1 Introduction

Corruption is rampant in many poor countries. As such, anti-corruption policies continue to be a central component of development strategies. For example, since 1996, the World Bank alone has supported over 600 anti-corruption programs.

Unfortunately, this is one area where research has lagged policy. Research on corruption faces two important obstacles— one empirical and one theoretical. On the empirical side, the primary challenge is measurement. Corruption, by its very nature, is illicit and secretive. How does one study something that is defined in part by the fact that individuals go to great lengths to hide it? What does one do to deal with the fact that attempts to measure corruption may cause the actors involved to either reduce their illicit behaviors during the periods of measurement, or find new ways to obscure their behaviors? If we cannot accurately measure corruption, how can we test between different theories, measure its impacts, or even produce suggestive correlations?

In recent years, some progress has been made along this dimension. In particular, while the previous generation of corruption measures were mainly based on the perception of corruption by participants (with various assorted problems in interpreting these measures), the current generation of studies have focused on collecting and reporting objective information, obtained either from direct measurement or from other information.

The theoretical challenge comes in part from the need to go beyond thinking corruption as more than a generic form of moral hazard in organizations to the point where we can map different manifestations of corruption to different underlying environments, where the word environment is interpreted to cover both usual focus of the corruption literature—the nature of the monitoring and the punishments as all as the intrinsic motivation of the bureaucrats (for example how corruption fits into their moral compass)—but also what is less emphasized, the nature of the particular economic decision that the bureaucrats are participating in. This is important for two reasons: First, from the point of view of empirical research, differences in the nature of corruption in different economic settings is an important source of testable predictions. Second, from the point
of policy design, it is vital that we are able to think of how changing the environment might be an effective substitute for simply adjusting the punishment (which may or not be feasible).

An example might make the second point clearer: Bandiera, Prat and Tommaso (2009) study waste in government procurement in Italy, a country that often rated as one of the most corrupt in Europe. Using detailed data, they show that different branches of government pay very different prices for the exact same product (down to the brand and color). These differences can be as large as fifty percent or more. In fact, they estimate that the government could save up to 2 percent of GDP if most purchase officers paid the same price that was obtained by the most frugal officers.

They also show, however, that the price differences are a function of where the purchase officers buy. They can either get their supplies from the market or from an approved supplier, Consip. Consip charges a publicly announced price which leaves no scope for kick-backs. Going to the market, by contrast, potentially allows the buyer to negotiate his own deal, which might include something extra for him. If this were true the least corrupt officers, those paying the lowest prices on the market, would use Consip. In fact, the data suggests the opposite. When a new item gets added to Consip’s list of available items, the bureaucrats paying the highest prices turn to Consip. Moreover these purchasing officers were also the ones that were, by all accounts, the best monitored—the centralized bureaucracies rather than the more autonomous hospitals and universities. This suggests a different narrative. These officers pay much higher prices than others not only because of kickbacks (though that is there too). Instead the issue is of justifiability. Buying from the official supplier requires no justification—and no effort. Bandiera, Prat and Tommaso (2009) argue that a major source of the waste here is the fear of being prosecuted for corruption. Bureaucrats pay high prices to avoid any taint of corruption. Notice that under this logic, changing the bureaucratic rules to give the bureaucrat a fixed procurement budget but full discretion—so that he can even pocket any money he saves, may generate both less waste and less corruption—he might be pocketing more money now, but that is perfectly legal, and being free to keep the money may give him a strong reason to avoid waste.

This does not imply that full discretion is always a good idea. Think of the allocation of hospital beds. If need is not related to ability to pay, giving a bureaucrat full discretion about how to allocate beds may lead to a large proportion of them going to those who do not really need them. Making stringent rules about how the bureaucrat is supposed to allocate the beds will generate corruption as the greediest bureaucrats bend the rules to make more money, but also, potentially, a better allocation, since the more honest bureaucrats will stick to the rules.

The more general point is that corruption is the result of the task that the bureaucrat is assigned to carry out. We can usually get rid of it by setting the appropriate task (giving discretion), but that is not always a particularly desirable from society’s point of view. However, the optimal response to the possibility of corruption may often be to change the nature of the task. Note
that the change in the task may not always reduce corruption: it might just address the misallocation or degradation of services that corruption often causes, i.e. the fact that the beds were going to the rich or that the wait for a bed was unacceptably long.

Starting from the premise that the corruption that we observe may be the result of the choice of a task that the bureaucrat is assigned to also gives us a way to generate testable implications. In particular, we will be able map the particular problem the government is trying to solve into a vector of outcomes (e.g. bribes, lines, misallocation of beds, etc.). The answers we are looking for are of the form: are the waits likely to be longer when the government is trying to target hospital beds to the very poor rather than the less poor? How about the bribes?

This repositioning of the corruption literature away from a purely crime and punishment approach towards a more "task" focused approach connects it more closely to the literature on the internal economics of organizations that has emerged over the last two decades. This literature explicitly recognizes that most organizations use bureaucratic mechanisms similar to the ones associated with government bureaucrats for many of their internal decisions, which creates scope for corruption (Tirole (1986)). However, there is a lot to be gained from focusing on the specific characteristics of the kinds of settings that governments often deal with. For example, one source of corruption in government is the fact that governments are expected to deliver goods and services to those who cannot pay their full value. This is less of an issue within profit-making organizations. We will however return to the relationship between corruption in government and similar issues in private firms in sub-section 3.5.

This chapter is an attempt to highlight the progress made in the corruption literature over the last decade or so, with a focus on the doors this opens for future research. It aims to be more forward looking than backward looking, less a comprehensive review of corruption research than a guide to where it appears to be headed.\footnote{Summarizing a literature as large and multi-disciplinary as corruption poses unique challenges. In this chapter, we have erred on the side of being forward-looking, trying to paint a picture of where (we feel) this literature is headed. Though we have aimed to cover all the important existing literature, some gaps are an unfortunate necessity in order to keep an overview within a manageable length. Our apologies to authors whose work we could not cover in as much detail.} It provides a theoretical framework to illustrate the tasks approach and an overview of the tools that are now available for empirically analyzing corruption. It then lays out the open questions that we feel are both interesting and within reach.

We start with a discussion of what we mean by corruption. The key point is that corruption involves breaking rules, not just doing something that is unethical or against the collective interest. This will lead us naturally to think of the task that the bureaucrat has been assigned (which includes the rules). This is the subject of the next section, where we develop a simple theoretical framework for thinking about corruption and its many manifestations. We then discuss strategies for measuring corruption. The penultimate section describes
some attempts to test theories of corruption using these measures. We conclude
with a discussion of the main areas where we feel are important for future research.

2 Defining Corruption

We define corruption as the breaking of a rule by a bureaucrat (or an elected
official) for private gain. This definition includes the most obvious type of
corruption—a bureaucrat taking an overt monetary bribe in order to bend a
rule, thereby providing a service to someone that he was not supposed to. How-
ever, it would also encompass more nuanced forms of bureaucratic corruption.
For example, it would include nepotism, such as if a bureaucrat provided a gov-
ernment contract to a firm owned by his or her nephew rather than to a firm
that ought to win a competitive, open procurement process. This definition
would also cover the case of a bureaucrat who “steals time”: he or she may, for
example, not show up to work, but still collect his or her paycheck.\[2\]

Under this definition, the rules define what is corrupt. As a result, the
same act can be classified as corruption in one setting, but not in another
one. For example, in many countries—the United States, India, etc.—a citizen
can obtain passport services more quickly if they pay a fee. While this act
would not be considered corruption in these countries, it would be in others
where no such provision in the law exists. On the other hand, many important
political economy issues may not necessarily be considered corruption under this
definition. For example, a government official providing patronage to supporters
may have important ethical and allocative implications, but this act would not
necessarily be corruption if no formal rule is technically broken.\[3\]

While the definition of corruption used in this paper is similar to those used
by others in the literature, there are important distinctions. For example, our
definition is quite similar to the definition discussed by Svensson (2005)—“the
misuse of public office for private gain”—and also to Shleifer and Vishny (2001),
who define corruption “as the sale by government officials of government prop-
erty for personal gain.” All three definitions imply that the official gains per-
sonally from their particular position. Moreover, as Shleifer and Vishny (2001)
define property quite loosely as including both physical assets (e.g., “land”) and
assets that have an option value (e.g., “a business license”), their definition
encompasses many of the same acts of corruption discussed in this paper and
in Svensson (2005). However, there are slight differences in what qualifies as
corruption across the definitions. For example, suppose we assume that a gov-
ernment official has the final say over whom to allocate a government contract
to. They may choose to sell it to their nephew, and gain great personal hap-
piness from doing so. Thus, this may be considered corruption under Svensson

\[2\] Quite often, we see the same forms of corruption in the nonprofit sector, where a social
good is being provided, and the private and social value may not necessarily coincide. The
models presented in this paper would naturally extend to the non-profit sector.

\[3\] To see a deeper discussion of political corruption, see Pande (2007).
(2005) and Shleifer and Vishny (2001). However, if the official has the final say, and has not broken any official rules, this would not be considered corruption under our definition, despite being morally questionable.

We have chosen to use this definition for a combination of pragmatic and conceptual reasons. Pragmatically, the emphasis on breaking formal rules (as opposed to moral or ethical ones) sidesteps the need to make subjective ethical judgments and thereby avoids the need to have a deeper discussion of cultural differences. The emphasis on all kinds of gain rather than just money, sidesteps a measurement problem: bribes by their very nature are hard to measure, whereas rule breaking is easier to measure. Conceptually, these distinctions are also in line with the framework we describe below.

3 A Formal Framework for Understanding Corruption

The challenge of modeling corruption comes from the very definition of corruption. Corruption, as we say above, is when the bureaucrat (or elected official) breaks a rule for private gain. This immediately raise some questions since the rules themselves are chosen by the government. Specifically, why have these rules in place which we know are going to be violated? Why not change the rules so that there is no incentive to violate them? This leads to an ancillary question: can you change the rules costlessly and eliminate corruption without affecting anything else that you care about?

To understand these issues, we begin by thinking about the underlying task. Our model of tasks is simple, and yet it captures many of the tasks bureaucrats (and also the private sector) typically carry out. We focus on an assignment problem. A bureaucrat must assign a limited number of slots to applicants. The applicants differ in their social valuation of a slot, their private valuation of it, and also their capacity to pay for it. This simple set-up captures many important cases. Consider a profit-maximizing firm selling a good. In this case, the slot is the good and private and social values coincide perfectly. Next, consider the case of a credit officer at a government bank assigning loans. Here, the private ability to pay may be the lowest precisely among those who have the highest social returns from the loan. This potential for divergence between private and social returns is not incidental—it may be the reason why the government was involved in providing the loans in the first place. However, it is also the reason why there is corruption.

The bureaucrat’s task here goes beyond just allocating the slots: he may also face rules about what prices he can charge for them and whether he can engage in "testing" to determine an agent’s type. The government sets both these rules and the incentives facing the bureaucrat.

While this framework does encompass many of the models of bureaucratic...
misbehavior in the literature, we make no claims of generality for it. We make a large number of modeling choices that are pointed out along the way. These are made mostly in the interests of simplicity and clarity, but we recognize that many of them can also have substantive implications.

3.1 Setup

We will analyze the problem of a government allocating “slots” through a bureaucrat who implements the allocation process. There is a continuum of slots with size 1 that need to be allocated to a population of size $N > 1$. Agents have differing private and social values for slots. Specifically there are two types of agents: $H$ and $L$ with mass $N_H$ and $N_L$ such that $N_H + N_L > 1$. The social value of giving a slot to type $H$ is $H$ and $L$ for type $L$. We assume that $H > L$. Private benefits can be different, with each group valuing their slots at $h$ and $l$. Agents' types are private information and unknown by either the government or the bureaucrat, though the bureaucrat has a technology for learning about type that is called "testing," which we describe below. Agents also differ in their ability to pay for a slot, which we denote by $y_h$ and $y_l$: because of credit constraints agents may not be able to pay full private value so $y_h \leq h$ and $y_l \leq l$.

