Political Corruption Traps

Marko Klašnja*
Andrew T. Little†
Joshua A. Tucker‡

February 19, 2016

Abstract

Academics and policy makers recognize that there are serious costs associated with systemic corruption. Stubbornly, countries or regions with high levels of corruption often seem resistant to attempts to make politics cleaner, remaining stuck in "corruption traps." Most existing theories concentrate on mutually reinforcing expectations of corrupt behavior among a fixed set of bureaucrats or politicians, implying that changing such expectations can lead to lower corruption. We develop several models that more fully characterize the *political* nature of corruption traps by also analyzing the behavior of voters and entrants, as well their interaction with incumbent politicians. We show that corruption traps can arise through strategic behavior of each set of actors, and particularly so through their interrelations. By linking politician, voter, and entrant behavior, we provide an explanation for why simply trying to change expectations among one set of actors is likely insufficient for eliminating corruption traps.

Key Words: Corruption, Voting, Multiple Equilibria, Persistence, Corruption Traps

*Edmund A. Walsh School of Foreign Service and Government Department, Georgetown University marko.klasnja@georgetown.edu

†Department of Government, Cornell University andrew dot little at cornell dot edu

‡Wilf Family Department of Politics, New York University joshua dot tucker at nyu dot edu

1 Introduction

The study of corruption – its causes, effects, and persistence – is a matter of great interest to both the scholarly and policy communities. Corruption is recognized as a threat to economic development, democratic consolidation, and human dignity. Consequently, the desire to reduce corruption has become an important goal of international financial institutions, civil society groups, and individual nations' foreign policies.

Two stylized facts related to corruption have been influential in how academics, policymakers, and others think about whether and how it is possible to attain these goals. First, the prevalence of corruption varies significantly across different contexts. This variation can be seen cross-nationally (e.g. Treisman, 2000; Montinola and Jackman, 2002), but also *within* countries as well: for example, there are large differences in corruption across American states (Alt and Lassen, 2003; Glaeser and Saks, 2006) and cities (Banfield and Wilson, 1963; Menes, 2006); between Southern and Northern Italy (e.g. Banfield and Banfield, 1958; Golden and Picci, 2005); and across states in India (Transparency International India, 2005). Such variation gives reformers hope that corruption can indeed be reduced, that "high corruption" countries or regions can be turned into "low corruption" countries or regions.

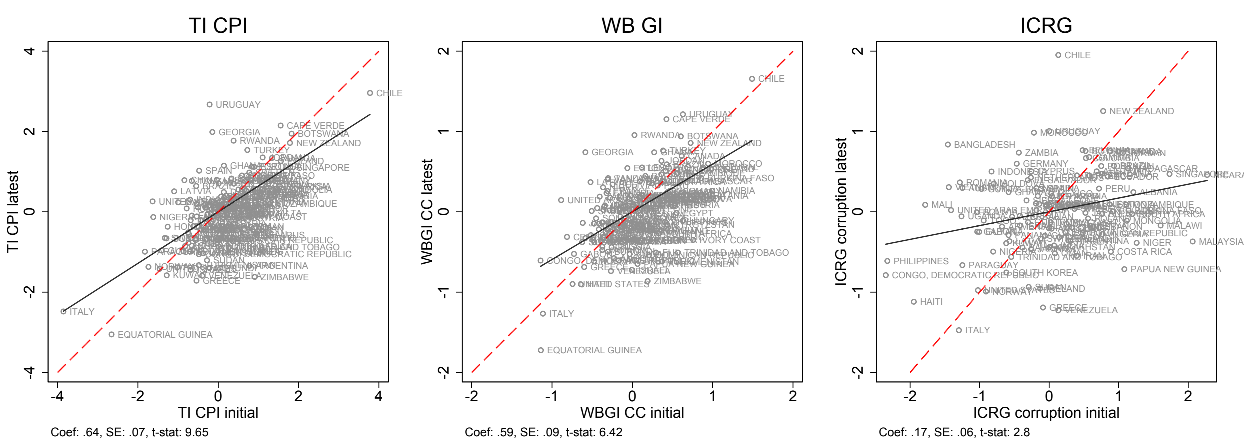
However, it is also the case that high or low levels of corruption tend to be very persistent: we rarely observe high corruption countries or regions becoming low corruption countries or regions, or visa-versa (Becker et al., Forthcoming; Damania, Fredriksson, and Mani, 2004; Fisman and Miguel, 2007).[1] While it is difficult to measure persistence with existing data, Figure 1 shows that after controlling for the common correlates of corruption,[2] the correlation

[1] Two important exceptions are the small city-states of Singapore and Hong Kong; see Quah (1994). We should note that we are not taking a very long view – even countries like the United States had considerably more corruption in the 19th Century. However, we believe it is not controversial to say that very few countries have moved much on corruption relative to other countries in the past several decades.

[2] The covariates, taken from Treisman (2007), are: GDP (PPP) in 1999, geographical size (logged area), British, French and other colonial legacy, British, French, Scandinavian, and German legal legacy, the share

between a country's initial and latest score on the well-known perceptual indices – spanning between 20 and 30 years – is quite strong.[3] The correlation is positive, statistically significant, and quite strong, especially for the Transparency International Corruption Perception Index and World Bank Control of Corruption index.

Figure 1: Persistence in corruption perception indices



Each panel is a partial residual plots showing the correlation between the initial and the latest value of the indicated corruption perception indices (the Transparency International's Corruption Perception Index, the World Bank's Control of Corruption Governance Indicator, and the International Country Risk Guide's (ICRG) corruption index, from left to right respectively), after controlling for the common correlates of corruption. The dashed red line is the 45-degree line.

One plausible explanation for this persistent variation in corruption levels is that the dynamics underlying corruption allow for multiple equilibria. The existence of multiple equilibria can lead to differences among similar countries and regions driven by whether they play the "low-corruption" or "high-corruption" equilibrium; if countries or regions consistently play the same equilibrium, these differences will persist over time.

---

of Protestants, Catholics and Muslims in 1980, ethnolinguistic fractionalization in 1985, length of democracy, dummies for world regions, and a dummy for federal states.

[3] Pritchett, Woolcock, and Andrews (2010) argue that even if today's countries with the weakest governance were to improve in line with the past experience of the 20 fastest improving countries, getting to the level of a median country such as Tanzania or Guatemala would take between 15 and 35 years.

This explanation has received significant theoretical attention. Most of the previous accounts, however, focus on interactions between a fixed set of bureaucrats or politicians, proposing that incentives to engage in corrupt behavior may increase as government officials come to believe that other government officials are also corrupt (e.g. Andvig and Moene, 1990; Lui, 1986). These mutually reinforcing beliefs can then lead to multiple equilibria: a low-corruption equilibrium where bureaucrats refrain from corruption when they believe that others will do so, and a high-corruption equilibrium where they engage in corruption in the belief that others will be corrupt as well.

We build on this approach, but highlight two important shortcomings of the existing literature. First, we argue that to describe the *politics* of multiple equilibria of corruption more fully, it is important to focus not just on a fixed set of politicians or bureaucrats, as in most of the previous models, but also on voters and potential candidates for political office (“entrants"), as well as the interaction among these three sets of actors. In principle, corrupt politicians can be voted out of office, but the empirical work on voting behavior has shown that in some contexts voters are willing to condone or overlook corruption (Banerjee and Pande, 2007; Klašnja and Tucker, 2013) while in other contexts corrupt politicians are punished (Ferraz and Finan, 2008). Moreover, multiplicity of equilibria may derive not just from the variation in incentives of incumbent politicians to be corrupt, as in most previous models, but in the variation of the type of candidates willing to run for office. When the majority of entrants are likely to be corrupt, voters may have little choice.

A second shortcoming of much of the existing literature is that high-corruption equilibria – driven mainly by expectations of how corrupt others will be – suggest a particular policy prescription of how to root out corruption: an intervention that rapidly changes the corruption expectations of a fixed set of politicians. While such interventions are not costless, they seem feasible.[4] And yet, this policy implication seems inconsistent with the fact that

[4] Abbink and Serra (2012, p.6) summarize this common policy prescription well: “As a result of the

corruption is so persistent across countries and regions. By focusing on voters and entrants as well as politicians, and on the interaction among them, our models exhibit more of the strategic complementarities that lead to multiple equilibria than past models suggest, highlighting why acting on politicians' expectations alone may not be enough to reduce corrupt behavior.[5]

We begin by presenting a simple model where politicians, who vary based on their predisposition to engage in corrupt behavior, choose a level of corrupt activity when the benefits from doing so are a function of the degree to which other politicians choose to engage in similarly corrupt behavior. This model has two important properties that motivate our subsequent analysis. First, the relative value of holding office for a politician is higher for those predisposed to corruption, and in line with previous work on multiple equilibria of corruption, the size of this "politician corruption differential" is increasing in the level of corruption. Second, as politics becomes more corrupt, citizens can become more tolerant of corrupt politicians, and thus less apt to vote them out of office.

We then proceed to show how these results affect behavior by voters and potential candidates. In our main model, voters observe the ability and corruption predisposition of their current representative and decide whether to keep or replace her. If voters become more tolerant of corrupt representatives (or start to prefer corrupt representatives) as others are more corrupt, there can again be multiple equilibria: one equilibrium where voters are gen-

strategic complementarities existing in the costs associated with corruption, multiple corruption equilibria exist and a country could converge to the bad equilibrium, i.e., a "corruption trap" from which it is extremely difficult to escape. The main policy implication of these theories is that when corruption becomes systemic, i.e., when the bad equilibrium is reached, a "big push" is required to escape the trap, that is, a combination of policies acting on the actual probability of detection, the penalty associated with corruption, the wages of public officials, and individuals' general beliefs about the extent of corruption among both public officials and ordinary citizens."

[5] This sentiment is echoed by Daron Acemoglu and James Robinson: "[...] corruption is a way for many economists and policymakers to talk about bad political outcomes without talking about politics. As long as the discussion is not about politics, there can always be a simple, non-political solution, often designed and operated by some impartial clever politicians, advisers or economists." http://whynationsfail.com/blog/2012/11/13/a-couple-of-economists.html.

erally willing to retain corrupt politicians, and one where voters generally replace corrupt politicians (as it is harmful to have a corrupt representative if few others are corrupt). We call the former situation a "political corruption trap."

