2)^2} \end{bmatrix} \begin{bmatrix} \frac{dS_1}{dw} \\ \frac{dS_2}{dw} \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2(\rho_1 R_1)}{\partial S_1 \partial w} \\ -\frac{\partial^2(\rho_2 R_2)}{\partial S_2 \partial w} \end{bmatrix}$$

The solution is

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

By Corollary 2 $\frac{\partial^2(\rho_1 R_1)}{(\partial S_1)^2} \frac{\partial^2(\rho_2 R_2)}{(\partial S_2)^2} - \frac{\partial^2(\rho_1 R_1)}{\partial S_1 \partial S_2} \frac{\partial^2(\rho_2 R_2)}{\partial S_1 \partial S_2} > 0$ in the equilibrium. From here it is easy to show that

$\frac{dS_j}{dw} < -1$ iff $\frac{\partial f_{2j}}{\partial S_k} A_k - \frac{\partial f_{2k}}{\partial S_k} A_j < 0$ where $A_j = a_j g[1 - p - L_j + pv] \frac{\partial \mathbf{E}[U_i(\tau_j^*, G_j^*)]}{\partial G}$.

7.9.2 Effect of wages when only one candidate steals.

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

Thus,

Lemma 9 *For any* U,

If only candidate j *steals in the equilibrium,*
$\frac{dS_j}{dW} < -1$ *iff* $L_j - pv < \frac{1-p}{[1+(a_k \frac{\partial \mathbf{E}[U_i^k(G_k^o)]}{\partial G} / a_j \frac{\partial \mathbf{E}[U_i^j(G_j^*)]}{\partial G})]}$.