There is a generic testing technology to detect types that the bureaucrat can use. If used on someone of type $L$ who wants to pass for a period of time $t$, the probability that he will fail the test (i.e. get an outcome $F$) is $\phi_L(t)$, $\phi'_L(t) \geq 0$. The corresponding probability for a type $H$ who wants to pass is 0; they always pass if they want to (i.e. get the outcome $S$). Either type can always opt to deliberately fail. The cost of testing for $t$ hours is $\nu t$ to the bureaucrat. The cost of being tested for the person for $t$ hours is $\delta t$. A simple example of testing would be a driving exam to verify that one can drive.

Testing is going to be the only costly action taken by the bureaucrat in our model. We assume, for one, that the bureaucrat does not put in any effort to give out the slots. We feel that we capture much of what is relevant about bureaucrats shirking through this device, but there are no doubt some nuances we are missing.

3.1.1 Possible Mechanisms

The basic problem for the bureaucrat is the choice of a mechanism. The bureaucrat announces a direct mechanism that he can commit to ex ante. Each mechanism constitutes of a vector $R = (t_x, p_{xr}, \pi_{xr})$, where $t_x$ is the amount of testing for each announced type $x = H, L$, $\pi_{xr}$ is the probability that someone gets a slot conditional on announcing a type $x = H, L$ and getting a result $r = F, S$, and $p_{xr}$ is the price they will pay in the corresponding condition. We will restrict ourselves to winner-pay mechanisms here, mechanisms where you do not pay when you do not receive a slot. For analysis of the more general case

\footnote{We recognize that the actual mechanism used will often be very different from the direct mechanism. We discuss some of the issues this raises in the concluding subsection.}
where you may have to pay a non-refundable "entry fee" to enter the bidding, see, for example, Banerjee (1997).

Since the bureaucrat only chooses direct mechanisms, any $R$ is supposed to satisfy the incentive constraints:

$$\pi_{HS}(h - p_{HS}) - \delta t_H \geq \pi_{LS}(l - p_{LS}) - \delta t_L$$
$$\pi_{LS}(l - p_{LS})(1 - \phi_L(t_L)) + \pi_{LF}(l - p_{LF})\phi_L(t_L) - \delta t_L \geq \pi_{HS}(l - p_{HS})(1 - \phi_H(t_H)) + \pi_{HF}(l - p_{HF})\phi_H(t_H)$$

Moreover, the clients are allowed to walk away. This is captured by the participation constraints:

$$\pi_{HS}(h - p_{HS}) - \delta t_H \geq 0$$
$$\pi_{LS}(l - p_{LS})(1 - \phi_L(t_L)) + \pi_{LF}(l - p_{LF})\phi_L(t_L) - \delta t_L \geq 0.$$ 

There is also a total slot constraint:

$$N_H\pi_{HS} + N_L\pi_{LS}(1 - \phi_L(t_L)) + N_L\pi_{LF}\phi_L(t_L) \leq 1$$

Finally, there is affordability: agents cannot pay more than they have:

$$p_{Hr} \leq y_{H}, r = F, P$$
$$p_{Lr} \leq y_{L}, r = F, P.$$ 

Define $R$ to be the set of values of $R$ that satisfy these conditions.

### 3.1.2 Rules

The government chooses a set of rules for the bureaucrat which take the form $R = (T_x, P_{xr}, \Pi_{xr})$; $T_x$ is the set of permitted values for $t_x$, $P_{xr}$ is the set of permitted prices and $\Pi_{xr}$ is the set of permitted values of $\pi_{xr}$ for $x = H, L$ and $r = F, S$. While we assume that the government does not observe every individual’s type, we do allow $P_{xr}$ and $\Pi_{xr}$ to depend on the buyer’s type. The idea is that if there is a gross misallocation of slots or large-scale bribery by one type, there may be some way for the government to find out (the press might publish a story saying that the hospital beds were all being occupied by those who are getting cosmetic surgery by paying high prices or the government might sample a few people who got the slots). However, we do not assume that being able to observe violations of $P_{xr}$ automatically implies being able to observe violations of $\Pi_{xr}$: it may be easy to find out that some people are being charged higher than permitted prices without learning anything more generally about how the slots are being allocated.

We assume that $R$ is feasible in the sense that there exists at least one $R = (t_x, p_{xr}, \pi_{xr}) \in R$ such that $t_x \in T_x$, $p_{xr} \in P_{xr}$ and $\pi_{xr} \in \Pi_{xr}$. If $R$ is not a singleton, then the bureaucrat has discretion.

The government also chooses $p$, which is the price that the bureaucrat has to pay to the government for each slot he gives out. Assume that this is strictly
enforced so that the price is always paid. This assumption can be relaxed easily, but offers no new insights.

In specific examples, we will make specific assumptions about what the government can contract on which will give structure to \( \mathcal{R} \). For example, if \( t_x \) is not contractable, then the rules will not say anything about it—in other words \( T_x \) will be \([0, \infty] \times [0, \infty]\).

### 3.1.3 The Bureaucrat’s Choice

For each mechanism \( R \in \mathbb{R} \cap \mathcal{R} \) the bureaucrat’s payoff is:

\[
N_H \pi_H S(p_{HS} - p) + N_L \pi_L S(p_{LS} - p)(1 - \phi_L(t_L)) + N_L \pi_L F(p_{LF} - p) \phi_L(t_L) - \nu N_H t_H - \nu N_L t_L.
\]

However if \( R \) is in \( \mathbb{R} \cap \mathcal{R}_c \), we will assume that a bureaucrat pays a cost of breaking the rules, which we will refer to as \( \gamma \). Hence, the bureaucrat’s payoff for any \( R \) in \( \mathbb{R} \cap \mathcal{R}_c \) is:

\[
N_H \pi_H S(p_{HS} - p) + N_L \pi_L S(p_{LS} - p)(1 - \phi_L(t_L)) + N_L \pi_L F(p_{LF} - p) \phi_L(t_L) - \nu N_H t_H - \nu N_L t_L - \gamma \tag{1}
\]

We will assume assume that the cost \( \gamma \) is unknown to the government when setting rules, though it knows that it is drawn from a distribution \( G(\gamma) \). A corruptible bureaucrat is one for which \( \gamma \) is finite.\(^6\) As a result, we will write \( R(\mathcal{R}, \gamma) \) as the mechanism chosen by a bureaucrat with cost of corruption \( \gamma \) when the rule is \( \mathcal{R} \).\(^7\)

### 3.1.4 The Government’s Choice

We assume that the bureaucrat is the agent of what we call the government (but others have called the constitution-maker), a principal whose preference is to maximize the social welfare generated by the allocation of the slots. This is partly an artifact of the way we model things. What is key is that the bureaucrat has a boss whose objectives are different from his and who is in a position to punish him. Otherwise, he would never have to break any rules since he, in effect, makes his own rules. The assumption that his boss cares only about social welfare is convenient but not necessary. Much of what we have to say would go through if the principal cares less about the bureaucrat’s welfare and more about that of the other beneficiaries than the bureaucrat, which may be true even if think of the principal as the standard issue, partly venal, politician. After all, the politician cares about staying in power and making the bureaucrat happy may not be the best way to do so. Of course, it is possible that the bureaucrat is the one who cares the beneficiaries and is trying to protect them from his boss. This is an interesting and not necessarily unimportant possibility that we do not investigate here. More generally, a set up like ours deliberately rules out the more interesting strategic possibilities that

---

\(^6\)This formulation is quite specific in that the cost of violating the rules is independent of the extent of violation.

\(^7\)We assume that when indifferent the bureaucrat chooses what the government wants.
arise in models of political economy, in order to focus on the implementation
issues that arise even without them. The government therefore maximizes:

\[
\max \left[ N_H \pi_H S(R(R, \gamma))H + N_L \pi_L S(R(R, \gamma))(1 - \phi_L(t_L(R(R, \gamma))))L \\
+ N_L \pi_L F_L(t_L(R(R, \gamma)))L - (\nu + \delta)N_H t_H(R(R, \gamma)) - (\nu + \delta)N_L t_L(v) \right] dG(\gamma)
\]

by choosing \( R \) subject to

\[3.1.5 \text{ Interpretation} \]

We intend this to be the simplest model that can illustrate all of the features
we are interested in. Specifically, it allows the bureaucrat to have multiple
dimensions of malfeasance:

- Corruption is when the bureaucrat breaks the rules.
- Bribes-taking is when the bureaucrat charges higher prices than those that
  are mandated.
- Shirk ing is when the bureaucrat fails to implement mandated testing.
- Red-tape is when the bureaucrat implements more than the mandated
  amount of testing.
- Allocative inefficiency is when the wrong people get the slots relative to
  what is in the rules or some slots remain unallocated, when the rules
  require that all of the slots are given out.

The fact that the government does not observe \( \gamma \) is the reason why there is
corruption in equilibrium. If the government knew the particular bureaucrat’s
\( \gamma \), it would know what the bureaucrat would choose given the rules that
are set, and hence be able to set rules that are not broken. However, when \( \gamma \) can take
different values, it has to choose between rules that give the bureaucrat lots of
discretion that will almost never be broken, and rules that are more rigid (to
get the bureaucrat to act in the social interest) and therefore will get broken by
some bureaucrats precisely because they are more stringent.

While simple, the problem goes beyond the standard resource allocation
problem under asymmetric information in two important ways. First, we do not
assume that the private benefit to the person who gets the slot is necessarily
the social benefit. Such a divergence is characteristic of many situations where
the government is involved. For example, society wants to give licenses to good

\[8\text{ It is also worth making clear that the assumption of welfare maximization, while stan-}
\[dard, is quite particular. The government could care, for example, about the distribution of}
\[welfare between the bureaucrats and the potential beneficiaries. In this case the government}
\[may prefer an inefficient outcome because it achieves distributional outcomes better and may}
\[therefore create a more complex set of trade-offs than are permitted here. We will return to}
\[this issue in the concluding subsection.}
drivers ($H > 0$) and not to bad ones ($L < 0$), but the private benefits of getting a license are probably positive for both types. Or suppose the slot is not going to jail. $H$ types are innocent. $L$ types are not. $H > 0$ is the social benefit of not sending an innocent to jail. $L < 0$ is the social benefit of not sending a criminal person to jail. However, the private benefits are positive for both types: $h \leq l$ but $h, l > 0$. Second, we allow the potential beneficiaries to have an ability to pay that is less than their private benefits (or willingness to pay) ($l > y_L$ or $h > y_H$). This is conventionally treated as being equivalent to the beneficiary being credit constrained, but it is worth emphasizing that this covers a range of situations (including the credit-constraint case). For example, consider the person who wants to take his child to the hospital to be treated but his permanent income would not cover the cost. He would however be willing to pay his entire income (less survival needs, say) to save his child’s life and also would be willing to additionally stand in line for four hours a day every morning. In this case, his total willingness to pay (in money and time) is clearly greater than his ability to pay. Clearly if he could freely buy and sell labor, this case would reduce to the standard credit constraint case, but given the many institutional features that govern labor markets, this would be an extreme assumption.\footnote{Another example that exploits a different rigidity in labor markets is the following: there is a woman who is not allowed by her family to work is willing to walk three miles every day to make sure that her child gets an education. Her ability to pay (assume that the rest of the family does not care about education) is clearly less than her willingness to pay.}

On the other hand, the formulation embodies a number of important simplifying assumptions. We impose, for example, that rule-breaking of any type has the same cost, which obviously need not be the case. For example, when a bureaucrat and an agent collude in such a way that the agent is better off than under the official rules, there will probably be less chance of being caught than when the bureaucrat attempts to make the agent worse off. We also do not deal with distributional issues—the government’s preferences are indifferent to who, between the bureaucrat and his various types of clients, gets to keep how much money. We will return to why this may be an important issue later in the paper.