Finally, we draw on Caselli and Morelli (2004) in deriving a model that examines a similar complementarity between potential entrants, where those predisposed to be corrupt are more apt to run when they expect others to be corrupt as well. Importantly, we also show that when the dynamics from the voter and politician models are combined with the dynamics of the entry model, the potential for a political corruption trap is generally even greater.

These results highlight that political corruption traps arising in our models may be harder to escape than the high-corruption equilibria in most of the existing literature. When the corruption trap is driven by the political forces we model, policy interventions would need to act not just on politicians' expectations, as prescribed by most previous models, but also on the expectations of voters and potential entrants to the political arena. We believe that this is a harder problem than coordinating the expectations of bureaucrats or politicians alone. Moreover, our results also suggest that even changing the expectations of all three sets of actors simultaneously may be insufficient to reduce corruption unless it is possible to remove many corrupt politicians simultaneously. This suggests that political and electoral institutions that affect the way politicians are elected may play a part in sustaining political corruption traps. For example, staggered elections and multiple tiers of representation may make it harder to replace a corrupt political elite wholesale, potentially perpetuating a high-corruption equilibrium through the dynamics we model.

2 Literature

In most theoretical studies that have previously examined the variation in the level and persistence of corruption through the multiplicity of corruption equilibria, the multiplicity arises through strategic complementarities stemming from the interplay between heterogeneous incentives to engage in corruption, some form of externality from committing a corrupt act, and the beliefs about whether others will engage in corruption. In Andvig and Moene (1990), the high incidence of corruption lowers the search costs of the briber for a willing bribee, and lowers the probability of detection to the bureaucrats from engaging in corruption (see also Lui, 1986). In Mishra (2005), the cost of compliance increases as others become increasingly non-compliant, lowering the benefits from being "clean" (see also Mauro, 2004). In Cadot (1987), the high prevalence of corruption implies that higher-level bureaucrats are more tolerant of lower-level corruption; this in turn further encourages lower-level bureaucratic corruption since getting caught is less likely to result in being removed from office.[6]

We build on and depart from these models in two important ways. First, as mentioned above, these models leave out much of the *politics* of multiple corruption equilibria. Multiple equilibria arising through the strategic complementarities among bureaucrats (Cadot, 1987; Lui, 1986), between bureaucrats and citizens (Andvig and Moene, 1990), or among firms and buyers (Mishra 2005; see also Tirole 1996; Hauk and Saez-Marti 2002), miss the fact that the control of corruption is fundamentally a political process. Bureaucrats are ultimately accountable to politicians,[7] who in democracies are accountable to voters.[8]

The fact that corrupt politicians can be voted out of office could in principle make it easier

[6]Others, such as Tirole (1996) and Hauk and Saez-Marti (2002) also focus on a limited set of actors (firms and buyers), but model the prevalence of corruption as path-dependent. They show that if the actions and reputations of individuals are not perfectly observable then past corruption can create a general sense of suspicion toward individuals and the groups to which they belong; this in turn lowers the incentives of the new group members to remain honest or to transmit a norm of honesty to the new generations, thus "ratcheting up" corruption in the future.

[7]Corruption traps may also be influenced by a distinct set of agency problems in the relationship between bureaucrats and politicians, as evidenced by bureaucratic resistance to public sector reforms, lobbying geared towards bureaucrats, and bureaucratization of politics. For simplicity and tractability, we abstract from these questions and leave them for future research.

[8]We limit the applicability of our models to democracies with meaningfully competitive elections. For the study of anti-corruption efforts in autocracies, see for example Hollyer and Wantchekon (2015).

to eliminate the high corruption equilibria. However, empirical research on voters' reactions to corruption shows that not all electorates unconditionally punish corruption.[9] Klašnja and Tucker (2013) conduct a survey experiment examining voters' reactions to economic performance and corruption in a low-corruption country (Sweden) and a high-corruption country (Moldova). They find that in Sweden voters react negatively to corruption regardless of the state of the economy, whereas in Moldova voters react negatively to corruption *only* when the state of the economy is also poor; when economic conditions have improved, corruption is considerably less important.[10] Elsewhere, scholars have similarly found that in higher-corruption countries, voters are often insensitive to corruption (Banerjee and Pande, 2007; Manzetti and Wilson, 2007; Zechmeister and Zizumbo-Colunga, 2013), sometimes even when given precise information about it.[11] Voters in low corruption countries tend to punish corruption more consistently (Hirano and Snyder, 2012; Peters and Welch, 1980; Reed, 1999; Welch and Hibbing, 1997).[12]

The second important departure in the models we present below is that we focus on how the multiplicity of corruption equilibria is driven by the *types* of politicians – both those in office and those willing to enter politics. The closest model to ours in this sense is Caselli and Morelli (2004), who argue that the current share of "bad" politicians will affect the relative returns to office-holding for other bad politicians. As a result, there can be two equilibria:

[9]The persistence of corruption as a function of voter behavior may be mediated by the ability of political parties to increase the salience of political corruption, or their reliance on vote-buying and patronage. We abstract from these issues and leave them for future research.

[10]Consider for example the "Toblerone Affair," during which Swedish MP Mona Sahlin was forced to withdraw her candidacy for leader of her party after a series of revelations that grew out of the fact that she had purchased a *candy bar* (we can only assume for personal consumption) with her official government credit card. (`http://blogs.reuters.com/global/2009/05/11/expenses-they-order-this-matter-differently-in-sweden/`).

[11]For example, de Figueiredo, Hidalgo, and Kasahara (2012) find no effect of an informational intervention about an incumbent's corruption on voters' electoral behavior in a mayoral race in Brazil, while finding an effect of a similar information dissemination about the challenger's involvement in corruption. Some studies, however, do find more consistent evidence of voters punishing corruption in high-corruption countries (Chong et al., 2015; Ferraz and Finan, 2008).

[12]However, even voters in low-corruption countries do not always sanction corruption (Barberá, Fernández-Vázquez, and Rivero, Forthcoming; Chang, Golden, and Hill, 2010).

one where politics is dominated by good politicians, encouraging other good politicians to run, and one where bad politicians dominate politics. Our paper builds on this argument by primarily focusing on the strategic complementarities in tolerating corruption among voters, while also exploring the interaction between voters' selection decisions and entry dynamics.[13]

Several other important previous contributions have also investigated the importance of political incentives and heterogeneity in politician types for persistence of corruption or the quality of accountability more broadly. The possibility of multiple accountability equilibria arising out of the interaction between voters and politicians has also been explored in recent related papers by Ashworth, Bueno de Mesquita, and Friedenberg (2014) and Svolik (2013). Both papers show that multiple equilibria can arise due to a trap of bad or pessimistic expectations on the part of voters: a bad equilibrium where politicians exert little effort when little is expected of them, and a good equilibrium where politicians exert high effort when it is expected.[14] Svolik (2013) also considers the effect of voter and incumbent behavior on entry decisions, showing that in a good equilibrium, bad entrants are discouraged from running for office. A central conclusion of these models is that, like much of the literature on corruption, accountability could be improved by changing expectations. While we echo this result in our analysis as well, we focus more closely on how daunting this task of changing expectations may be to bring about a reduction in corruption. By linking our study more directly with the literature on strategically-induced multiplicity of equilibria in models of corruption mentioned above, we generate predictions and policy implications not addressed by these studies.[15]

[13]The details of our argument differ from Caselli and Morelli (2004) in various other ways. Most significantly, the reason why traps with bad politicians are hard to escape in Caselli and Morelli (2004) is that politicians in office manipulate the value of office-holding in the future, and hence good politicians will choose a value that encourages more good politicians to run while bad politicians choose a value conducive to more bad politicians running. Our model does not rely on this type of argument; instead, entry decisions in our model depend primarily on the existing share of corrupt incumbent politicians.

[14]Meirowitz and Tucker (2013) make a related argument about protest against "bad" leaders: once citizens are sufficiently pessimistic about the pool of candidates, they simply conclude that politicians are crooks and do not pay the cost to protest in an effort to replace the current incumbent with a different candidate.

[15]Also, the details of our analysis differ in other ways from these papers. Most notably, the multiplicity

3 Formal Models

3.1 Politicians

We first present a model of the decision to engage in corrupt behavior among politicians. The model distills several ideas from the extant literature on corruption into a simple form, which will allow us to build more complex models that also account for voter behavior and candidate entry.

There are $N > 1$ politicians indexed by i, as would be the case in a legislature with N single member districts. Politicians are characterized by an ability $a_i > \underline{a} > 0$, where $\underline{a}$ is the "minimum" ability level, and a predisposition for corruption $c_i \in \{0, 1\}$. Politicians with $c_i = 0$ are not predisposed towards corruption (or "clean"), which we model by giving them no incentive to engage in corrupt behavior, while those with $c_i = 1$ will have incentives to be corrupt. This heterogeneity could be microfounded by having varying costs or benefits to corrupt behavior due to ability or distaste for illegal activity. Denote the proportion of $c_i = 1$ politicians with $\rho \in [0, 1]$. The distribution of ability does not affect the equilibrium analysis in this section.

Politicians have an amount of effort to exert equal to their ability, which they allocate between corrupt activities $x_i \geq 0$ and other, non-corrupt, activities $y_i \geq 0$. That is, they choose x_i and y_i subject to a "budget constraint" $x_i + y_i \leq a_i$. Let the utility for politician i be:

$$u^p(x_i, y_i; c_i) = y_i + c_i g(x_i, \overline{x})$$

in both Ashworth, Bueno de Mesquita, and Friedenberg (2014) and Svolik (2013) derives in large part from a voter's ability or willingness to use a politician's action to infer their hidden type. In our analysis, incumbent's type is observed by the voter, and so multiplicity arises from other sources. Also, the model in Ashworth, Bueno de Mesquita, and Friedenberg (2014) considers voter-politician relations in isolation; our results discussed below show that an accountability trap in one constituency can change the type of politician running in other constituencies. Further, in Svolik (2013), an important feature that gives rise to multiplicity is the infinite horizon of the interaction between the actors, which is not the case in our analysis.