### 3.2 A Useful Typology

Before jumping into the analysis of this model, it is helpful to underline some of the different possibilities that can arise in our framework. The following typology will prove to be particularly handy. The labels of the cases should be self-explanatory, but, in any case, more explanation will emerge as we analyze each case.\footnote{The fact that we have only four cases is an artifact of the assumptions that all $H$ types are identical in their willingness and ability to pay and likewise for $L$ types. However, the basic distinction we are trying to make here is between the case where $H$ types are willing and/or able to pay more and the case where they may not be (captured by $h \leq l$ and $y_H \leq y_L$). The situation where a large fraction of $L$ types are willing and/or able to pay more than a large fraction of $H$ types is qualitatively very similar to the case where all $L$ types are able/willing to pay more than all $H$ types.}
3.2.1 Examples of Case I: Bids Aligned with Value

This is the case where social and private value rank orderings align. While the pure market case, \( H = h = y_H, L = l = y_L \), belongs to this category, it is ultimately broader than that because while the rank ordering may align, the actual ability to pay may not match social value. Some of the other cases that fall into this category include:

1. Choosing efficient contractors for road construction: Type \( H \) are the more efficient contractors. For the same contract, they make more money (\( h > l \)). Since the contractors will be paid for their work, the price that they pay to get the contract can be seen as just a discount on how much they will eventually get paid rather than out of pocket expense. It is plausible therefore that \( y_H = h \) and \( y_L = l \).

2. Allocating licenses to import goods to those who will make the socially optimal use of them: In an otherwise undistorted economy, the private benefits should be the same as the social benefits, like in the road construction case, but in this case there may be credit constraints because you pay first and profit from them later. However, it is plausible that the type \( H' \)'s should be able to raise more money for than the type \( L' \)'s. Thus, \( y_H < h = H > L = l > y_L \) and \( y_L < y_H \).

3.2.2 Examples of Case II: Unwillingness to Pay

This is the least likely of the four cases and so we will not spend much time on it. However, one possible example is a merit goods like subsidized condoms against HIV infection: \( H \) are high risk-types. They like taking risks. Hence \( h < l \). However, they may as well be richer (say because they can afford to buy sex): \( y_H > y_L \).

3.2.3 Examples of Case III: Inability to Pay

In this case, there is alignment of values: the high type values the good more than the low type, but there is an inability to pay.

1. How to allocate hospital beds? The \( H \) types really need the beds (say, rather than those who just need cosmetic surgery). The social valuation probably should be the private valuation in this case: \( H = h > L = l > 0 \). However, there is no reason to assume that the \( H \) types can afford to pay more. We capture this by assuming: \( y_H = y_L = y \).

2. How to allocate subsidized food grains targeted towards the poor? Presumably the \( H \) types are the poor who benefit more from subsidized
food grains and the social benefit is plausibly just the private benefit \((H = h > L = l > 0)\). However, the poor may not be able pay as much for the grains as the non-poor: \(y_H < y_L\).

3. How to allocate government jobs to the best candidates? The private gains from getting the job may be higher for the \(H\) types (because the jobs offer so much more rents than the next best alternative to anyone, and the better candidate may get more out of the job). However, everyone is constrained in how much they can pay for the job up-front \((y_H = y_L = y)\).

3.2.4 Examples of Case IV: Values Misaligned

In this case, there is simple misalignment: those who we would like to give the slot to value it the least.

1. Law enforcement: This is the example we already mentioned where the slot is not going to jail: \(H > 0 > L, y_H = y_L = y, h = l > 0\).

2. Driving Licenses: We discussed the set-up of this example previously. However, this example would fall under Case IV if bad drivers value the license more since they are more likely to be picked up by the police: \(H > 0 > L, y_H = y_L = y, h < l\).

3. Procurement: The government wants to procure a fixed number of widgets and has a fixed budget for it (as in Bandiera, Prat and Tommaso (2009), say). Suppose there are high quality firms and low quality firms. It is socially efficient to procure widgets from the high quality firm, even though these firms have higher costs. In this case the slot is the contract, which needs to be allocated among firms. The gains from getting the contract are obviously higher for the low quality firm, which has lower costs. So \(l > h\). As long as these firms not credit constrained that would also mean that \(y_H = h\) and \(y_L = l\).

3.3 Analyzing the Model

This very simple model nevertheless allows for a very rich variety of possibilities and situations. Here we confine ourselves to some illustrative examples of the kind of incentive issues that can arise within this framework, the corresponding patterns of rules chosen, and the violations of the rules. Specifically, we will focus on a set of special cases of Case I, Case III and Case IV, which yield many of the insights we are looking for and conclude with a brief discussion of the other cases.

3.3.1 Analysis of Case I

In this case, private and social rankings are aligned. Assume in addition that \(N_H < 1\) but \(L > 0\), so that it is optimal to give the leftover slots to \(L\) types.
To solve the government’s problem, let’s start with mechanism design problem. Consider the following candidate solution (we drop the success or failure subscripts wherever a particular type is not being tested):

\[
\begin{aligned}
    p_H &= y_L + \epsilon, p_L = y_L \\
    \pi_H &= 1, \pi_L = \frac{1 - N_H}{N_L} \\
    t_H &= t_L = 0
\end{aligned}
\]

Notice that the low types would not want to pretend to be high types. They cannot pay \(p_H\). What about the high type? If they pretend to be the low type they could pay \(\epsilon\) less but they would receive the slot with a probability less than 1. As long as

\[
h - (y_L + \epsilon) \geq \frac{1 - N_H}{N_L} (h - y_L),
\]

the high types would prefer to pay the price and be guaranteed a slot. We can always set \(\epsilon\) low enough to ensure that this is the case. Therefore, the mechanism is incentive compatible for small enough \(\epsilon\). Since both types are getting positive expected benefits, the participation constraints are also satisfied. It is feasible since the ratio \(\frac{1 - N_H}{N_L}\) was chosen precisely to exhaust the total number of slots. Finally, it is affordable, as long as \(\epsilon\) is small enough since \(y_H > y_L\) in this case.

Define \(E\) to be set of \(\epsilon\) such that this mechanism is in \(\mathcal{R}\).

This is also social welfare maximizing since every \(H\) type gets a slot and every slot is used up and no one gets tested. The key question is whether the bureaucracy will want choose this mechanism for some \(\epsilon \in E\)–if he will choose it then the government’s problem is solved. However, it is possible that he might prefer an alternative mechanism.

Given our assumption that there is a fixed cost of breaking the rules, if he is corruptible and chooses to break the rules, he will choose the mechanism that maximizes his payoff given by 1. Therefore, he will want to maximize the amount of revenue he can extract. The mechanism already allows him to extract all possible revenues from type \(L\). To maximize his payoff (within this class of mechanisms), he will set \(\epsilon\) to its maximal value in \(E\). That is he will set:

\[
p_H = p_H^* = \min\{y_H, y_L + (h - y_L)\frac{N - 1}{N_L}\}.
\]

Let us, with some abuse of terms, call the "auction mechanism" the following:

\[
\begin{aligned}
    p_H &= p_H^*, p_L = y_L \\
    \pi_H &= 1, \pi_L = \frac{1 - N_H}{N_L} \\
    t_H &= t_L = 0
\end{aligned}
\]

However, in this scenario, he is not extracting all of the rents from the type \(H's\) since \(p_H\) might be lower than \(y_H\). What are other mechanisms that could potentially give him higher payoffs?
One is the class of "monopoly mechanisms": Set
\[ p_H = \tilde{p}_H \leq y_H, p_L = y_L, \]
\[ \pi_H = 1, \pi_L = \min \{ (h - \tilde{p}_H) / (h - y_L), 1 - \frac{N_H}{N_L} \} \]
\[ t_H = t_L = 0 \]
These mechanisms are constructed so that the probability of getting the slot as an L type is low enough that no H type will want to pretend to be an L type. No L type can afford the slot at the H type’s price, so that incentive constraint also does not bind. By construction, these mechanisms also satisfy the slot constraint, as well as the participation constraint and the affordability constraint. However, they generate an inefficient outcome as some slots are wasted.

Obviously, this class of mechanisms will only interest the bureaucrat if
\[ (h - y_L) \frac{N_L - 1}{N_L} + y_L < \tilde{p}_H \leq y_H. \]
The condition that it makes more money than the auction mechanism is that:
\[ N_H (\tilde{p}_H - p) + N_L (h - \tilde{p}_H) (y_L - p) \]
is increasing in \( \tilde{p}_H \), since for \( \tilde{p}_H = (h - y_L) \frac{N_L - 1}{N_L} + y_L \), this is exactly the payoff from the auction mechanism. The relevant condition is therefore:
\[ N_H > N_L (y_L - p) / (h - y_L). \]
If this condition holds, the monopoly mechanism that maximizes the bureaucrats earnings will have \( \tilde{p}_H = y_H \). Otherwise, the auction mechanism dominates.

Finally, the last alternative we will consider is the "testing mechanism":
\[ p_H = \min \{ y_H, h - (h - l) \frac{1 - N_H}{N_L} \}; p_{LS} = p_{LF} = y_L. \]
\[ \pi_H = 1, \pi_{LS} = \pi_{LF} = \frac{1 - N_H}{N_L} \]
\[ t_H = 0, t_L = \max \{ 0, \frac{1}{\delta} \min \{ (h - y_L) \frac{1 - N_H}{N_L} - (h - y_H), (l - y_L) \frac{1 - N_H}{N_L} \} \} \]
The exact construction of this mechanism is less obvious so let us look at it in a bit more detail. The idea of this mechanism is to use testing just to reduce the rents of the self-declared L types so that H types would not want to pretend to be L types. It is inefficient because testing is wasteful. Since H types are more likely to pass a test than L types, it would be counterproductive to reward "passing" since our goal is to discourage H types from pretending to be L types. Rewarding failing the test also does not work since H types can always fail on purpose. Therefore, there is no advantage on conditioning on test outcomes.
To see that testing relaxes type H’s incentive constraint, note that now it becomes:
\[ (h - p_H) \geq (h - y_L) \frac{1 - N_H}{N_L} - \delta t_L. \]
Clearly $p_H$ can go up when $t_L$ goes up, which is why the bureaucrat might want it. However, there is obviously no point in driving $t_L$ past the point where $p_H = y_H$. This defines one limit on how large $t_L$ should be:

$$(h - y_H) - (h - y_L) \frac{1 - N_H}{N_L} \geq \delta t_L$$

Another limit comes from the fact that, by imposing testing, the $L$ type is being made worse off. So $t_L$ must satisfy $IR_L$:

$$(l - y_L) \frac{1 - N_H}{N_L} \geq \delta t_L$$

As long as $IR_L$ is not binding, raising $t_L$ always pays off in terms of allowing $p_H$ to be raised. Once it binds, it is possible to continue to increase $t_L$ by reducing $p_L$ below $y_L$. However, this will never pay off since reducing $p_L$ also forces the bureaucrat to reduce $p_H$. Setting $\delta t_L = (l - y_L) \frac{1 - N_H}{N_L}$ and plugging this into type $H$’s incentive constraint gives us the limit on how high we can drive $p_H$ by testing $L$ types

$$p_H \leq h - (h - l) \frac{1 - N_H}{N_L}.$$ 

Putting these observations together explains why we construct the testing mechanism in this way.

It is also worth observing that $t_L = 0$ when $y_L = l$. This is a consequence of the fact that when $IR_L$ is binding, red-tape will never be used; this tells us that the fact that the bureaucrats clients are unable to pay the full value of what they are getting is key to getting red-tape (that is why they pay in "testing" rather than money).

These three mechanisms do not exhaust the class of feasible mechanisms. For example, it may be possible to combine the testing and the monopoly mechanisms. However, it is easy to think of situations where each of them may be chosen by some bureaucrats, depending on the rules that the government sets and other parameters. This is mainly what we need understand, i.e. the trade offs that this model generates.

**Scenario 1** Suppose it is the case that $(h - y_L) \frac{N - 1}{N_L} + y_L \geq y_H$. Then, the auction mechanism extracts as much rents as possible. The government can give the bureaucrat full discretion (no rules) and expect the optimal outcome. It can then set $p$ to appropriately divide the surplus between itself and the bureaucrat. The bureaucrat choose the auction mechanism.