The first term captures the benefits to non-corrupt activity, which for simplicity are equal to the effort spent. Politicians with $c_i = 0$ get no benefit from corrupt behavior, while those who are predisposed toward engaging in corrupt behavior get a partial payoff that is a function of their corruption choice (x_i) and the average level of corruption chosen by others ($\overline{x}$), captured by the $g(x_i, \overline{x})$ term. We use the common notation where g_1 is the partial derivative of g with respect to its first argument (in this case x_i), g_2 the partial derivative of g with respect to the second argument, g_{11} is the second derivative with respect to the first argument, g_{12} is the cross-partial derivative, etc. Our key assumptions about the returns to corruption are the following:

Assumption 1. *g is continuous and twice-differentiable, with (i) $g(0, \overline{x}) = 0$, (ii) $g_1 > 0$, (iii) $g_{11} < 0$, (iv) $g_2 > 0$, and (v) $g_{12} > 0$.*

Part (i) implies there is no return to corruption if the politician allocates no effort to this end. Parts (ii)-(iii) imply the returns to corruption are increasing and concave in the effort allocated; that is, there are diminishing marginal returns to corrupt behavior.

Parts (iv)-(v) imply that the return to corruption and the marginal return to more corrupt behavior are both increasing in how corrupt other politicians are. This is consistent with the literature on strategic complementarities in corrupt behavior described previously in Section 2. The marginal returns to corruption may be increasing in the overall corruption level because the higher proportion of corrupt officials decreases the search and transaction costs of the bribers, thus increasing the demand for bribes (Andvig and Moene, 1990).[16] Also, the probability of detection may decrease as overall corruption increases if there is a fixed budget available for monitoring (Lui, 1986), or because it is more difficult to find honest officials who would supply the information (Cadot, 1987). Moreover, the costs of violating the norm of not being involved in corruption may decline as an increasing number of others

[16] Related, Corbacho et al. (Forthcoming) show that citizens are more likely to engage in corruption when they are induced to perceive a higher overall level of corruption.

do not adhere to that norm (Mishra, 2005).[17]

We search for a symmetric Nash Equilibrium of this game, which is characterized by a function mapping the corruption level and ability of each politician to an allocation of effort (x_i, y_i), where these strategies are mutual best responses.

As $c_i = 0$ types get no benefit from corruption, they allocate all of their effort to non-corrupt activities. The payoff for a corrupt politician is increasing in both x_i and y_i, so they always use all of their effort budget: $x_i + y_i = a_i$. Hence we can write the optimization problem for a corrupt politician as selecting x_i to maximize:

$$u^p(x_i; c_i, a_i) = a_i - x_i + g(x_i, \overline{x})$$

Treating the corruption level of others as fixed, this gives a first order condition:

$$g_1(x_i, \overline{x}) = 1$$

Note that this is independent of a_i, so as long as the corrupt politicians always choose an interior $x_i \in (0, a_i)$ (which is ensured by the assumptions below), we can characterize the equilibrium strategy solely by the corruption level chosen by those with $c_i = 1$, and call this x^*. Since proportion ρ of politicians are corrupt, the equilibrium condition for an interior x^* is:

$$g_1(x^*, \rho x^*) = 1 \tag{1}$$

The marginal benefit to allocating an additional unit of effort to non-corrupt activities

[17] It is possible that corruption levels of different politicians could be strategic substitutes if, for example, there is a fixed amount of money to steal over which politicians compete. However, in our view the theoretical and empirical evidence suggests complementarities are more likely to dominate, or at least characterize enough types of corrupt behavior in many contexts for there to be value in writing a model with this assumption.

is always one, so this condition states that corrupt politicians get the same marginal benefit from allocating more effort to corruption, given the fact that other corrupt politicians also choose x^*. The derivative of $g_1(x, \rho x)$ with respect to x is:

$$\frac{\partial g_1(x, \rho x)}{\partial x} = g_{11}(x, \rho x) + \rho g_{12}(x, \rho x)$$

which may be positive or negative since $g_{11} < 0$ but $g_{12} > 0$. In words, this is because there are diminishing returns to corrupt behavior when keeping the behavior of others fixed, but in a symmetric equilibrium increasing the corruption level of others also increases the returns to corruption. If the second effect dominates for some part of the parameter space, it is possible for this expression to be increasing. In the extreme, if this expression is increasing everywhere, there will never be an interior corruption level: corrupt politicians will either allocate none or all of their time to corrupt activities. In fact, both of these allocations may be an equilibrium choice due to these strategic complementarities, which is analogous to how other models result in a "corruption trap" with multiple equilibria.

To bring attention to a mechanism through which high and low corruption equilibria emerge *for other reasons*, we assume that the diminishing returns effect of higher corruption levels is larger than the strategic complementarity effect ($g_{11}(x, \rho x) < -\rho g_{12}(x, \rho x)$), so that $\frac{\partial g_1(x,\rho x)}{\partial x} < 0$. To reduce the cases to consider, we also assume $g_1(0, 0) > 1$ (which ensures $c_i = 1$ types will engage in some corrupt behavior even if no other politician does) and $g_1(\underline{a}, \rho\underline{a}) < 1$ (which ensures $c_i = 1$ types do not allocate all of their time to corruption) for all ρ, where $\underline{a}$ is the minimum ability level. This gives the following:

Proposition 1. *The politician model has a unique equilibrium level of corruption x^* chosen by the $c_1 = 1$ types characterized by equation 1. In equilibrium:*

(i) corrupt politicians get a higher payoff than non-corrupt politicians,

(ii) the optimal level of corrupt activity for those predisposed to corruption is increasing in

ρ, and

(iii) the relative payoff of a corrupt politician is increasing in ρ.

All proofs can be found in the Supplemental Appendix.

For later derivations, it will be useful to express the difference between the equilibrium payoff to a corrupt politician and non-corrupt politician – or the *politician corruption differential* – as:

$$d^p(\rho) = [a_i - x^* + g(x^*, \rho x^*)] - a_i = g(x^*, \rho x^*) - x^*$$

Proposition 1 implies that $d^p(\rho)$ is positive and increasing in ρ. A convenient aspect of our setup is that while more able politicians get a higher payoff in general, the a_i terms drop out when considering this differential, meaning $d^p(\rho)$ is not a function of politician ability.

As we will focus on voter behavior in the next section, now consider the payoff to a citizen represented by politician i. Suppose the payoff to being represented by a politician making effort allocation x_i, y_i is:

$$u^v(x_i, y_i) = y_i + bg(x_i, \overline{x})$$

where b scales the citizen's (or, mnemonically, "(v)oter's") utility from the corrupt politician's payoff. For negative b, the citizen's payoff is decreasing in the corruption payoff earned by the politician. Voters may be directly harmed by the corrupt behavior of their representatives (in addition to the foregone effort to more productive ends) for several reasons, such as the potential negative effects of corruption on outcomes like investment (Mauro, 1995) and human capital accumulation (Reinikka and Svensson, 2004), and given the fact that citizens frequently view corruption as an important problem facing their country.[18]

[18]See for example: http://www.gallup-international.bg/en/Publications/71-Publications/181-

However, even if politics being corrupt generally harms citizens, they may sometimes *benefit* when their representative is corrupt, implying $b > 0$.[19] While $b \geq 1$ is unlikely (i.e. that the voter benefits above and beyond the politician's own corrupt payoff; we will not consider it hereafter), $b \in (0, 1)$ implies that the voter's payoff is increasing in the corruption payoff earned by the politician, but scaled downward. One way to interpret this is that for every $b + 1$ dollars earned, the politician keeps 1 for herself and distributes $b < 1$ to the citizens. And indeed there are instances in which reelection-seeking politicians would funnel some of the corrupt rents to their electorate or political parties with vote buying campaigns and patronage (Singer, 2009). For example, *coronéis*, local political bosses in Brazil, have long used corrupt revenues extracted locally or received from politicians higher up in the hierarchy to broker votes through vote-buying, patronage and influence (Gingerich, 2014; Mainwaring, 1999).[20] Research by Transparency International indicated that in Poland and Ukraine, out of a 5% kickback, about 4.5% ends up in private accounts, while the rest is possibly channeled for political purposes (Transparency International, 2004, p.23).

In equilibrium, the payoffs from having a non-corrupt and corrupt representative are then:

$$u^v(c_i = 0; x^*) = a_i$$

$$u^v(c_i = 1; x^*) = a_i - x^* + bg(x^*, \rho x^*)$$

`Corruption-Tops-the-List-as-the-World%E2%80%99s-Most-Important-Problem-According-to-WIN-Gallup-International%E2%80%99s-Annual-Poll`.

[19] While we assume that a $c_i = 0$ type politician is predisposed to corruption, in reality a corrupt politician need not necessarily be the one that personally profits from bribes, but one who is willing to work within the confines of a corrupt system. For example, Mayor Daley of the infamous Chicago machine, President Harry Truman, a loyal member of the notoriously corrupt Pendergast machine in Kansas City, Manmohan Singh, the 14th Prime Minister of India during such high-profile corruption scandals as the 2G spectrum case, or Dilma Rousseff, the current president of Brazil and the former board chairman of Petrobras before the eruption of the ongoing Petrobras scandal, have all been thought to be personally honest.

[20] High corruption may also reinforce the demand for clientelism. de la O (2015) demonstrates that when corruption is high, voters in Mexico are more likely to believe that vote selling is prevalent in their communities and more willing to sell their own votes.

Analogous to the politician corruption differential, define the *voter corruption differential* as:

$$d^v(\rho) = u_i^v(c_i = 1) - u_i^c(c_i = 0) = bg(x^*, \rho x^*) - x^*.$$

As with the politician corruption differential, this is not a function of a_i: voters always prefer more able politicians, but the ability does not affect the relative value of having a corrupt or clean representative.

If b is small or negative, the citizen prefers a representative that does not engage in corrupt behavior, i.e., $d^v < 0$. On the other hand, as $b \to 1$, the citizen payoff is the same as the politician payoff, and since corrupt politicians get a higher payoff, this means the citizen prefers a corrupt representative. By the continuity of the payoff function there must be a critical value of b between zero and one that determines whether the citizen prefers a corrupt politician:

Proposition 2. *There exists a* $\hat{b} \in (0, 1)$ *such that the payoff to having a corrupt politician is higher than having a non-corrupt politician if and only if* $b > \hat{b}$.

The intuition behind this result is that even if citizens only get a small benefit from their representatives' corrupt behavior, they may prefer this fraction of the proceeds to having a completely clean politician in a highly corrupt environment.

By highlighting this contingency we do not mean to imply that voters usually or even often strictly prefer a corrupt politician. Still, this possibility is consistent with a widespread view that when corruption is prevalent, corrupt politicians may be better able to achieve results than honest politicians. For example, in Russia in the early 1990s, getting a business started often required bribing an entire chain of administrators and authorities, from the local legislature, to the local executive branch, the central ministry, the fire authorities, the water authorities, and so on (Shleifer and Vishny, 1993); it is easy to imagine that a corrupt politician could be better able to navigate such an environment than an honest one.[21] In a generally corrupt system of private infrastructure construction through city-granted franchise in the U.S. at the end of the 19th Century, streetcar lines were built more quickly and covered more miles in those U.S. cities where private streetcar companies were able to speculate in real estate along the streetcar lines (such as the dealings of Henry Huntington in Los Angeles, and Francis Newland in Maryland; Jackson, 1985). Aravind Adiga, a Booker prize winning Indian writer, put this notion poignantly in the context of Indian politics:

> "Even worse than the fake honest politician, is the genuinely honest one. In India, corruption is a form of patronage; a politician or bureaucrat who takes the bribe then has to let it cascade among a series of lesser bureaucrats and elected officials, who will make sure that the pet projects are completed. An honest politician has no goodies to toss around. This limits his effectiveness profoundly [...] Would you really want an honest man representing you in parliament? The neighbouring constituency may get that new cricket stadium you were supposed to."[22]

To be clear, our modeling approach does not assume that b is either positive or negative, only that it is reasonable to expect that b could be either positive or negative depending on the circumstances. Further, what is more important for the possibility of political corruption traps is how the relative value of having a corrupt representative *changes* as ρ increases. This is given by:

$$\frac{\partial d^v(\rho)}{\partial \rho} = -\frac{\partial x^*}{\partial \rho} + b\frac{\partial g(x^*, \rho x^*)}{\partial \rho}$$

[21] Sayings in several languages echo this notion: "el que no tranza, no avanza," Spanish for "he who does not cheat, does not get ahead;" "soldi e amicizia vincono la giustizia" Italian, loosely meaning that money and friendship beat justice, and "Не подмажешь не поедешь," Russian, loosely meaning "no grease, no go" (all quoted in Simpser 2013, p. 3).

[22] "Indians' worst fear: the honest politician," *The Guardian*, 30 July 2008, http://www.theguardian.com/commentisfree/2008/jul/30/india.scamsandfraud.

There are two competing effects on voter welfare. First, as the level of corruption increases, representatives for whom $c_i = 1$ allocate less effort to non-corrupt activities (the $-\frac{\partial x^*}{\partial \rho}$ term), reducing the relative value of having a corrupt politician. However, the returns to corruption increase, and, if citizens enjoy some of the fruits of corruption, they can benefit from this as well. Whether making politics more corrupt makes having a corrupt representative better or worse depends on which of these effects dominates.

When $b \leq 0$, citizens get no benefit from corruption, and so the only effect of increasing ρ is that corrupt politicians allocate less effort to non-corrupt activities, which hurts their constituents. As a result, $d^v(\rho)$ is decreasing. On the other hand, if $b = 1$, the citizens payoff is exactly the same as the politician, and part (iii) of proposition 1 is equivalent to stating that $d^v(\rho)$ is increasing. In particular, the relative value of having a corrupt representative is increasing in ρ if and only if citizens get a high enough benefit from corrupt activity:

Proposition 3. *If $b > \tilde{b}$ for some $\tilde{b} \in (0, 1)$, the citizen payoff to having a corrupt politician is increasing in ρ*

This result would only be reinforced if the returns to non-corrupt activities are diminishing as more politicians are corrupt and allocate their effort to corrupt behavior. That is, in this formulation the payoff to having a clean politician is constant, and the above result highlights the case when the citizen payoff to having a corrupt representative is higher than this constant (and potentially increasing in the level of corruption). As discussed above, it is plausible that the returns to non-corrupt activity are lower in a very corrupt environment, also raising the value of having a corrupt representative if others are corrupt as well.

It is also plausible that when corruption is rare citizens benefit little from corruption, but when corruption is more pervasive and open citizens benefit more from it. This need not even mean that citizens prefer corruption politicians to non-corrupt politicians in general, but simply that they are more likely to view corruption as less of a concern *vis a vis* other factors affecting the vote decision (see for example Klašnja and Tucker (2013)) when corruption is more prevalent.

3.2 Voters

As in section 3.1, in our voter model there are N districts. Each district has an incumbent politician, a challenger, and a representative voter. This environment could correspond to an electoral system where legislators are elected by plurality vote in single-member constituencies. In some countries, such legislators are vehicles for corruption. For example, while corruption appears to increase in magnitude up the political hierarchy in India, even state legislators on average extract substantial rents, as judged by their relative wealth accumulation (Fisman, Schulz, and Vig, 2014). However, legislators are often not the main corrupt actors, as unlike in the U.S. they are frequently not given a large amount of resources to work with, particularly in PR systems, where backbenchers are often little more than party foot soldiers (e.g. Golden and Picci, 2015).

Even in PR systems (and elsewhere), our model can still describe the choice of directly elected local executives, such as governors or mayors. Directly elected local executives are quite common, and there is a wealth of evidence of their involvement in corruption. For example, Ferraz and Finan (2011) estimate more than \$550 million in corruption from federal transfers (about 15% of all transfers) in Brazilian municipalities in 2004 alone. Coviello and Gagliarducci (2014) document widespread manipulation of public procurement in Italian municipalities, especially by longer-tenured mayors. In Indonesia, prominent local politicians running on anti-corruption platforms have been highlighting the entrenched local corruption, such as the former mayor of Soko and governor of Jakarta (and now president) Joko Widodo, Bu Risma in Surabaya, or Ridwan Kamil in Bandung. The Romanian integrity agency and its anti-corruption directorate initiated more than 1,000 corruption cases against county and local politicians between 2008 and 2013 (Klašnja, 2015). Cases of corrupt local governments abound in China, such as in Shenyang, the capital of Liaoning Province, where almost an entire local administration was engaged in malfeasance, from the mayor, deputy mayor, the chiefs of the tax, price control, land resource, construction, and state assets bureaus, to the head of the city's prosecutor's office (Sun, 2004, p. 127). More than two dozen municipalities in Hungary were recently involved in suspicious tenders for public lighting renovation, awarded uniformly to a firm belonging to the Prime Minister's son-in-law.[23]

In the first stage of the model, voters simultaneously decide whether to keep their incumbent or replace her with a challenger. Those who are elected then play the politician-level game described in Section 3.1. Again, we assume politicians are characterized by an ability $a_i \geq \underline{a} > 0$ and corruption level $c_i \in \{0, 1\}$. Assume politician abilities are drawn from a distribution F that admits a density f, such that $f(a) > 0$ for all $a > \underline{a}$. We assume that voters know the type of their incumbent representative but not the potential replacement (or "challenger").[24] Let the probability of a challenger being corrupt be given by a constant q.

The strategy for a politician is a mapping from their type to a corruption level x_i, and a strategy for a voter is a mapping from the type of their incumbent to a decision to keep the incumbent or vote them out in favor of the challenger.

Given the equilibrium politician behavior described above, we can write the citizen payoff as:

$$u^v(a_i, c_i; \rho) = a_i + d^v(\rho)c_i.$$

The first term implies that the voter's payoff is increasing in the ability of the politician they select. When d^v is negative, voters prefer a non-corrupt politician, and when d^v is positive they prefer a corrupt politician. As demonstrated in section 3.1, d^v can be either positive or

[23]See for example: http://www.direkt36.hu/en/2015/03/11/tiborcz-istvan-es-az-elios-innovativ-zrt-sikerei-ledes-kozvilagitasi-kozbeszerzeseken/.

[24]If voters also (partially) observe the type of the challenger their re-election decision will depend on the type of the challenger as well. However, citizens will still be more apt to keep incumbents who are more able and (sometimes) corrupt, so this is unlikely to change our main conclusions.

negative, and either increasing or decreasing in ρ.

The expected payoff to choosing a challenger is then $\mathbb{E}[a] + qd^v(\rho)$, and hence the voter prefers a replacement to an incumbent politician with (a_i, c_i) if:

$$a_i + d^v(\rho)c_i \leq \mathbb{E}[a] + qd^v(\rho)$$

In particular, this implies that the voter prefers a corrupt incumbent to the challenger if $a_i \geq \mathbb{E}[a] - (1-q)d^v(\rho)$ and a clean incumbent to the challenger if $a_i \geq \mathbb{E}[a] + qd^v$.

Our equilibrium definition in this section is:

Definition An equilibrium to the voter model comprises a symmetric corruption choice x^* and voter thresholds $\hat{a}_0^v$ and $\hat{a}_1^v$ that jointly solve:

$$x^* = \arg\max_x a_i - x + g(x, \rho\overline{x}) \tag{2}$$

$$\hat{a}_1^v = \mathbb{E}[a] + (1-q)d^v(\rho) \tag{3}$$

$$\hat{a}_0^v = \mathbb{E}[a] + qd^v(\rho) \tag{4}$$

where

$$\rho = \frac{q(1 - F(\hat{a}_1^c))}{q(1 - F(\hat{a}_1^e)) + (1-q)(1 - F(\hat{a}_0^e))} \tag{5}$$

Equation 2 implies that once in office politicians choose the corruption level derived in Section 3.1 (which can be summarized with a corruption choice x^* by all $c_i = 1$ types). Equations 3-4 imply that voters behave optimally given the strategies of politicians and other voters. Equation 5 implies that the proportion of corrupt politicians is equal to the proportion of retained politicians that are corrupt. One way to interpret this condition is that in the long-run all the politicians in office are good enough to be retained. For example, if all districts start with a random politician and then continually replace their incumbent based on the decision rules elaborated above (and incumbents never leave for exogenous reasons), the long-run proportion of corrupt politicians will be given by equation 5.