**Scenario 2** Suppose it is the case that $(h - y_L) \frac{N - 1}{N_L} + y_L < y_H$. Assume $\pi_{xr}$, $p_{xr}$ and $t_x$ are contractible. The rules do not impose any restrictions on the choice of $t_x$. Also, assume that the bureaucrat has no cost of testing: $\nu = 0$ and that it is possible to extract maximal rents from the type $H$ by testing the $L$
type, which will be true when:\footnote{As long as $y_L < l$, this is consistent with the condition $(h - y_L) \frac{N_h - 1}{N_L} + y_L < y_H$ imposed above.}

\[ y_H \leq h - (h - l) \frac{1 - N_H}{N_L}. \]

Suppose first that the government sets no rules. Since the bureaucrat pays no cost for testing and testing allows him to extract maximal rents, he will choose the testing mechanism described above as a way to create artificial scarcity.

One alternative for the government is to set the rules so that the maximum price the bureaucrat can charge is $(h - y_L) \frac{N_h - 1}{N_L} + y_L$ and all testing is disallowed. For those bureaucrats not prepared to break the rules, the optimal mechanism in this case will be the auction mechanism (since they were deviating from it precisely in order to charge the $H$ type a higher price, which is now not allowed).

However, those who have a low cost of breaking the rules (low $\gamma$) will deviate from the auction mechanism and choose the testing mechanism or the monopoly mechanism.\footnote{It is true that in our model we would get the same result with either a rule on testing or a cap on the price, but this reflects our extreme assumption that breaking one rule is the same as breaking them all. An epsilon extra cost of breaking two rules instead of one would make it strictly optimal to have both rules.} The testing mechanism tends to extract less money from each $L$ type (because they also pay the cost of being tested) but more $L$ types get slots. Which of the two will be chosen will depend on the parameter values—for example, an increase in $y_H - y_L$ keeping $l - y_L$ fixed, will make the monopoly mechanism relatively more attractive (intuitively when the $H$ type can pay relatively more, the cost of including the $L$ type goes up). If the monopoly mechanism is chosen there will be no red-tape, but large bribes (price above the maximum allowed price). If the testing mechanism is chosen, we will observe both bribery (price above the maximum allowed price) and red-tape. Nevertheless, from the social welfare point of view this outcome is strictly better than the no rules outcome since a fraction of the bureaucrats (those with high $\gamma$) choose the auction mechanism. What is particularly interesting here, though, is that the rules themselves are now affected by the potential for corruption. A different set of rules make sense when the bureaucrats are more corruptible.

\textbf{Scenario 3}  
Suppose it is the case that $(h - y_L) \frac{N_h - 1}{N_L} + y_L < y_H$. Assume $p_{xr}$ and $t_x$ are contractible but $\pi_{xr}$ is not.\footnote{Within our model, since the bureaucrat always pays the government for the slots, the government actually knows how many slots he used up and therefore should able to contract on $p_{xr}$, However, it is easy to think of an extension of the model where there is a state of the world where the demand for slots is lower and the government does not observe this state.} However, let $\nu$ be very high so that the bureaucrat is not prepared to use red-tape.

In the absence of any rules the bureaucrat will either choose the auction mechanism or the monopoly mechanism. We already generated the condition under which the monopoly mechanism makes more money:

\[ N_H > N_L \frac{(y_L - p)}{(h - y_L)}. \]
Interestingly, this condition is less likely to hold if \( p \) is lower—the government may be better off not charging the bureaucrats for the slots. However, even with \( p = 0 \), it is possible that the above condition holds (especially if \( y_L \) is very low), and the bureaucrat, unconstrained, would choose the monopoly mechanism. Suppose that this is the case. Then, the no rules outcome will leave many slots unallocated.

The government may prefer to set a rule where the prices that can be charged are capped by \((h - y_L)\frac{N_H - 1}{N_L} + y_L\). Then, the bureaucrats who have high \( \gamma \) will choose the auction mechanism while the low \( \gamma \) bureaucrats will choose the monopoly mechanism. There will be bribery because the monopoly price is higher than the price cap.

### 3.3.2 Analysis of Case III

Let us focus on one special case where \( L > 0, N_H < 1, h > l, y_H = y_L \) (which are the assumptions under which this case is analyzed in Banerjee (1997)) and to limit the number of cases, let \( y_H = y_L = y < l \) and \( \phi_L(t) = 0 \), i.e. no one ever fails the test.

In this case, once again there is an auction mechanism:

\[
\begin{align*}
p_H &= y, p_L = p_L^* \\
\pi_H &= 1, \pi_L = \frac{1 - N_H}{N_L} \\
t_H &= t_L = 0
\end{align*}
\]

where \( p_L^* \) is such that:

\[
l - y = \frac{(1 - N_H)}{N_L}(l - p_L^*).
\]

This mechanism implements the efficient outcome because the high types, though they cannot pay more, value the slot more \((h > l)\) and hence would rather pay the high price (all they can afford) and ensure a slot rather than risk not getting one at the low price. The logic of auctions still works.

However, now consider an alternative "testing mechanism":

\[
\begin{align*}
p_H &= y, p_L = y \\
\pi_H &= 1, \pi_L = \frac{1 - N_H}{N_L} \\
t_H &= t_H^*, t_L = 0
\end{align*}
\]

where \( t_H^* \) is given by

\[
l - y - \delta t_H^* = \frac{1 - N_H}{N_L}(y - l).
\]

As in scenario 2, testing only happens when \( l - y > 0 \).

\[\text{\footnote{This only works if } y \text{ is high enough. Otherwise } p_L \text{ might have to negative.}}\]
And a third "lottery mechanism," given by:

\[
\begin{align*}
    p_H &= y, p_L = y \\
    \pi_H &= \pi_L = \frac{1}{N} \\
    t_H &= 0, t_L = 0
\end{align*}
\]

The bureaucrat charges everyone \( y \) and simply holds a lottery to allocate the slots.

**Scenario 4** Suppose that \( \pi_x, p_xr, t_x \) are all contractable and \( \nu = 0 \).

What would happen if the government set no rules? The bureaucrat would always prefer the lottery, with very significant misallocation of slots.

Now, suppose the government sets the rules so that the bureaucrat is required to choose:

\[
\pi_H = 1, \pi_L = \frac{1 - N_H}{N_L}
\]

but there is no rule for what prices he can charge or the amount of testing. Every bureaucrat will choose the testing mechanism since it gives them the same payoff as the lottery without breaking any rules.

Suppose the government wants to stop this unnecessary testing. Then, it can set rules so that the bureaucrat is required to set the auction mechanism:

\[
\begin{align*}
    p_H &= y, p_L = p_L^* \\
    \pi_H &= 1, \pi_L = \frac{1 - N_H}{N_L} \\
    t_H &= t_L = 0
\end{align*}
\]

This mechanism will be chosen by those bureaucrats who have high enough \( \gamma \). However, the low \( \gamma \) bureaucrats will choose the testing mechanism and there will be both bribery and red-tape.\(^\text{15}\)

Alternatively the government could choose the lottery as the rule. All bureaucrats would then choose it and there would be no corruption and no red-tape but the outcome that everyone chooses would involve misallocation. However, since there is no testing, this outcome might be better than the outcome from the testing mechanism if the cost of being tested, \( \delta \), is high enough. Moreover, the testing mechanism is only better because it makes the high \( \gamma \) people choose the optimal mechanism. Therefore, if most bureaucrats face a low value of \( \gamma \), representing a government that cannot enforce the rules very well, then the lottery mechanism is likely to dominate.

**3.3.3 Analysis of Case IV**

Let us restrict our attention to the specific situation where \( N_H > 1, y_L = l > h = y_H, L < 0 \). The goods are scarce, but the private valuation of the high

\(^{15}\)Once again we assume that the bureaucrat (at least weakly) prefers to break one rule rather than two.
types is lower than that of low types. The low types should ideally not get the slots.\footnote{For examples, see Laffont-Triola (1993).} The analysis in this section draws upon Guriev (2004). Consider the following "testing+auction" mechanism:

\[
\begin{align*}
    p_{HS} &= p_H^*, p_{HF} = p_L = l \\
    \pi_{HS} &= \frac{1}{N_H}, \pi_{HF} = \pi_L = 0 \\
    t_H &= t_H^*, t_L = 0
\end{align*}
\]

where \(t_H^*\) and \(p_H^*\) solve the two equations.\footnote{We assume that solution with \(p_H^* \geq 0\) exists, which is true when \(l - h\) is not too large.}

\[
\begin{align*}
    h - \delta t_H^* - p_H^* &= 0 \quad (2) \\
    (1 - \phi_L(t_H))(l - p_H^*) - \delta t_H^* &= 0. \quad (3)
\end{align*}
\]

It is easy to check that this mechanism satisfies all of the constraints. Of particular interest are the type \(L\)'s truth-telling constraint:

\[(1 - \phi_L(t_H))(l - p_H^*) - \delta t_H^* \leq 0\]

that says that type \(L\) voters weakly prefer not getting the slot to pretending to be a type \(H\) and getting it with some probability. It is clear that for this to hold, it must be that \(t_H^* > 0\), since without testing the \(L\) type always want it if the \(H\) type does. Testing is necessary in this case.

It is also the mechanism that implements the optimal allocation (only \(H\) types get the slots) with the least amount of testing.

However, the bureaucrat may consider other mechanisms. One possibility is a straight auction:

\[
\begin{align*}
    p_H &= p_L = l \\
    \pi_H &= 0, \pi_L = 1/N_L \\
    t_H &= 0, t_L = 0
\end{align*}
\]

Another is a lottery. No one gets tested but the allocation is all wrong-only \(L\) types get the slots:

\[
\begin{align*}
    p_H &= p_L = h \\
    \pi_H &= 1/N, \pi_L = 1/N \\
    t_H &= 0, t_L = 0
\end{align*}
\]

**Scenario 5** Suppose that \(x_r, p_{xr}\) and \(t_x\) are contractable and \(\nu > 0\). Consider what would happen absent any rules. The auction mechanism maximizes the bureaucrat’s earnings without any testing and will be chosen. Now suppose the government sets rules about \(t_x, \pi_{xr}\) and \(p_{xr}\) exactly at the level given by the
testing-auction mechanism. Bureaucrats who do not want to break the rules will then choose the testing-auction mechanism. The ones who are prepared to break the rules will choose the auction mechanism. There is bribe-taking, shirking and misallocation of resources (kind of like in Bandiera, Prat and Tommaso (2009)).

However, the government could also give up on trying to get the ideal testing and auction mechanism implemented. It could set the rules corresponding to the lottery mechanism. The advantage of this mechanism is that the bureaucrats are making more money from the slots since \( h > p_H \) and spending less effort testing and hence the gains from deviating are smaller. The disadvantage is that some slots go the type \( L \)s even if the bureaucrat is not corrupt, but the upside is that fewer of them deviate from the rules and give all of the slots to the type \( L \)s.

3.4 Interpretation of results

The above analysis makes clear the essence of our approach. Governments are interested in setting rules when the laissez-faire outcome does not maximize social welfare. Put simply: in this model, governments only interfere to improve upon an inefficient situation. Corruption, however, results when these rules do not extract maximum surplus for the bureaucrat. Sometimes the rules allow the bureaucrats to extract surplus exactly as he wants (e.g. our Case I) but in many other cases it may not. The task assigned to the bureaucrat and the rules are chosen by a government cognizant of the possibility for corruption. In several of the cases, it is clear that the presence of corruptible bureaucrats changes the rules and tasks. The government chooses those rules, and the task assigned to the bureaucrat, recognizing that the rules will sometimes be broken. The overall outcome is still improved by setting those rules. This is the essence of tasks approach.

However, the model also offers a number of more specific insights. The first observation is that, red tape goes hand-in-hand with bribery. Given that testing is costly, there is no reason to overuse it unless there is extra money to be made.

However, there are two distinct reasons for using it. When the willingness to pay is aligned with ability to pay and social valuation (our Case I, specifically Scenario 2), red-tape is faced by \( L \) types, i.e those who have a low probability of getting the good, and is designed to create some artificial scarcity and extract more rents for the bureaucrat (along the lines suggested by Shleifer-Vishny (1994)). In other words, the purpose of the red-tape is to screen in the high willingness to pay types. When ability to pay is not related to the willingness to pay (our Case 3, Scenario 4) then red-tape emerges because even corrupt bureaucrats prefer to generate as efficient an allocation as possible, conditional on not making less money. The red-tape is then on the \( H \) type and the purpose is to screen out the low willingness to pay types.