While this is a non-linear system with multiple endogenous variables that interact in complex ways, we can solve the for the equilibria of the model with a relatively simple algorithm that focuses on the proportion of politicians with a predisposition for corruption:

1. Propose an equilibrium proportion of corrupt politicians $\rho^* \in [0, 1]$.
2. Derive the optimal corruption level x^* and the resulting voter corruption differential $d^v(\rho^*)$.
3. Given $d^v(\rho^*)$, compute the entry rules $\hat{a}_0^v$ and $\hat{a}_1^v$ that generate a proportion of corrupt entrants by equation 5.
4. If this proportion is equal to ρ^*, then ρ^* (and the corresponding strategies derived in steps 2-3) comprise an equilibrium; if not there is no equilibrium with proportion of corrupt politicians ρ^*.

The analysis in the previous section accomplishes step 2, so what remains is steps 3-4.

To express the solution concisely, first define the probability of a randomly drawn politician being clean and of high enough ability to be "kept" as:

$$q_0^k(\rho) = (1-q)(1 - F(\mathbb{E}[a] + qd^v(\rho)))$$

and the analogous probability for a corrupt politician:

$$q_1^k(\rho) = q(1 - F(\mathbb{E}[a] - (1-q)d^v(\rho)))$$

We can then write the equilibrium condition as:

$$\rho = \frac{q_1^k(\rho)}{q_0^k(\rho) + q_1^k(\rho)} \equiv \overline{q}(\rho). \tag{6}$$

The right-hand side of equation 6 is continuous and bounded by $(0, 1)$, and so there must be at least one $\rho^* \in (0, 1)$ that meets equation 6. To determine whether intersection (and hence equilibrium) is unique, differentiate the right-hand side with respect to ρ, giving:

$$\frac{\partial \overline{q}}{\partial \rho} = \frac{(q_0^k(\rho) + q_1^k(\rho))\frac{\partial q_1^k}{\partial \rho} - q_1^k(\rho)\left(\frac{\partial q_1^k}{\partial \rho} + \frac{\partial q_0^k}{\partial \rho}\right)}{(q_0^k(\rho) + q_1^k(\rho))^2} = \frac{q_0^k(\rho)\frac{\partial q_1^k}{\partial \rho} - q_1^k(\rho)\frac{\partial q_0^k}{\partial \rho}}{(q_0^k(\rho) + q_1^k(\rho))^2} \tag{7}$$

Referring back to the definitions of $q_i^k(\rho)$, this can be written as:

$$\frac{\partial \overline{q}}{\partial \rho} = \frac{\partial d^v(\rho)}{\partial \rho} \frac{q(1-q)(q_0^k(\rho) f(\mathbb{E}[a] - (1-q)d^v(\rho)) + q_1^k(\rho) f(\mathbb{E}[a] + q d^v(\rho)))}{(q_0^k(\rho) + q_1^k(\rho))^2}$$

All terms other than $\frac{\partial d^v(\rho)}{\partial \rho}$ are always positive, so the sign of this expression is the same as the sign of $\frac{\partial d^v(\rho)}{\partial \rho}$. If $\frac{\partial d^v(\rho)}{\partial \rho}$ is negative – which again implies that when more politicians are corrupt, it is worse to have a corrupt representative, – then $\frac{\partial \overline{q}}{\partial \rho}$ is negative, which implies equation 6 must have a unique solution. However, if $\frac{\partial d^v(\rho)}{\partial \rho}$ is large enough – which implies that the relative benefit of having a corrupt politician increases quickly as there are more additional corrupt politicians – then the slope of this equation can be greater than one, which raises the possibility of multiple intersections and hence multiple equilibria, one with a relatively low share of corrupt politicians and one with a relatively high share of corrupt politicians.

Recall $\frac{\partial d^v(\rho)}{\partial \rho}$ being positive does not *require* that voters sometimes strictly prefer a corrupt politician. It is possible that $d^v(\rho)$ is always negative but the derivative is positive, meaning that voters always dislike corruption but become more tolerant of corrupt representatives when their prevalence increases.

It is difficult to provide additional analytical results about when such multiplicity of equilibria is possible. Therefore, we conclude this section with illustrations that highlight conditions under which the multiplicity occurs. In all of the illustrations that follow, ability is drawn from an exponential distribution with scale parameter 1. In the initial illustrations, the proportion of corrupt challengers is $q = .5$; later we allow it to be a function of ρ. What varies in the figures is the voter corruption differential d^v, which we take as exogenous.[25]

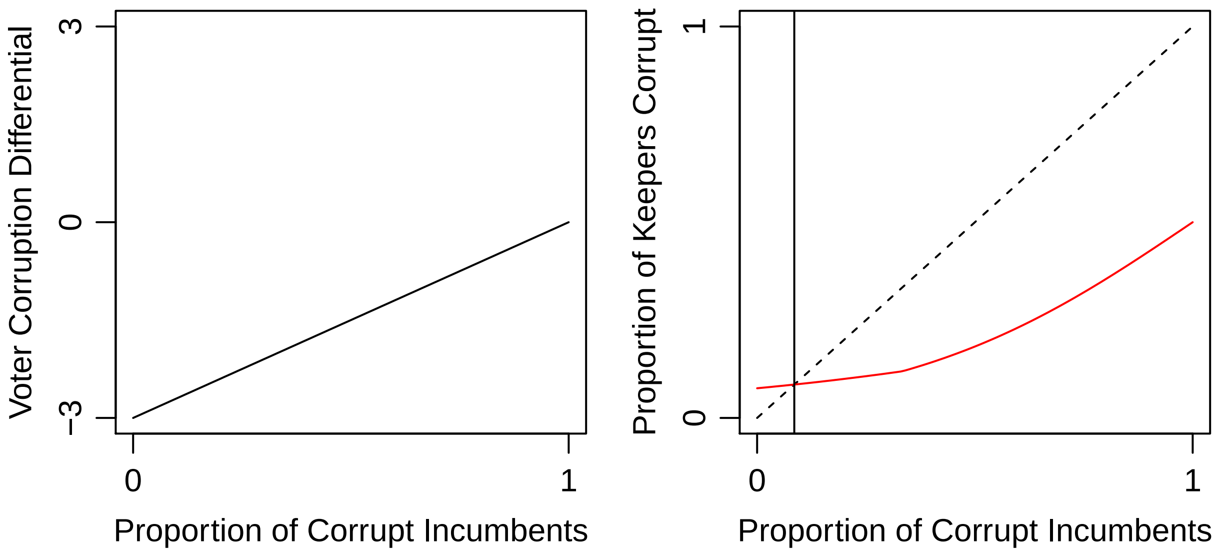
In the main text we mainly focus on the case where d^v is linear, and hence can be written as: $d^v(\rho) = b_0 + b_1\rho$. In this formulation, b_0 can be interpreted as the voter corruption differential when no others are corrupt, and b_1 is how the relative value of having a corrupt politician changes as ρ increases. That is, if $b_1 < 0$, the relative value of having a corrupt politician becomes worse as there are more corrupt politicians, while if $b_1 > 0$ the relative value of having a corrupt representative increases as there is more corruption. As proposition 4 states, there will always be a unique equilibrium if $b_1 \leq 0$, and hence we only include illustrations where $b_1 > 0$.

In the first illustration (Figure 2), $b_0 = -3$ and $b_1 = 3$, which implies that it is always bad to have a corrupt politician, but the degree to which it is bad diminishes as there are additional corrupt politicians. In this case, the proportion of kept politicians that are corrupt is increasing in the proportion of others that are corrupt, but the slope of this curve is never steep enough for there to be multiple intersections. So, there is a unique equilibrium with a low level of corruption.

Next, consider a case where $b_0 = -3$ and $b_1 = 6$ (Figure 3). This implies that citizens prefer a non-corrupt representative when less than half of the other politicians are corrupt, but prefer a corrupt representative when more than half of others are corrupt. Now the slope of the right-hand side of the equilibrium condition is steep enough that there is an

[25] That is, we do not specify a primitive g function and solve the politician model, as it is the resulting d^v that matters for voter behavior.

Figure 2: Illustration of voter model with one stable proportion of corrupt incumbents



equilibrium with a very low proportion of corrupt politicians (and corrupt politicians are generally voted out of office) and one with a very high proportion of corrupt politicians (and *non-corrupt* politicians are generally voted out of office).[26]

To generalize beyond these examples, Figure 4 illustrates when there are multiple equilibria as a function of b_0 and b_1 for linear corruption effects. The parameters with multiple equilibria are represented by the shaded area. For this set of parameters, there are multiple equilibria when (1) the slope b_1 is highly positive, and (2) the intercept is negative and around half the slope. This implies that the relative value of having a corrupt representative must be quickly increasing in how many others are corrupt, and the effect of having a corrupt politician must be close to zero when half of politicians are corrupt. Where this intercept lies depends on the prior probability of corruption q, but the general fact that the value of

[26]The "middle" equilibrium is unstable, in the sense that a slight increase in the proportion of corrupt politicians would lead there to be even more corruption. On the other hand, in the stable equilibrium, an increase in the proportion of corrupt politicians to $\rho^* + \epsilon$ would lead to a proportion of re-elected corrupt politicians that is less than $\rho^* + \epsilon$.

Figure 3: Illustration of voter model with multiple equilibria

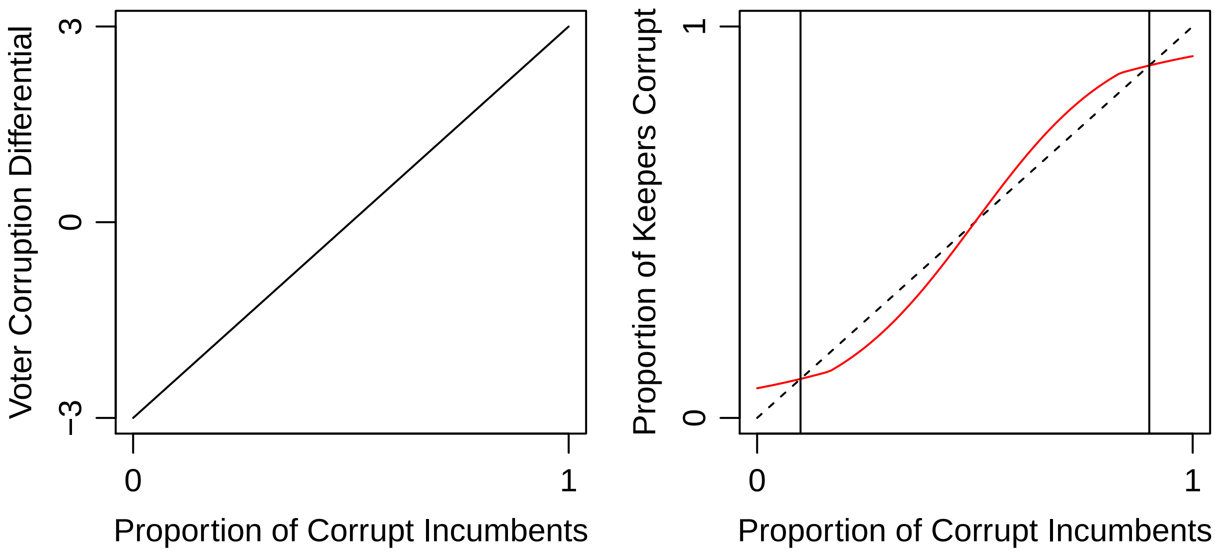
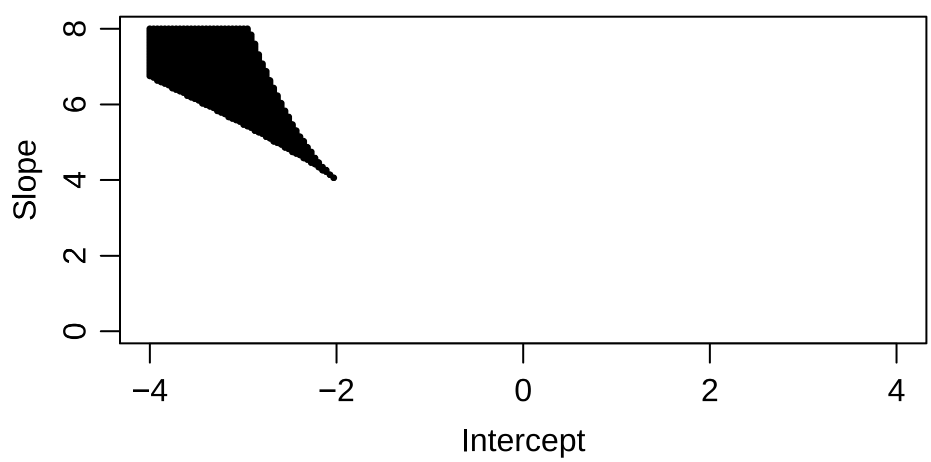


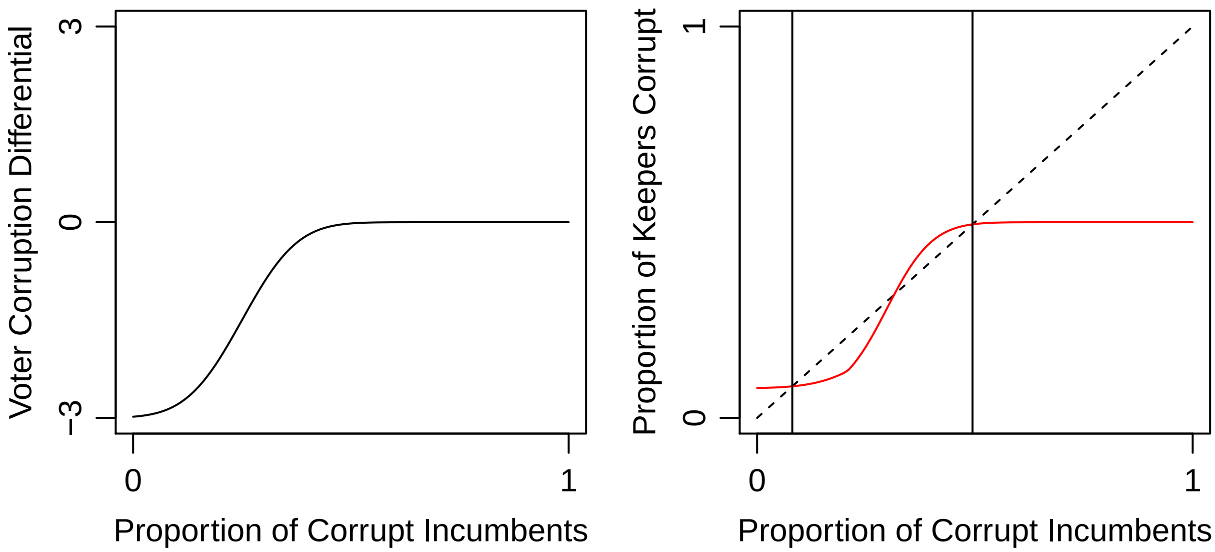
Figure 4: Range of linear voter corruption differentials that result in multiple equilibria



having a corrupt politician must sometimes be positive and sometimes negative always holds in the linear case.

Finally, to demonstrate that it is not necessary for voters to prefer a corrupt representative

Figure 5: Illustration with multiple equilibria when voters always prefer a clean politician



in order to get a political corruption trap, we present one example with a nonlinear voter corruption differential. As illustrated by the right panel of Figure 5, in this illustration the penalty for having a corrupt politician drops (i.e., $d^v(\rho)$ is negative but increasing) for low levels of corruption, but once corruption is common voters preference for a clean politician becomes very small. The right panel demonstrates that this leads to multiple equilibria, one where there are very few corrupt politicians and one where about half of politicians are corrupt and voters make little distinction on this dimension. See the appendix for further discussion of when non-linear voter corruption differentials lead to a corruption trap.

Summarizing:

Proposition 4. *There is at least one equilibrium to the voter model, and if* $\frac{\partial d^v(\rho)}{\partial \rho} \leq 0$ *the equilibrium is unique. If* $\frac{\partial d^v(\rho)}{\partial \rho} > 0$ *there may be multiple equilibria to the voter model.*

The possibility of multiple equilibria in the voter model provides an explanation of corruption traps that is inherently political. When there are equilibria with both a low and high level of corruption and the higher corruption equilibrium is played, the accountability mechanism does not necessarily solve problems of corruption and can perpetuate political corruption traps: voters are either accepting of corrupt politicians or avoid clean politicians under the mutual expectation that voters in other districts will do the same. That is, the accountability mechanism requires that citizens are willing to take costly actions to remove "bad" politicians from office. In the case of corruption, if voters in district i are unwilling to vote corrupt politicians from office, voters in district j may not be willing to do so either. In the case where it is better to have a corrupt representative when others are corrupt as well (i.e., d^v is increasing in ρ), voters may even "hold accountable" *clean* politicians who avoid corruption.

3.3 Voter Behavior with Endogenous Entry

Similar dynamics can lead to multiplicity of equilibria due to strategic complementarities among potential entrants to the political arena. The Appendix contains a model where in each district a number of potential entrants decide whether or not to run for office, with the expectation that the winners will play the game described in section 3.2. Since the general logic of this model is similar to the voter model (and the results are of a similar flavor as Caselli and Morelli 2004), we leave the formalization in the Supplementary Appendix.

The two key results from endogenizing entry decisions are: (1) the proportion of entrants who are pre-disposed to corruption is increasing in the expectation of how many others will be corrupt, and (2) this complementarity can lead to a multiplicity of equilibria for similar reasons to the voter model. Therefore, corruption traps can arise individually at the level of incumbents, voters, and entrants.

Our final formal analysis demonstrates how even if the complementarities in engaging in and tolerating corruption at the individual levels do not lead to a political corruption trap, the *interaction between the three* can lead to multiple equilibria and hence a high corruption trap.

To do so, we now let the proportion of entrants that are corrupt be a function of ρ (as is predicted in proposition 1 in the Supplemental Appendix). That is, we now write $q(\rho)$, and assume q is increasing, implying that more challengers are corrupt when more incumbents are corrupt.

With this modification, the probability of entering and being retained for a corrupt and non-corrupt politician is:

$$q_0^k(\rho) = (1 - q(\rho))(1 - F(\mathbb{E}[a] + q(\rho)d^v(\rho)))$$
$$q_1^k(\rho) = q(\rho)(1 - F(\mathbb{E}[a] - (1 - q(\rho))d^v(\rho)))$$

Differentiating with respect to ρ gives:

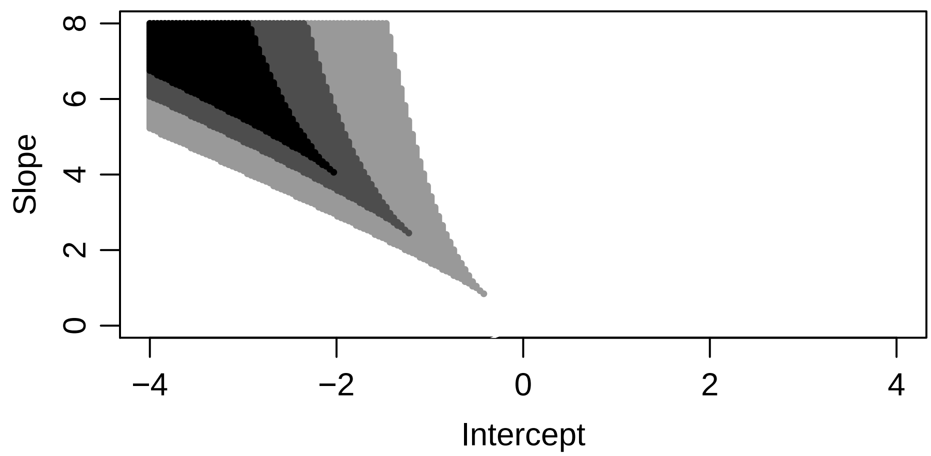
$$\frac{\partial q_0^k}{\partial \rho} = -(1 - q(\rho))(q(\rho)g'(\rho) + p_c'(\rho)d^v(\rho))f(\mathbb{E}[a] + q(\rho)d^v(\rho)) - (1 - F(\mathbb{E}[a] + q(\rho)d^v(\rho)))q'(\rho)$$
$$\frac{\partial q_1^k}{\partial \rho} = q(\rho)(1 - q(\rho))g'(\rho) - q(\rho)d^v(\rho))f(\mathbb{E}[a] - (1 - q(\rho))d^v(\rho)) + (1 - F(\mathbb{E}[a] - (1 - q(\rho))d^v(\rho)))q'(\rho)$$

If $q' = 0$, these reduce to the derivatives derived above. While it is difficult to isolate the impact of non-constant entry as q' will also change q globally, increasing q' at any fixed ρ will decrease $\frac{\partial q_0^k}{\partial \rho}$ and increase $\frac{\partial q_1^k}{\partial \rho}$. Referring back to equation 6, this will tend to make the slope of $\overline{q}$ steeper, making it easier to have multiple equilibria.