The second point is that red-tape only emerges when \( y_L < l \). Moreover, it is easy to check that in both cases red-tape is increasing in the gap between the willingness to pay and the ability to pay. The intuition is simple: It is precisely
because there is this gap that it is possible to impose red-tape. If there were no such gap, the client would simply walk away if he was faced with a lot of red-tape. This makes clear why people do not have to go through red-tape when they try to buy most marketed goods (which are goods for which \( l = y_L \)).

 Governments are associated with red-tape, in this view, because governments often supply goods to people whose ability to pay is less than their willingness to pay. For that reason, this is particularly a problem for governments serving the poor.

A third point, which is related to the first, is that red-tape does not have to be a result of scarcity. Scarcity may have a positive or a negative effect on red-tape depending on whether we are trying to discourage the \( H \) type from pretending to be an \( L \) type (our Scenario 2, where scarcity reduces red-tape) or the opposite (our Scenario 4, where scarcity increases red-tape).

A fourth point comes from thinking about the correlation between red-tape and bribery. This correlation is emphasized by La Porta et al. (1999) who, looking at cross-country data, argue that the fact that testing is positively correlated with bribery is evidence that a lot of it is unnecessary and hence red-tape. As we already observed, we only get red-tape when there is bribery. However, we cannot assume that there is always more red-tape when there is more bribery—that depends on what is the underlying source of variation as well as the underlying economic problem. For example, both in Scenario 2 and in Scenario 4, the incentive to move away from the auction outcome towards the testing outcome is always stronger when \( l - y_L \) is larger and therefore there will be more red-tape and more corruption when \( y_L \) is lower, as long as the government sets rules that correspond to the socially efficient mechanism. On the other hand, as already remarked, in Scenario 2, when \( y_H - y_L \) goes up keeping \( l - y_L \) fixed, corrupt bureaucrats will switch from testing to monopoly. There is an increase in the size of bribes and a fall in red-tape.\(^{18}\)

A fifth point follows from Scenario 5. Bribery can be associated with red-tape, but it is also associated with shirking, which is kind of the opposite.

A sixth point follows from the observation in Scenario 4 that a rightward

\(^{18}\)We would also see a negative correlation if we were prepared to go outside the world of our model and assume that the government has other unmodeled reasons that do not have to do with promoting efficiency for setting a price cap. We would also see a negative correlation if we were prepared to go outside our model and assume that the government has other reasons that do not have to do with promoting efficiency for setting a price cap. For example, the government may be sensitive to distributional or ideological concerns—government bureaucrats should not be seen to be getting too rich, even if that is what efficiency demands. Or there may be some political or symbolic argument for making all public goods free or very cheap. This might lead the government, to set the price cap below \( y_L \). Take the extreme case where the good is supposed to be free (the cap is zero). In this case the government would actually want the bureaucrat to test those who apply for the slots (since the price mechanism will not do any screening). Here a bureaucrat who is corrupt might actually do some good: all he needs to do is to raise the price for the \( H \) types to \( y_H + \varepsilon \) and scrap the red-tape and social welfare would unambiguously go up. In such cases, bribe-taking would be an antidote to red-tape and they would move in opposite directions. This is the old idea that corruption greases the wheels and increases efficiency (Huntington (1968)), but note that it can only happen when the government is not using its choice of rules to maximize welfare.
shift in the distribution of $\gamma$ will lead the government to switch from rules that correspond to the lottery mechanism to rules that correspond to the testing mechanism, which will lead to greater bribery and more red-tape, but less misallocation and therefore worthwhile. In other words, greater state capacity might be associated with more red-tape and greater bribery. Conversely, the fact that there is no bribery cannot be interpreted as evidence that all is well.

A seventh observation is that corruption might force the government to give up on trying to maximize the revenues it gets from the sale of the slots. In Scenarios 1 and 3, we saw that the government should set the price of the slots to the bureaucrat as low as possible to reduce corruption.

Finally, it is clear that the government often faces a choice between more desirable mechanisms that are more subject to corruption and less desirable mechanisms that are less so. This is exactly the choice we saw in Scenario 4, but also in Scenario 5. Corruption is therefore an outcome of this choice.

### 3.5 Firms and Governments: A Digression

The broad framework developed above would apply, *mutatis mutandis*, to any situation where the principal cares about the assignment of "slots" to the right people but not how much money he makes from the slots, but the slot allocation is implemented by an agent who does not care about who gets the slots but how about how much money he makes.

However, much of what is interesting here also relies on two additional assumptions. First, the private valuation of those who get the slots is not necessarily equal to the value the principal puts on giving them the slots. Second, the private valuation may not be the same as the ability to pay.

These assumptions are quite natural in the context of thinking about governments and similar organizations such as non-profit organizations. However, in the conventional way of looking at market transactions (think of the principal as the manufacturer, the agent as the retailer and the clients as buyers) none of these three defining assumptions probably apply. We now recognize that there are lots of transactions involving firms where the firm is not seeking to maximize short-run profits: This may be because of multi-tasking, reputation or signalling, for example, or because the transaction is internal to the firm (who should be promoted, for instance). In such cases, our first two assumptions are likely to be satisfied. However, the third key assumption, about the divergence between the willingness to pay and the ability to pay is less obviously applicable.

Given this, it is Cases I and IV that probably fit the private sector best. We know that there can be corruption in both of these cases (Scenario 2 or Scenario 5) and this is true even if the ability to pay and the willingness to pay are the same ($y_L = l, y_H = h$). However, as already observed, there will not be red-tape in such cases, though there may be shirking.

However, there are potentially other sources of corruption that do not arise in our model that would arise in private firms. This is because there may be conflicts over the division of the surplus with a profit-maximizing principal—for variety of different reasons (such as the agents may be credit constrained), the
firm may want to share enough of its realized revenues with its agents to align their incentives. This may be even more of an issue for private firms than in governments, though as we will discuss in the next sub-section, it is central to some things that governments do such as tax collection.

There are also many reasons that are not in our model for why the government may have more corruption than private firms. One big difference between firms and the government is that even when a lot of the transactions are within the firm, there is still an ultimate principal (or principals) whose business is maximizing profits. This could place a limit on the level of rents that could possibly be captured. Consider corruption in promotion decisions. As we said earlier, this is a context where we might expect corruption in both sectors. In the government context, there are jobs that one would like to capture because they offer rents in terms of bribes from customers, e.g. the job of a policeman. Thus, there is both the capacity to engage in corruption and a willingness to do so (there are big rents). In the firm context, because promotions are also not allocated through a market mechanism, there is room for corruption. However, the upside gains from capturing the job are far more limited: the customers are paying market prices and hence there are no rents to be had from them. At best, one gets the rents from a higher paying job and, in most cases, these are nowhere near the rents to be had from having access to customers who are willing to pay a bribe.

A related point is that governments and firms are held accountable through different mechanisms. Governments have to convince voters to reelect them but the welfare of the voters depends on a combination of many outcomes. As a result, the electorate cannot simply use single metrics such as revenues or taxes collected to evaluate performance. Firms however are (much more easily) evaluated along a single dimension—profitability. This too generates room for corruption in government that does not exist in firms. For example, consider procurement procedures. In firms, there is a clear mapping from corruption in procurement to lower profitability, which the boss typically does not like. Whereas in the government the resultant rise in costs would probably get submerged somewhere in the general budget, and while the loss quality might be noticed by some voters, they would most likely place small weight on it in their decision of who to vote for. Indeed, we might speculate that those who are in favor of corruption in government have a stake in arguing against any single metric of performance.

In summary, firms have less slack at the edges than governments. This lack of slack limits the room for corruption within the organization. This, of course, does not completely eliminate it. Nor does it mean that all firms face little slack. There are surely poorly governed firms or firms for whom monopoly power or other factors generates slack. We are merely speculating that qualitatively, governments experience much more slack when it faces the customers and this slack translates into more room for corruption within the organization.
3.6 Limitations of our Framework

We end this section by discussing some of the limitations of our framework.

3.6.1 Monitoring

A clear limitation of our framework is made clear from the discussion in the previous subsection about the differences between external incentives faced by governments and private firms. It tells us very little about monitoring structures, other than the obvious point that better monitoring would make things better. The emphasis on monitoring in theorizing corruption goes back, of course, to the classic paper by Becker and Stigler (1974). However, while the insights from this paper remain fundamental, we see this more as a paper on agency problems generally and not particularly about corruption.

There is also now a literature that emphasizes the institutional aspects of monitoring. There is clearly a choice of who monitors (the superior bureaucrats, the community, the voters?) and how intensively (should there be pro-active public disclosures of public accounts, of the performance of individual bureaucrats, etc.?). However, while Bardhan and Mookherjee's (2006) model captures the important and basic idea that communities may have an informational advantage in monitoring and therefore in controlling corruption, the theoretical literature on this subject has not advanced very much beyond this point, though there is a lot of interesting empirical material available.

There is also an important theoretical literature that emphasizes the endogeneity of the effectiveness of monitoring. Lui (1986) makes the point that corruption may be harder to detect when everyone else is corrupt. Tirole (1996) shows that within a model where experimentation is costly, even when enough people are corrupt, everyone acts as if everyone is corrupt, which removes the incentive to be honest. There is also the idea that people feel less bad about being corrupt when everyone else is corrupt. As far as we are aware, there is very little rigorous empirical research based on these ideas, though they are obviously important.

3.6.2 Reinroducing Distributional Concerns: Understanding Extortion

Another limitation of our framework is that it assumes that the government maximizes total social welfare. However, ignoring the distribution of welfare between the beneficiaries and the bureaucrats is clearly wrong in many instances. One obvious example is tax collection. Tax collection is all about who pays. In such cases, the government might prefer an inefficient outcome because it achieves the distributional outcome better and may therefore create a more complex set of trade-offs than are permitted here.

Relatedly, the fact that the outside options in our model are set up to allow people to walk away from getting the slot is also potentially problematic, especially because it says that the outside option does not vary by type. Consider
a tax collection example. Suppose you are trying to get a tax exempt certificate because you have no money. An undeserving tax-payer (a type $L$), who actually can afford to pay the taxes, does have the option to walk away: if he does not have the certificate he can always pay the taxes and be done with it. The deserving tax payer (the type $H$), on the other hand, cannot pay the taxes. As a result, if he tries to walk away from trying to get the certificate, he risks prison. His outside option is worse than that of the type $L$.

More generally, our assumption about outside options limits the possibility of extortion. However, it has long been recognized that one reason many government functions cannot be privatized is because they open up possibilities of extortion. We need framework that can help us think better about these issues.

3.6.3 Screening on Multiple Dimensions

In our framework, the assumption that there is only one dimension of asymmetric information is very restrictive. The bureaucrat may want to know about both the beneficiary's type (because misallocation is punished) and his ability to pay (because he wants to make the more money) and they may not be perfectly correlated. A very simple example of what can happen in such situations is worked out in Banerjee (1997), but the general multi-dimensional screening case is not well-understood.

3.6.4 Modelling Corruption

Our framework also embodies one very specific view of why there is corruption in equilibrium. Corruption comes off the fact that the cost of violating rules varies across bureaucrats. However, as discussed at some length in Tirole (1986), there are other reasons. For example, the government may recognize that in some states of the world, the bureaucrat and his clients are in position to cut a profitable private deal that the government would like the prevent, but lacks the information to be able to do so. Why not then simply recognize that this deal will happen and make it legal? One possible answer is that there many other states of the world where the same deal would be available, but in these other states of the world, the government is in a position to detect such behavior and prosecute the bureaucrat and thereby prevent the transactions. However, if the courts cannot distinguish between these states of the world where these private transactions are proscribed and the other states of the world where there are not, the bureaucrat could always claim that the transaction was allowed and get away with it. By banning all transactions between the bureaucrat and his client, the government is creating the possibility that the rules will be violated from time to time, but gaining in terms of being able to prosecute the bureaucrat if he goes too far.
4 Measuring Corruption

Measuring corruption is the primary challenge in the empirical literature. Without robust measures, the theories cannot be tested, the magnitudes of corruption cannot be quantified, anti-corruption policies cannot be assessed, and so forth. However, measuring corruption is challenging, and even today, there are relatively few studies that are able to credibly describe the extent of the problem. As Bardhan (1997) notes in his review of corruption, “our approach in this paper is primarily analytical and speculative, given the inherent difficulties of collecting (and hence nonexistence) of good empirical data on the subject of corruption.”