In particular, Figure 6 looks at when there are multiple intersections with a linear q for three cases: first, for $q = .5$ (as before), then for $q = .3 + .4\rho$, and finally for $q = .1 + .8\rho$. Note that for all three of these entry functions, $q(.5) = .5$, however for the latter two the proportion of corrupt incumbents has an increasingly large effect on the proportion of corrupt entrants.

As the proportion of corrupt entrants becomes more heavily dependent on the number

Figure 6: Range of linear corruption effects with multiple equilibria when entry is a function of ρ



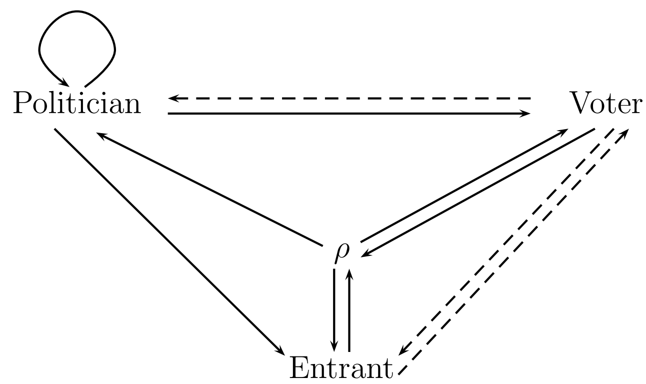
The black area corresponds to a case where p_c is constant in ρ, and the dark and light grey areas correspond to the cases when p_c is slightly increasing in ρ and sharply increasing in ρ, respectively.

of corrupt incumbents (lighter shaded areas in Figure 6), the range of voter corruption differential functions with multiple equilibria increase. This captures the idea that when both voters and potential entrants condition their behavior on expectations of how many other politicians are corrupt, it is easier to construct cases where there are equilibria with both very few corrupt politicians and nearly pervasive corruption. That is, there is a wide range of the parameter space where there would be no political corruption trap with constant entry, but allowing for the pool of entrants to vary based on the level of corruption can generate such a trap.

3.4 Summary

The overarching conclusion of the formal models we have presented is that decisions made by incumbent politicians, potential entrants to politics, and voters all interact in a way that can lead to political corruption traps, i.e., the presence of a high-corruption as well as the

Figure 7: Relationships between the three levels of analysis



ρ is the proportion of politicians predisposed towards corruption.

low-corruption equilibrium.

Figure 7 summarizes these interactions. For each of the three levels, we can think of the actors in question inter-dependently behaving in a way that may lead to more corruption: (1) politicians explicitly choosing to engage in more corrupt behavior, (2) voters being more tolerant of or even approving of corrupt politicians, and (3) those predisposed to corruption being more apt to enter politics. Solid arrows indicate relationships which we have formalized in the previous sections or the Appendix. The dashed arrows represent the relationships we informally discuss below.

In our models, these three levels primarily interact through the channel of what proportion of politicians are predisposed towards corruption (ρ). As ρ increases, politicians choose a higher level of corruption (Proposition 1; solid arrow from ρ to politician), voters can be more accepting of corruption (Equation 7; solid arrow from ρ to voter), and more corrupt potential entrants are relatively more likely to enter than clean potential entrants (Proposition 1; solid arrow from ρ to entrant). The solid arrows from entrant/voter to ρ follow from the definition of ρ in these models: more corrupt entrants and more voter tolerance toward corruption directly lead to more corrupt politicians.

The solid arrows from politician to entrant and politician to voter indicate that even for a fixed proportion of incumbents predisposed to corruption, more corrupt behavior (i.e. greater x^*) has analogous effects on voters and entrants. That is, while the voter and entrant models primarily focused on the proportion of corrupt incumbents (ρ), what really matters is the aggregate level of corruption (ρx^*). So, we could equivalently write the voter and politician corruption differentials as $d^p(\rho x^*)$, and the same results that formalize the ρ to entrant and ρ to voter lines imply the lines from politician to entrant and politician to voter. Finally, the insights from Section 3.3 show the combined effect of ρ on the behavior of entrants (solid arrow from ρ and entrant) and voters (solid arrow from ρ to voter).

There are other potential sources of positive feedback among these levels of analysis not modeled here. For example, if voters are less apt to punish corrupt behavior, then politicians will be more free to engage in corruption (dashed line from voter to politician). Further, those predisposed to corruption are more apt to run if behaving in a corrupt manner is less likely to lead to voter punishment (dashed line from voter to entrant). Finally, if voters think all entrants will be corrupt, there is no incentive to vote out corrupt incumbents for this reason (dashed line from entrant to voter).

We can also relate our argument to the existing literature on multiple corruption and accountability equilibria in the context of Figure 7. Most of the existing formal work on multiple corruption equilibria looks at the circular arrow from politician to politician (also formalized here in Proposition 1; Andvig and Moene 1990; Cadot 1987; Lui 1986; Mauro 2004).[27] The main model in Caselli and Morelli (2004) presents a somewhat different argument for the arrows between entrant and ρ, where candidates of higher quality are more likely to run for office when other high-quality candidates run; they also argue that politicians

[27] Formally, even for a fixed ρ the incentives to commit fraud are increasing in how much fraud others commit, i.e., x^*.

make policy decisions that make candidates more like them (good or bad) run in the future, capturing an alternative channel from politician to entrant. Ashworth, Bueno de Mesquita, and Friedenberg (2014) and Svolik (2013) characterize differently the arrows between politician and voter, arguing that voters' pessimistic expectations can fail to incentivize politicians who could otherwise be induced to perform well. Svolik (2013) also models differently the arrows between voter and entrant by showing that bad entrants are discouraged from running when voters pay the cost of monitoring politicians and potential entrants' outside options are good.

In sum, the findings from our models, summarized in Figure 7, suggest that there are strong strategic complementarities in many of the decisions among voters, entrants and incumbents that determine the level of corruption in a polity. This raises the problematic possibility that interactions between these three sets of actors create even more opportunities for a political corruption trap. We conclude by informally discussing why we believe such political corruption traps may be particularly hard to escape.

4 Discussion

The most common implication of the existing account of strategically-induced corruption traps is that a way to escape to a lower corruption equilibrium is to change the expectation among the fixed group of politicians (or bureaucrats). That is, believing that all others will switch to a lower corruption equilibrium can be a self-fulfilling prophecy. In our models, if all politicians are up for re-election at once, the type of corruption trap that exists when there are multiple equilibria can also be interpreted as solvable by coordinating expectations as in the canonical literature. That is, if all the voters anticipate that all other voters – including those in other constituencies – will switch to using the selection rules associated with the low corruption equilibrium rule and are patient enough to wait until the corruption level reaches its lower steady state, it will be optimal to do so. If endogenous entry is also considered, then potential entrants to politics would need to coordinate their expectations as well.

We think this is unrealistic for three reasons. First, it may require voters to be very patient, as it may take multiple replacements to find a representative that is both competent and non-corrupt. The results of Svolik (2013) and Meirowitz and Tucker (2013) suggest that such patience may be unrealistic, since repeated poor performance by incumbents may lead the voters to abandon the monitoring of politicians, discourage them from protesting against the regime, and even question the merits of democratic rule.

Second, given the importance of the prevalence of corruption among incumbents for the possibility of a corruption trap in our models, coordinating the expectations of all actors and voter patience may not be enough, unless the entire political elite in power is replaced. While our equilibrium definition – as with any standard solution concept – assumes that such coordination of expectations occurs, inducing rapid shifts in expectations across a wide variety of actors is likely to be very difficult. As Golden (2010) notes:

> the implosion of the Italian party system on the heels of the Clean Hands investigations [in the early 1990s] is the only known historical instance in any democratic nation of the electoral repudiation by voters of an entire corrupt national political elite (pp. 80-81).

This suggests that the persistence of corruption may in part be a consequence of political and electoral institutions that influence how many politicians are up for reelection at any point in time. While rigorously formalizing this in our models proves difficult,[28] staggered election timing and the many incumbency advantages may make it even more difficult to

[28] The main challenge is that our solution concepts assume a long-run distribution where all politicians are up for election at the same time. Having staggered elections requires voters (and politicians and entrants) to make more complex forward-looking choices, which we leave to future research.

replace politicians *en masse*. Therefore, our results suggest that exploring the effects of electoral and political institutions on the *persistence* of corruption may be a valuable way to learn more about how countries can escape these traps. Existing studies of the institutional antecedents of corruption primarily focus on explaining the effects on the *level* of corruption, rather than its persistence over time (e.g. Lambsdorff, 2006; Pellegrini and Gerlagh, 2008). This is potentially a valuable area for future research.

References

Abbink, Klaus, and Danila Serra. 2012. "Anti-Corruption Policies: Lessons from the Lab." In *New Advances in Experimental Research on Corruption*, ed. Danila Serra and Leonard Wantchekon. Vol. Research In Experimental Economics Volume 15. Emerald Group Publishing.

Alt, James E., and David Dreyer Lassen. 2003. "The Political Economy of Institutions and Corruption in American States." *Journal of Theoretical Politics* 15 (3): 341–365.

Andvig, Jens Chr., and Karl Ove Moene. 1990. "How Corruption may Corrupt." *Journal of Economic Behavior & Organization* 13 (1): 63–76.

Ashworth, Scott, Ethan Bueno de Mesquita, and Amanda Friedenberg. 2014. "Accountability Traps." Manuscript.

Banerjee, Abhijit, and Rohini Pande. 2007. "Parochial Politics: Ethnic Preferences and Politician Corruption." Manuscript.

Banfield, Edward C., and James Q. Wilson. 1963. *City politics*. Vol. 335 Harvard University Press Cambridge, MA.

Banfield, Edward C., and Laura Fasano Banfield. 1958. *The Moral Basis of a Backward Society*. Free Press New York.

Barberá, Pablo, Pablo Fernández-Vázquez, and Gonzalo Rivero. Forthcoming. "Rooting Out Corruption or Rooting For Corruption? The Electoral Consequences of Corruption Scandals in Spain, 2003-2011." *Political Science Research and Methods*.

Becker, Sascha, Katrin Boeckh, Christa Hainz, and Ludger Woessmann. Forthcoming. "The Empire is Dead, Long Live the Empire! Long-run Persistence of Trust and Corruption in the Bureaucracy." *The Economic Journal*.

Cadot, Olivier. 1987. "Corruption as a Gamble." *Journal of Public Economics* 33 (2): 223–244.

Caselli, Francesco, and Massimo Morelli. 2004. "Bad Politicians." *Journal of Public Economics* 88 (3–4): 759 - 782.

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

Chong, Alberto, Ana L. de la O, Dean Karlan, and Leonard Wantchekon. 2015. "Does Corruption Information Inspire the Fight or Quash the Hope? A Field Experiment in Mexico on Voter Turnout, Choice and Party Identification." *Journal of Politics* 77 (1): 55-71.

Corbacho, Ana, Daniel W. Gingerich, Virginia Oliveros, and Mauricio Ruiz-Vega. Forthcoming. "Corruption as a Self-Fulfilling Prophecy: Evidence from a Survey Experiment in Costa Rica." *American Journal of Political Science*.

Coviello, Decio, and Stefano Gagliarducci. 2014. "Tenure in Office and Public Procurement." Manuscript.

Damania, Richard, Per G. Fredriksson, and Muthukumara Mani. 2004. “The Persistence of Corruption and Regulatory Compliance Failures: Theory and Evidence.” *Public Choice* 121 (3-4): 363-390.

de Figueiredo, Miguel F. P., F. Daniel Hidalgo, and Yuri Kasahara. 2012. “When do Voters Punish Corrupt Politicians? Experimental Evidence from Brazil.” *Manuscript* (May).

de la O, Ana L. 2015. “How Governmental Corruption Breeds Clientelism.” In *Mexico's Evolving Democracy: A Comparative Study of the 2012 Election*, ed. Jorge I. Dominguez, Kenneth F. Greene, Chappell H. Lawson, and Alejandro Moreno. Baltimore, MD: Johns Hopkins University Press.

Ferraz, Claudio, and Frederico Finan. 2011. “Electoral Accountability and Corruption: Evidence from the Audits of Local Governments.” *The American Economic Review* 101 (4): 1274.

Ferraz, Claudio, and Frederico S. Finan. 2008. “Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes.” *Quarterly Journal of Economics* 123 (2): 703-745.

Fisman, Raymond, and Edward Miguel. 2007. “Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets.” *Journal of Political Economy* 115 (6): 1020–1048.

Fisman, Raymond, Florian Schulz, and Vikrant Vig. 2014. “The Private Returns to Public Office.” *Journal of Political Economy* 122 (4): 806–862.

Gingerich, Daniel W. 2014. “Brokered Politics in Brazil: An Empirical Analysis.” *Quarterly Journal of Political Science* 9 (3): 269–300.

Glaeser, Edward L, and Raven E Saks. 2006. "Corruption in America." *Journal of Public Economics* 90 (6): 1053–1072.

Golden, Miriam A. 2010. "Puzzles of Political Corruption in Modern Advanced Democracies." In *Democracy and Accountability: Globalized Political Responsibility*, ed. Hideko Magara. Tokyo: Fukosha.

Golden, Miriam A, and Lucio Picci. 2005. "Proposal for a New Measure of Corruption, Illustrated with Italian Data." *Economics & Politics* 17 (1): 37–75.

Golden, Miriam A., and Lucio Picci. 2015. "Incumbency Effects under Proportional Representation: Leaders and Backbenchers in the Postwar Italian Chamber of Deputies." *Legislative Studies Quarterly* 40 (4): 509–538.

Hauk, Esther, and Maria Saez-Marti. 2002. "On the Cultural Transmission of Corruption." *Journal of Economic Theory* 107 (2): 311–335.

Hirano, Shigeo, and James M. Snyder. 2012. "What Happens to Incumbents in Scandals?" *Quarterly Journal of Political Science* 7: 447–456.

Hollyer, James R., and Leonard Wantchekon. 2015. "Corruption and Ideology in Autocracies." *Journal of Law, Economics and Organization* 31 (3): 499–533.

Jackson, Kenneth T. 1985. *Crabgrass Frontier: The Suburbanization of the United States.* New York, NY: Oxford University Press.

Klašnja, Marko. 2015. "Corruption and the Incumbency Disadvantage: Theory and Evidence." *The Journal of Politics* 77 (4): 928–942.

Klašnja, Marko, and Joshua A. Tucker. 2013. "The Economy, Corruption, and the Vote: Evidence from Experiments in Sweden and Moldova." *Electoral Studies* 32: 536–543.

Lambsdorff, Johann Graf. 2006. "Consequences and Causes of Corruption: What Do We Know from a Cross-Section of Countries?" In *International Handbook on the Economics of Corruption*, ed. Susan Rose-Ackerman. Northampton, M.A.: Edward Elgar.

Lui, Francis T. 1986. "A Dynamic Model of Corruption Deterrence." *Journal of Public Economics* 31: 1-22.

Mainwaring, Scott P. 1999. *Rethinking Party Systems in the Third Wave of Democratization: the Case of Brazil*. Stanford University Press.

Manzetti, Luigi, and Carole J. Wilson. 2007. "Why do Corrupt Governments Maintain Public Support?" *Comparative Political Studies* 40 (8): 949–970.

Mauro, Paolo. 1995. "Corruption and Growth." *The Quarterly Journal of Economics*: 681–712.

Mauro, Paolo. 2004. "The Persistence of Corruption and Slow Economic Growth." *IMF staff papers*: 1–18.

Meirowitz, Adam, and Joshua A. Tucker. 2013. "People Power or a One-Shot Deal? A Dynamic Model of Protest." *American Journal of Political Science* 57 (2): 478–490.

Menes, Rebecca. 2006. "Limiting the Reach of the Grabbing Hand: Graft and Growth in American Cities, 1880 to 1930." In *Corruption and Reform: Lessons from America's Economic History*, ed. Edward L. Glaeser and Claudia Goldin. Chicago, IL: University of Chicago Press.

Mishra, Ajit. 2005. "Persistence of Corruption: Some Theoretical Perspectives." *World Development* 34: 349-358.

Montinola, Gabriella A., and Robert W. Jackman. 2002. "Sources of Corruption: A Cross-Country Study." *British Journal of Political Science* 32: 147-170.

Pellegrini, Lorenzo, and Reyer Gerlagh. 2008. "Causes of Corruption: a Survey of Cross-National Analyses and Extended Results." *Economics of Governance* 9: 245-263.

Peters, John G., and Susan Welch. 1980. "The Effects of Charges of Corruption on Voting Behavior in Congressional Elections." *The American Political Science Review* 74 (3): 697-708.

Pritchett, Lant, Michael Woolcock, and Matthew Andrews. 2010. "Capability Traps? The Mechanisms of Persistent Implementation Failure." *Center for Global Development Working Paper* (234).

Quah, Jon S. T. 1994. "Controlling Corruption in City-States: A Comparative Study of Hong Kong and Singapore." *Crime, Law and Social Change* 22 (4): 391–414.

Reed, Steven A. 1999. "Punishing Corruption: The Response of the Japanese Electorate to Scandals." In *Political Psychology in Japan: Behind the Nails which Sometimes Stick Out (and Get Hammered Down)*, ed. Ofer Feldman. Commack, NY: Nova Science.

Reinikka, Ritva, and Jakob Svensson. 2004. "Local Capture: Evidence from a Central Government Transfer Program in Uganda." *Quarterly Journal of Economics* 119 (2): 679–705.

Shleifer, Andrei, and Robert W Vishny. 1993. "Corruption." *The Quarterly Journal of Economics* 108 (3): 599–617.

Simpser, Alberto. 2013. "The Inter-generational Persistence of Attitudes toward Corruption." Manuscript.

Singer, Matthew M. 2009. "Buying Votes with Dirty Money: The Relationship between Clientelism and Corruption." Paper presented at the 2009 Annual Meetings of the American Political Science Association, September 3-6, Toronto.

Sun, Yan. 2004. *Corruption and Market in Contemporary China*. Cornell University Press.

Svolik, Milan. 2013. "Learning to Love Democracy: Electoral Accountability, Government Performance, and the Consolidation of Democracy." *American Journal of Political Science* 57 (3): 685–702.

Tirole, Jean. 1996. "A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality)." *The Review of Economic Studies* 63 (1): 1–22.

Transparency International. 2004. "Global Corruption Report 2004." London, UK: Pluto Press.

Transparency International India. 2005. "India Corruption Study." Transparency International India and Center for Media Studies.

Treisman, Daniel. 2000. "The Causes of Corruption: a Cross-National Study." *Journal of Public Economics* 76 (3): 399–457.

Treisman, Daniel. 2007. "What Have We Learned about the Causes of Corruption from Ten Years of Cross-National Empirical Research?" *Annual Review of Political Science* 10: 211–244.

Welch, Susan, and John R. Hibbing. 1997. "The Effects of Charges of Corruption on Voting Behavior in Congressional Elections, 1982-1990." *The Journal of Politics* 59 (1): 226-239.

Zechmeister, Elizabeth J., and Daniel Zizumbo-Colunga. 2013. "The Varying Political Toll of Concerns about Corruption in Good Versus Bad Economic Times." *Comparative Political Studies* 46 (10): 1190–1218.