The measurement challenges are driven by several problems. First, the very fact that corruption is illegal makes people reluctant to talk about it out of fear of getting into trouble and, possibly, also shame. Second, the existing literature—because of the theory it draws on and how it defines corruption—takes a transactional view of corruption. Measurement means finding out who bribed who and by how much. This is inherently harder to quantify. Third, the traditional narrow focus on monetary transactions also makes it more difficult. When a government official benefits by stealing “time”—he or she decides not to show up for work—random spot-checks can be very revealing (see Chaudhuri et al. (2006) for an example). Finally, measurement systems will evoke responses which make the measurement system less reliable. If the government has a monitoring system in place, people will adjust to it and find ways around it. As a result, these measurement systems will underestimate corruption.

Despite these difficulties, there were many early attempts at measuring corruption, relying on rich qualitative data, occasionally backed up by numbers. These studies gave the first real evidence about the channels through which corruption occurred and the possible methods to eliminate it. Wade (1984) in particular provides a detailed description of how irrigation engineers reap revenue from the distribution of water and contracts in a village in South India. The most fascinating aspect of this study was the documentation of a fairly formal system in which the engineers redistributed revenue to superior officers and politicians. In order to obtain transfers to lucrative posts, the junior officers paid bribes based on expectations about the amount of bribe money that can be collected from the post. Thus, the value of the bribe payment in the transfer process was higher for jobs that had greater bribe extraction possibilities. In essence, the ability to take bribes in a job induces bureaucrats to bribe others to get it. This highly detailed study gives a glimpse into the pervasiveness of corruption in this area and the mechanics of how it operates. It also illustrates how corruption is interconnected throughout the entire organization and raises the possibility that rather than trying to stamp out one aspect of corruption, it may be necessary to invoke coordinated policies to stamp it out throughout the system.

However, like all case studies the study raises questions of generalizability. Is there as much corruption in other contexts? Under what set of circumstances do these systems come into being?

19There are exceptions. For example, Tran (2008) gathers a comprehensive set of internal bribery records from a firm in Asia to document the cost of bribe payments over time.
Other early studies focused on anti-corruption policies. For example, Klitgaard (1988) provides several case studies of successful elimination of corruption, such as in the Hong Kong Police Force, Singapore Excise Department, and the Bureau of Internal Revenue of the Philippines. In all these cases, the levers used are intuitive from an agency theory perspective—more intense or better monitoring, replacing individual actors, etc. They also all seem to involve a person at the top of each institution who was eager to implement these changes. On the one hand, they represent a vindication of an agency theory of corruption. On the other, they raise the more fundamental question: if these levers for eliminating corruption is within the choice set of governments, why are they not implemented? While motivating one to have hope that corruption can be fought, these examples leave lingering questions about why conditions were ripe for these interventions in these cases, but not elsewhere. Is what we observe due to particular institutional factors in these settings, or due to more generalizable features of how governments function? What particular combinations of institutions, policies, or conditions would lead to similar steps being taken elsewhere? Should we expect the same consequences of similar anti-corruption policies in different settings?

To solve these inherent problems of case studies, the next set of attempts to measure corruption tried to provide consistent measures of corruption across countries. However, given the difficulty of inducing people to talk about corruption, these measures focus on collecting perceptions of corruption, rather than the actual bribes paid or the actual theft of resources. This perceptions based approach has been carried out at quite a large scale, generating interesting cross-country and cross-time data sets. The Economist Intelligence Unit created one of the first of such datasets. The data collection effort consisted of factor assessment reports that were filled in by their network of correspondents and analysts which are then aggregated into risk factors for about 70 countries. This report included, for example, a question where the respondents report "the degree to which business transactions (in that country) involve corruption or questionable payments" on a scale of one to ten, where a high value implies good institutions. Other cross-country measures on subjective perceptions of corruption followed, including the Control of Corruption measures in the World Bank Governance Indicators (a description of the measures can be found in Kaufmann, Kraay and Mastruzzi (2004)), and measures by Transparency International. Each of these indicators uses a different methodology with its own advantages and disadvantages, which we do not go into here due to space constraints. The real advantage of such data is their breadth, which allows one to run large correlational studies. Mauro (1995) is an often-cited example of this

---

20 These are also called the Business International Indicators.
22 Svensson (2005) provides a thorough description of the differences between the most common cross-country indicators of corruption.
kind of study. He uses the Economist Intelligent measures in a cross-country growth regression equation to study the relationship between economic growth, corruption, and red tape.

While these perception-based measures of corruption provided evidence on which countries tend to report more or less corruption, many have pointed out their limitations. First, as Rose-Ackerman (1999) points out, it is unclear what the corruption indices actually mean, and what a particular rank implies about the type and level of corruption in a country. For example, in the Transparency International Corruption Perceptions Index for 2008, Brazil, Burkina Faso, Morocco, Saudi Arabia and Thailand all have the same index value. However, the value does not tell us what the form of corruption in these countries entails, nor does it tell us whether the types of corruption that we observe in these very different countries have different efficiency implications. For example, theoretical framework developed above, bribe-taking can actually promote efficiency if the problem is that the government is committed, for political or ideological reasons, to set the price cap below what it should be. Moreover, as we also point in the previous section, corruption often emerges as a result of government interventions designed to deal with some other distortion. These countries may have very different problems—why would we think that the gap between \( l \) and \( y_L \) or between \( h \) and \( y_H \) would be the same in Saudi Arabia and Burkina Faso, given their very different levels of per capita wealth? Or, to repeat a point we have made many times before, corruption may be a result of the government's attempt to fight some other form of misbehavior by its bureaucrats—for example, in Case I above, absent a price cap, the bureaucrat might choose the monopoly outcome and many slots may be wasted. A price cap, might move us to an outcome where there is more visible malfeasance by the bureaucrats (red-tape and bribe-taking), but less misallocation and less inefficiency overall. Greater corruption in one country could simply be a reflection of a greater willingness to fight corruption in that country. Since they do not give us information about the source of the corruption, these corruption indices actually tell us little about what types of governance interventions would help deal with these problems, or even whether we should reward/praise governments that have less corruption by these measures.

Second, perceptions may tell us little about the actual reality of situations since they are influenced by the way we see everything else. For example, perhaps when the economy is good, we perceive less corruption because we are more satisfied with the government. Olsken (2008), for example, compares the perception of corruption in a roads project with actual missing expenditures. He finds that while there is real information in perceptions, reported corruption is not particularly responsive to actual corruption. For example, “increasing the missing expenditures measure by 10 percent is associated with just a 0.8 percent increase in the probability a villager believes that there is any corruption in the project.” He also finds that the bias in perceptions is correlated with demographic characteristics, implying that perceptions of a non-random sample of the population may not adequately reflect real corruption levels. This is particularly problematic as many perception measures are not necessarily based on
random samples. For example, the measures from the Economist Intelligence Unit is based on the perceptions of foreign analysts, who may have different perceptions of corruption than the average citizen. Finally, and perhaps probably most importantly for our purposes, these data are most useful for cross-country or cross-geography analysis. They are less useful for testing micro-theories of corruption.

More recently, newer methods for measuring corruption have arisen that solve the small sample problem and move to more concrete measures of corruption. The first set of methods focus on refining survey and data collection techniques to improve the ability to assemble data on self-reported bribes/service delivery quality. For example, Svensson (2003) provides analysis from a dataset that provides information on bribes paid by firms in Uganda. To encourage truth-telling in the survey, the survey was conducted by a trusted employers' association, and also asked carefully worded, hypothetical questions to measure of corruption.23 Hunt (2007) utilizes the International Crime Victims Surveys and Peruvian Household surveys, both of which contain information on bribes to public officials if the individual has been the victim of a crime.24 Other studies collect “prices” paid for services that should be free. For example, Banerjee, et al (2004) collects fees paid at government health centers in India (that should mostly be free), while Antonessa, et al (2008) collected data on prices paid and quantities received from the public distribution system in India and compare them to the official prices for these commodities. The main benefit of these methods is a move to measures that have actual meanings. For example, using the measure in Banerjee, et al (2004), we can estimate the bribe amounts paid at the health centers, and use this information to understand how the bribes affect the allocations of health services. We can also use it as an outcome measure to study the effectiveness of anti-corruption policies in government health centers. However, these types of measures are limited if we believe that individuals are under-reporting bribes, and particularly if we believe that this under-reporting is biased by corruption levels. Moreover, most of the time, these measures are often limited to petty corruption since it is difficult to ask individuals about the larger bribes that they may have paid. Indeed, in many of these cases the reports are assumed to be reliable precisely because the agents do not know how much they should have paid, and therefore, do not see themselves as paying bribes. It is also clear that people might be more willing to reveal bribes that they paid in settings where the good they are getting is abundant (so that they are not displacing anyone else by paying the bribe).

The second method is the use of physical audits of governmental processes. For example, Chaudhury et. al. (2006) performed a multi-country study of teacher and health worker absence, where they performed spot checks to deter-

---

23For example, “Many business people have told us that firms are often required to make informal payments to public officials to deal with customs, taxes, licenses, regulations, and services, etc. Can you estimate what a firm in your line of business and similar size and characteristics typically pays each year?”

24Morgan (2008) also use the ICVS to determine what characteristics were associated with greater corruption.
mine whether bureaucrats were showing up for work.\textsuperscript{25} Similarly, Bertrand, et al (2006) followed individuals through the process of obtaining a driving license in India, and recorded all extra-level payments made and the rules that were broken in exchange for these extra-legal payments. Barron and Olken (2007) designed a study in which surveyors accompanied truck drivers on 304 trips along their regular routes in two Indonesian provinces. The surveyors observed the illegal payments the truck drivers made to the traffic police, military officers, and attendants at weighing stations.

One of the key challenges to the audit studies is whether the observed outcomes actually reflect corruption rather than some less intentional form of bureaucratic ineffectiveness, since often there is no smoking gun (bribe overtly paid, job left entirely unattended, etc.). For example, Bertrand et al (2006) find that there is a misallocation of licenses—people who cannot drive are able to get them. Could this be due to an “overloaded” bureaucrat who does not have time to screen license candidates or due to an “incompetent” bureaucrat who cannot distinguish between a good and bad driver? To understand this question, they collect detailed quantitative and qualitative data on how the bureaucrats behave through the licensing process. They document extreme behaviors (e.g. simply never administering a driving test) which would be hard to label “incompetence.”

Similarly, Duo, Hanna, and Ryan (2008) measure teacher absenteeism in rural India using audit methods. Does the fact that teachers often do not come to school imply that they are consciously breaking the rules for private gain? Or, are lives of the teachers so difficult that they just cannot make it to school often enough, despite trying as hard as they can? Their research design provides information that allows them answer these questions. Specifically, they evaluate a program that monitors the teachers and provides incentives to the teachers based on teacher attendance. They find that teachers are very responsive to the incentives. The very fact that the teachers respond to the incentives so strongly implies that the teachers were previously ignoring the rules, and were not just incapable of attending.

We refer to the third technique as cross-checking. The idea behind it is to compare official records of some outcome with an independently collected measure of that same outcome. One example of cross-checking is to compare how much money was released to the bureaucrat with how much the ultimate beneficiaries of the funding report having received. For example, Reinikka and Svensson (2005) compare data from records on central government disbursements and a public expenditure tracking survey to measure dissipation in a school capitation grant in Uganda. They find that the average school received only about 20 percent of central government spending on the program. Fisman and Wei (2004) compare Hong Kong’s reported exports to China at the product level with China’s reported imports from Hong Kong to understand the extent of tax evasion. Another way to conduct a cross check would be to collect records from the bureaucrat documenting how the government resources were spent in

\textsuperscript{25}Other papers that do this include Duo, Hanna, and Ryan (2008), who measure teacher absence, and Banerjee et al (2007), who measure absenteeism amongst medical providers.
achieving a task, and then to compare with an objective measure of how much it should have cost to conduct that same task. The difference between the two measures is, then, the estimate how much was “stolen.” Olken (2007) uses this method. He calculates corruption in the context of road projects, by comparing the actual expenditures reported with an independent measure of what it should have cost to build a road of that particular quality. To obtain the independent measure, he sampled each road to determine the materials and labor used, and then multiplied these by local prices. In the end, he finds that, on average, about 24 percent of expenditures across the project villages were missing.

As with any other method, this method has both strengths and limitations. Its innovation lies in the fact that it creatively allows the measurement of dissipated government resources without asking the actors involved if they have paid or received an actual bribe, reducing the chance of under or mis-reporting. Thus, it often allows us to move past petty corruption and perhaps look at larger scale corruption. However, it is difficult to understand whether the dissipated resources are actually corruption, or simply mis-measurement in the indicators or even just a sign of bureaucrat incompetence. For example, in Olken (2007), it remains possible (though unlikely) that road quality is mis-measured, or that the bureaucrats are not good at building roads? It is possible (though again unlikely) that the missing resources reflect that the bureaucrats are trying to reallocate funds to better uses. For example, in Renikka-Svensson (2005), it is possible (although again unlikely) that the resources that should have gone to the schools capitation program were actually spent on services that the community deemed more important, and did not end up in the pockets of government officials. Would this have necessarily been bad?

One way to get around these concerns is to look for correlations (motivated by theory) between the extent of dissipation and some other variable. For example, to show that the differences in reported shipments is corruption, and not just mis-measurement in the shipments, Fisman and Wei (2004) document that the differences are negatively correlated with tax rates on closely related products, suggesting that evasion takes place partly through mis-classification of imports from higher-taxed categories to lower-taxed ones, in addition to under-reporting the value of imports.

26 Other examples of cross-checking in the developing world include Hsieh and Moretti (2006) to estimate bribes in Iraq’s Food for Oil Program. Olken (2006) and Antonassavva et. al. (2008) measure theft in food distribution programs using the same methods, while Di Tella and Schargrodsky (2003) use it to measure corruption in hospital procurements.

27 Duggan and Levitt (2002) provide an interesting example of cross-checking in sumo wrestling. They basically show that a wrestler has a higher probability of winning than expected when the match is key to his rank. To distinguish match throwing from effort, they use theory as a guide: the effect goes away when there is greater media scrutiny, suggesting that it is not effort. Moreover, the next time the same two wrestlers meet, the opponent is more likely to win, suggesting that throwing future matches is a form of the bribe paid for winning a key match. Similarly Antanasovs et. al (2008) cross-check receipt of a BPL card (which in India identifies someone to be poor and allows them access to a set of redistributive programs) against actual income levels. They correlate the error rate with features such as caste of the recipient, their place in certain social networks etc., and argue that the correlations are what a simple theory of corruption would predict.
In addition to these methods that have been recently used in the literature, there are several innovative methods that are being explored in current studies. While the work is not yet published, methods such as these will surely contribute to our tools available. For example, Banerjee and Pande (2007) attempt to use second-hand data on which politicians have gotten rich since they entered politics to identify those who have profited from corruption.\textsuperscript{28} They find that this measure correlates strongly with other more direct (perceptual) measures of corruption (such as the answer to the question "do you think the politician used his office for private gain"). Banerjee, Olken and Pande (2008) are collecting data on public procurement records, and using information on number of qualified and不合格 bidders, characteristics of the winning bid, and the final price of contracts to understand whether officials are bending procurement rules.

All these methods however pick up the direct correlates of corruption, such as bribes and absence from work, rather than its more indirect ramifications. In some cases, such as teacher absenteeism, the direct consequence may be the most important. However, the point of our framework was to argue that in many situations, the bribe may just be the tip of the iceberg, with the more serious repercussions showing up in terms of misallocation and red-tape. The next section discusses an example of empirical research that tries to capture the bureaucrat's entire decision process and its various ramifications.

5 Understanding the Structure of Corruption

The focus of much of the empirical research has been based on measuring the extent of bribery or shirking and on how incentives affect these. Bertrand et al. (2007) differs from much of the empirical research by focusing on the entire resource allocation problem faced by the bureaucrat and therefore looking beyond just bribe-taking as the measure of welfare. The basic strategy of the paper is to experimentally vary the underlying types a bureaucrat faces and use the bureaucrat's responses to infer the allocative mechanism used by the bureaucrat.

Specifically, they compare three randomly chosen groups of license candidates. The first group was told to obtain a license as usual, the second group was given a large incentive to obtain the license in the minimum legal time allowed (30 days) and a third group was offered free driving lessons. Within our model the second group represents a situation where \( h \) and \( l \) have both gone up by the same amount. The third group represents a situation where some of those who were low types will now be a high type.

The driving license case, as noted above, corresponds to our Case IV. To reduce the number of possible cases, assume that \( y_H = h < l = y_L \) (which, as always, is to be interpreted as the case where a substantial fraction of \( L \) types

\textsuperscript{28}Di Tella and Weisbach (2007) provide a theoretical framework for thinking about unexplained wealth as an indicator for corruption.

32
are willing and able to pay more than a large fraction of $H$ types). Moreover, assume $N < 1$ but $L < 0$.

We will assume that the bureaucrat faces people drawn from this altered distribution of types. However, that because he does not know the environment has changed, both he and the government will use mechanisms that match the prior type distribution. This is because the bureaucrat would likely not have any direct knowledge that there an experiment is occurring. Of course, consumers can change their behavior as a result of the treatment: subjects who are more desirous of a license can offer higher bribes, for example. This may in principle allow the bureaucrat to learn of the experiment and adjust his overall strategy. In practice, we assume the samples involved are too small for the bureaucrat to change his behavior. In effect, in interpreting this experiment, we assume that it changes people's types rather than changing the bureaucrats anticipated type distribution.

It is easy to check that in Case IV, an equal increase in $h$ and $l$, without any change in the rules, will have a number of different effects. $L$ types will now want to apply even if they are assigned to a bureaucrat who is being "honest," i.e. a bureaucrat whose cost of breaking the rules are too high compared to the benefit to him from breaking the rules. Therefore, the fraction of those who apply to an honest bureaucrat and end up getting a license should go down. However, just because many more $L$ types apply, many more will (just by chance) end up passing and the average quality of those who get a license would go down. On the other hand, the gains from being corrupt would also be higher and this will reduce the fraction of bureaucrats who choose to be honest.

The "corrupt" bureaucrats, i.e. the ones who will opt to break the rules, of whom there are now more, will raise the price. The fraction of $L$ types getting through, conditional on being allocated to a corrupt bureaucrat, should not change but since there are now more corrupt bureaucrats, the average quality of driving among those who get a license should go down. Moreover, because there are now more corrupt bureaucrats, the average amount of testing should also go down.

By contrast, converting some of the high types in the population (the second treatment) to low types should improve the fraction of drivers who get a license when they get allocated to an honest bureaucrat, as well as the average quality of licensed drivers. Those who get allocated to a corrupt person, should get more or less the same treatment, with the licenses going to the highest bidders and perhaps some red-tape on those who want to pay less ($H$ types) to discourage those who are willing to pay more from claiming otherwise.

Empirically, individuals who were offered the incentive were 42 percentage points more likely to obtain it in 32 days or less. However, they paid about 50 percent more to obtain this licenses, and they were more likely to break a rule during the process (for example, they were 13 percent more likely to not take a driving test). In the end, these extra payments translated to a greater number of bad drivers on the road: those offered the incentive were 18 percent more likely to obtain a license despite not knowing how to operate a car. These results are entirely consistent with theoretical predictions discussed above, as long as
there are enough bureaucrats around who have a relatively low cost of breaking the rules. In particular, while the decline in the quality of driving among those who have a license would happen both with honest and corrupt bureaucrats, the reduction in testing can only happen in our model if the bureaucrat is corrupt.

It is worth pointing out that one would not observe the same pattern if the bureaucrats did not know that the distribution of $h$ and $l$ has shifted. In this case, if $h$ and $l$ go up by the same amount, both types would now want to be tested when they hit an honest bureaucrat. If they end up with a corrupt person, both types will just pay and get the license without being tested. The average price would not change and the fraction getting tested would go up. However, the quality of driving among those who get a license would go down since the $l$ types who hit the honest bureaucrat now want to be tested. This clearly does not fit the facts about price and testing.

The results of the driving lessons treatment are also broadly consistent with the theory. Those in this group are tested more often and more likely to have obtained a license based upon passing the test. This last fact, in particular, suggests that there are some honest bureaucrats around. They also pay less for the license, though much more than they should have legally. This is in fact exactly a prediction of the model we have, but only because we assumed (for convenience) that all $H$ types have a lower willingness to pay than all $L$ types. A more plausible argument is that the $H$ types actually have an incentive to shop around, i.e. go to multiple bureaucrats till they find one who is honest. In fact, they did observe shopping around in the data and it is therefore a plausible explanation for why the increase in the fraction of $H$ types reduces the average amount paid.

The paper also provides evidence that there is a lot of red-tape, i.e. pointless testing. They show that of those experiment participants who, at least initially, tried to follow the rules (get tested, not pay bribes, etc.), there is a higher success rate among those individuals found to be unqualified to drive based on the independent test (74 percent compared to 62 percent). In other words, the probability of getting a license is less than 1 even for those who can drive and it is not higher for them than for those who cannot drive. This suggests that passing the test is uncorrelated (at best) with driving ability. The testing serves no direct screening purpose.

\textsuperscript{29}However it also worth remarking that our assumption that bureaucrats are randomly assigned to applicants plays a very important role here. One alternative assumption would be that applicants can either choose to go through the official system knowing that an honest bureaucrat will be assigned to them with some probability, but otherwise they end up with a corrupt guy who will always fail you (because in the official channel there are no bribes—the corrupt bureaucrats find it at least weakly optimal to fail everyone who comes through the system). Or you can choose to go through the unofficial channel which guarantees that you pay a bribe and get the license. In the original equilibrium of this game, it is likely that all $H$ types will try the official channel while the $L$ types will go the other way. In this scenario, only $H$ types ever take the test and fail.

In this case, during the experiment, $H$ types will shift towards the corrupt route and therefore end up paying more, testing less and an increased probability of getting a license. However, counterfactually, average licensed driver quality will go up because the fraction of $H$ types who fail goes down.
Both these features—a probability of winning less than 1 and pointless testing—are consistent with Case IV, in the scenario where slots are abundant. However, to get pointless testing, it is important that $y_H < h$—otherwise you might as well raise the price all the way up to $h$. This divergence between the ability to pay and the willingness to pay, seems implausible in this context. The amount of money involved (around $25) is not large for the poor in Delhi, which is where the experiment was carried out.

A more plausible story might be that the applicants do not fully understand the rules of the game and therefore think that it is easier to use the official channel than it actually is, while the bureaucrat is not in a position to directly explain to them how things really work, and therefore is using the fruitless testing to signal to them that they need to readjust their expectations. This would be consistent with the fact that no one directly pays a bribe to the bureaucrat. Those who want to go the bribe route go to an agent, who facilitates the transaction. When someone actually approaches a corrupt bureaucrat through the official route, the bureaucrat does not ask for a bribe, and instead goes through the motions of what he is supposed to do, while presumably trying to make sure that the next time they go directly to the agent.

These type of empirical results are an intriguing complement to the theoretical framework we have laid out. They focus our attention on the allocative outcomes and not just the bribes. They focus on the details of testing and not only on the transfers made. In short, they illustrate the broader view on corruption that we advocate for in this paper.

6 Combatting corruption

There is a large and growing empirical literature that studies the effect of efforts to fight corruption. For example, Fisman and Miguel (2007), find that an increase in punishments for parking violations in New York City reduced the violations amongst the set of diplomats that were most likely to violate the rules. Using experimental techniques, Olken (2007) finds that the theft in road projects is greatly reduced by raising the probability of being caught. Dufo, Hanna and Ryan (2008) and Banerjee et al. (2007) find that strengthening incentives reduces absenteeism. Using a natural experiment in Buenos Aires hospitals, Di Tella and Schargrodsky (2001) find effects of both more stringent monitoring and higher wages on procurement prices.

Several more recent papers in this space have also tried to go a step further and understand whether or not a reduction in corruption due to monitoring and incentives improve the final allocation of services. In the case of the Di Tella and Schargrodsky (2003), less bribe-taking means better procurement prices, which is the outcome of interest. Olken (2007) looks at the effect of auditing not only on theft, but also on the quality of the roads that were built. Dufo, Hanna and Ryan (2008) study whether incentives can create distortions due to multi-tasking. In other words, they are concerned that in order to complete the task as specified by the incentives, the agent reduces their effort along other
dimensions. Specifically, they ask whether providing incentives for teachers to attend school will cause the teachers to compensate by teaching less. To answer this question, they measure not only teacher attendance as the final outcome, but also the learning levels of the children and find that the multi-tasking problem is certainly not large enough to outweigh the benefits of better incentives.

6.1 Institutional Structures for Monitoring

Corruption exists because there is not enough monitoring and incentives to eliminate it. However, then what determines the extent of monitoring and incentives?

One challenge with looking at this issue empirically is essentially methodological. What should we assume about the extent to which these rules are the product of optimization by the government? Governments may make rules to combat malfeasance by government officials, but how well does it understand the consequences of these rules for corruption?

There are two possible approaches to this issue. One is what one might call the political economy approach. This is the position taken in the theory section here. We assume some preferences for the government and figure out how the rules and the nature and extent of corruption should vary as a function of the underlying economic environment if the government was optimizing based on those preferences. We could then look for evidence for the comparative static implications of that model, and jointly test the model and our assumption about what is being optimized.

The alternative is to assume that the forces of political economy, while important, leave a significant amount undetermined and as a result changes in rules can often arise as pure organizational innovations, without changes in the "fundamentals." This approach leads naturally to an experimental approach towards studying the effects of the rules.

Banerjee, Munshi, Mookherjee and Ray (2001) implement a version of the first approach in the context of the governance of localized sugar cooperatives. They assume that the cooperative maximizes a weighted average of the profits of the various principals of these cooperatives – the different types of member-farmers who grow the cane that the cooperative turns into sugar – taking into account the fact that the management of the cooperative wants to siphon off as much of the profits as possible, and generate a set of comparative statics predictions that map the pricing of cane and the productivity of the cooperative onto the underlying mix of farmers in the cooperative. These implications are then tested and seem to be broadly confirmed.

However, corruption in these cooperatives is essentially private sector corruption, embodied primarily by the underpricing of cane. We are yet to come across a paper that combines this political economy approach with the more complex manifestations of corruption that we identified above.

More importantly, in many instances, our theories of political economy are simply too incomplete to be very useful guides to what rules actually get chosen.
The objective of a specific government at a specific point of time is some complex product of its long-term goals and its short-term compulsions and moreover, the way it chooses the rules must take account the compulsions of all future governments. While there is an interesting and growing literature on this subject, it is not clear that it is ready to be taken to the data.

It is also not clear how much governments understand about the consequences of the various policy choices that they are making or about the policy options that they have in hand. A more evolutionary approach to policy change, where changes happen because political actors are trying to solve some "local" problem but the solutions often have unanticipated and often global consequences, may be more descriptively accurate. Certainly it fits better with the kinds of stories that one hears about how change came to China.

An advantage of this second approach is that it permits us to think of policy changes as organizational innovations, and therefore at least initially exogenous in a way that technological innovations are usually thought of as being exogenous (in other words, the assumption is that the need to solve problems is a product of various forces of society, but the adoption of a particular solution at a particular point of time, less so). It also makes it clear that governments might choose bad rules (rules that go against its own objectives), because it does not understand the consequences of its choices.

A number of recent empirical exercises start from this point of view. For example, Besley, Rao, and Pande (2006) find that, in South India, there exists a relationship between holding village meetings (i.e. more community participation in the process) and better allocations of BPL (Below Poverty Line) cards, that provide privileged access to subsidies and government services. Bjorkman and Svensson (2008) study decentralization in an experimental context. Rather than imposing all centralized rules on health centers, community meetings to decide the most important rules that health centers should be following, and the mechanisms for the community to monitor the health centers. They find incredibly large impacts: infant mortality rates were cut by one-third. However Banerjee et al. (2008) who evaluate a similar decentralization program in India find that it performs no better than the civil service system based system of monitoring teachers.

It is difficult to think about what to make of these vastly different empirical findings given that we do not have a particularly good theory of how decentralization affects corruption and the distortions associated with corruption. How does decentralization change the kinds of rules that are optimal and the direction in which they get violated? How does the exact nature of decentralization factor into all this?

As it is, the presumption behind the empirical literature is that decentralization is a shift of control rights into the hands those who have more local information. The basic notion is that the community now has more information and therefore can limit the extent of malfeasance by the bureaucrat. We do see

---

30 Banerjee (2002) discusses many of these methodological issues in greater detail.
31 Bardhan and Mookherjee (2006) is a rare exception.
some evidence that the easy availability of information matters. In the driving license experiment, Bertrand et al. (2007) report that there are two things that bribes cannot get around: one is the requirement of showing some proof of address and the other is the requirement of waiting at least 30 days after making the initial application for a learner’s permit. Neither of these seem as important as being able to drive, especially in India where the driving license is not always accepted as an identification. However, violations of these rules is easy to observe, whereas the inability to drive properly is something that requires another test to verify. Therefore, these are the rules that get enforced.

However, the answer to the question of what rules gets violated in what way must also depend on who exercises which type of control rights and what information reaches whom, which all turns on the exact model of decentralization that gets adopted. More generally, theoretical work mapping the effects of alternative organizational forms on the choice of rules and corruption outcomes must be a high priority if this literature has to make progress.

7 Conclusion: Where should the literature go next?

We have already discussed a number of the gaps in the literature. In the previous section, we discussed the need to more tightly model “learning the system.” Below, we discuss other important gaps:

7.1 Corruption and Competition

Thinking about organizational forms naturally leads to the role of competition in reducing corruption, as emphasized by Rose Ackerman (1978). The way we modeled corruption takes as given the idea that the assignment of the applicants to the bureaucrats is random. This effectively places the applicants and the bureaucrat in a bilateral monopoly setting. However, the nature of competition between bureaucrats ought to be a policy choice governed by the nature of the underlying incentive problems. This is an area that needs further exploration.

In particular, competition is not always a plus. As pointed out by Shleifer and Vishny (1994) competition between a number of corrupt distinct and dis- coordinated authorities, who each has the power to block the "application" might be worse than a single monopolistic rent-seeker. Barron and Olken (2007), document this phenomenon using a unique data set that they collected in Indonesia of the bribes paid by truck drivers at road blocks. Reduction in the number of check-points along the road reduces the total amount of bribes collected from them. Credible evidence on the salutary effects of competition has so far been hard to found, though no doubt the right setting to look for them will emerge soon.
7.2 Implications of Illegality and Non-transparency

One reason that corrupt bureaucrats find it hard to coordinate with each other is that corruption is illegal. This essential non-transparency has a number of important implications that deserve further study. First, if the applicants for the slots differ in their ability to make illegal deals for either intrinsic or extrinsic reasons, then the playing field is no longer level, which introduces important distortions. This not merely a theoretical concern. Many countries have laws that forbid their firms to pay bribes in foreign countries, which could potentially act as a constraint on foreign investment in countries with high levels of corruption (see Hines (1995)).

Another fallout of this non-transparency that we already noted is the reliance on agents who provide intermediate bribe-taking. The theory how the use of agents alters the nature of corruption is yet to be developed. Barron and Olken (2007) provide an interesting insight into this. They observe that in Indonesia, truckers can either pay a bribe at every check-post, or pay a single bribe to an agent at the starting city. However, the contract with the agent tends to be very simple—the amount of the single bribe does not depend on the load carried by the truck—probably because of the same lack of transparency. This means that only the most overloaded trucks pay the fixed bribe, and the shape of the total bribe paid as a function of the truck's load is concave, whereas theory suggests that the optimal penalty function ought to be convex.

A third issue also came up earlier. We speculated that a lot of the drivers who try to get a license without paying an agent probably do not know the rules of the game. This happens because corruption is meant to be secret. In other words, understanding the process by which the real rules of the game become (or fail to become) common knowledge between the bureaucrat and the applicants should be an integral part of the study of corruption.

Bertrand, Djankov, Hanna and Mullainathan (2007) actually gather data about what individuals who are trying to obtain a drivers' license know about the licensing process. They find that not much is known, and more surprisingly, many applicants believe the official process is more onerous than it actually is. They also found in their qualitative work that discovering the actual rules was surprisingly difficult given that they change periodically.

This leads to them speculate whether the bureaucrats deliberately try to make the rules more complicated than they should be in order to extract more in rents. Antonassoa, Bertrand and Mullainathan (2008) find that individuals who are supposed to receive subsidized allocations of food-grains in India are mis-informed about their exact entitlements and the qualitative evidence in this case suggests that the shopkeeper often manufactures "rules" that increase the scope for his corruption, such as that all grain must be bought on one of two days.

Thinking about this issue leads us to an interesting theoretical possibility. Is it the possible that the government's attempts to change the rules, perhaps in

---

order to fight corruption, generates so much confusion among the citizens that corruption actually goes up?

7.3 Learning among Bureaucrats

The emphasis on learning brings up another important issue: While much of the work in the field has focused on innovations in fighting corruption, there has been little focus on the innovations in corruption. A change in policy/institutions may reduce the prevalence of corruption to start with, but over time, the bureaucrat may learn how to adapt to the new policy or institutions. For example, Camacho and Conover (2008) provide evidence that individuals were better able to game the eligibility rules for social welfare programs in Colombia as rules for eligibility became more known over time. More generally, how much of the knowledge regarding how to conduct corruption is general knowledge, versus knowledge to a specific institution?

7.4 Norms of Corruption

The idea that the rules may be important in establishing a simple norm that the courts can easily interpret, suggests a further thought. Perhaps the rules that the government makes for bureaucrats has a signalling role. The bureaucrat or the citizen uses them to infer the society’s preferences and therefore to figure out what they should and should not do. If the government formally allows its bureaucrats to extort money from its citizens, the citizens might take this as a signal that the moral standards of society are low, and therefore feel comfortable in extorting others. This could explain why governments continue to have rules on their books that are violated all the time.

On the other hand, a government that has rules on the books but does not manage to enforce them is also signalling something about its view of rules and rule-governed behavior that might spill over into other walks of life. For this reason and others, corruption may have a direct social cost, which is something our model does not take into account.

7.5 The social psychology of corruption

To fully understand how corruption (or lack of corruption) becomes the norm, there is a need to try to understand the psychology of when and where people feel more or less comfortable about engaging in corruption. For example, we often observe that there is a tendency to try to legitimize corruption. This could take the form of "excuse making," i.e. the bureaucrat not directly asking for a bribe, but instead discussing the costs of his or her time in providing a service to a citizen. Or alternatively, the citizen may suggest making a payment in kind, rather than a monetary bribe, to make the bureaucrat feel as if he or she is simply accepting a gift from a happy citizen, rather than engaging in an illegal act.
The concept of legitimization may be powerful part of our understanding why there is not as much corruption in the world as there could be. For example, even in the most corrupt countries, empirical antidotes suggest that the bureaucrats will often ask for a bribe to break a rule that goes against a given citizen, but will not threaten to punish him no reason. For example, traffic policemen often ask for a bribe if a citizen has committed a violation. However, they will not necessarily ask for a bribe if the person as done nothing wrong and yet it is not clear that the enforcement in these two cases is very different.

Locating the study of corruption in the broader context of how people relate to each other and to the state, may be very important in really getting a handle on why corruption exists in some settings, but not others.

8 Bibliography

References


