# A theory of informative red tape*

Sergei GURIEV†

September 2000

**Abstract**

We study emergence and interaction of red tape and corruption in a principal-bureaucrat-agent hierarchy. We assume that the red tape is informative: albeit imposing some cost upon the agent it also produces certain information about the agent. Therefore the socially optimal level of red tape is not trivial. Implementing the social optimum may be difficult if the bureacrat who operates the red tape is corrupt. The bureaucrat may extort bribes from the agent both ex ante (before setting the level of red tape) and ex post (after learning the information produced by the red tape). We show that if there is no threat of ex post corruption, the principal can implement the socially optimal level of red tape even if ex ante corruption is present. The threat of ex post corruption may lead to overproduction of the red tape, even though the ex post corruption does not occur in equilibrium.

*JEL Codes: D73, H11, K42*

*Keywords: informative red tape, corruption, three-tier hierarchy.*

---

*This is a substantially revised version of a working paper that was circulated under the title "A theory of informative red tape with an application to top-level corruption". I am grateful to Abhijit Banerjee, Richard Ericson, Arye Hillman, Mark Levine, Victor Polterovitch, Andrei Sarychev, Andrei Shleifer, Jean Tirole, Stephen Yeo, Katia Zhuravskaya, two anonymous referees and participants of M.I.T. Theory Lunch, EEA Congress in Berlin, EERC Conference in Moscow and seminar participants in Moscow and St.-Petersberg for helpful discussions and comments. Financial support from US Information Agency and Moscow Public Science Foundation and hospitality of M.I.T. Department of Economics where the work has started are gratefully acknowledged. All views expressed and remaining errors are my own.

†New Economic School and CEPR. Address: New Economic School, Nakhimovsky pr. 47, Moscow 117418, Russia. E-mail: sguriev@nes.ru

# 1 Introduction

In this paper we shall study emergence and interaction of different types of corruption and red tape in a hierarchy. Red tape and corruption are probably the most ancient and widespread diseases of government bureaucracy. They have been observed in all societies; there is also no reason to believe that they will soon disappear. Numerous attempts to fight either of them seem to have brought only limited results. One of the problems with corruption and red tape in bureaucracy is that they cannot be treated independently. Corruption in one part of hierarchy may stem from corruption in another part; excessive red tape may emerge due to potential corruption; bribes may be extorted because of potentially high red tape. When trying to make public bureaucracy more efficient, one should keep these interdependencies in mind in order to fight causes rather than consequences.

Economists have begun to study these issues only recently. The economic literature on corruption was initiated by Rose-Ackerman (1978). The first formal models analyzed corruption in a conventional principal-agent setting (see Bardhan (1997) for a survey). Later, Tirole (1986) has suggested to use a three-tier principal-supervisor-agent model as a more appropriate framework to study corruption in a hierarchy. Quite a few papers developed the analysis of the three-tier model further on (Kofman and Lawarree (1993), survey in Tirole (1992)). Some recent papers (Carrillo (2000), Bac and Bag (1998)) study four-tier hierarchies and show that there are even more issues in economics of corruption than the three-tier models suggest.

Red tape has received much less attention in the literature. An economic rationale of red tape is offered by Banerjee (1997) who analyzes red tape in general, and the link between red tape and corruption, in particular. In his model, government wants to provide a cash-constrained agent with a good which may be either of high or low value to the agent. Government prefers to give the good to agents of high type rather than to those of low type. The agent's type is her private information. In order to screen the agents government has to introduce red tape.

In this paper, we use a similar framework but there are two major distinctions. First, in Banerjee (1997), red tape is a pure cost imposed upon the agent. The bureaucrat can then use red tape to screen the agent's type. Another screening device is prices but as long as agents are cash-constrained, red tape is more effective. This model can therefore be applied to any means of non-monetary harassment which bureaucrats are authorized to use (e.g. queues). In this paper, we will study the case where the red tape is informative per se. A simple point that may justify the informativeness of the red tape is that red tape is essentially related to the social value of the good to be allocated. While applying for a welfare benefit from the government, an agent is usually requested to submit a document certifying her low income rather than to do some physical exercises or to take a course in economics though all are costly to the agent and can be used for screening.[1]

Second, we distinguish between private and social values of the good. This distinction may come from either externalities or agent-specific costs of provision of the good. Both are quite important for many government activities: we have in mind licenses, passports and visas, product quality certificates, targeted public transfers and welfare benefits, government contracts, and provision of certain public goods.[2] In the case where social values are negatively correlated with private values, neither prices nor any uninformative cost mechanisms can help screening the agents even if there are no cash constraints. The agents eligible for the good are willing to pay less (in both monetary and non-monetary means) than the agents who are not eligible. The situation can be improved with informative red tape.[3] There arises a trade-off between informativeness and cost of red tape. This trade-off can bring about imperfect screening of agents in the social optimum, when all types of agents apply for the good and get it with positive (though may be different) probabilities.

The informativeness of red tape may explain why red tape is considered to be a lesser evil than corruption. Although red tape may be very costly for economic agents,[4] nobody suggests to eliminate the red tape altogether: the socially optimal level of red tape is not trivial. The problem is that self-interested bureaucrats tend to overproduce red tape relative to the social optimum.

The model of informative red tape helps to study emergence and interaction of such kinds of corruption as ex ante corruption and ex post corruption in a principal-bureaucrat-agent hierarchy. Like in Tirole (1986), ex post corruption occurs when the bureaucrat finds out that the agent is of a socially undesirable type but agrees not to report this fact to the principal. (Shleifer and Vishny (1993) refer to this kind of corruption as 'corruption with theft'). Ex ante corruption emerges due to the bureaucrat's discretion to choose the level of red tape. The bureaucrat offers the agent to reduce red tape ex ante in exchange for a bribe. Anticipating the agent's willingness to pay, the bureaucrat threatens a very high level of red tape but then negotiates it down to a lower level. This type of corruption is also similar to ones analyzed in literature ('corruption without theft' in Shleifer and Vishny (1993)) but we show that the informative red

[1] There is an extensive literature on using costly signals for separating agents that began with Spence (1973).

[2] It may also apply to some activities in the private sector such as university admissions and appointments, and selection of papers for publication in academic journals.

[3] Wilson (1989) argues that complicated red tape is created because the society has a compassion for people who, under simpler mechanisms, would not get what the society thinks they deserve. Also, Bardhan (1997) emphasizes that deregulation may be socially harmful since regulation is usually introduced in order to achieve certain social objectives.

[4] See EBRD (1999), p.124 for estimating the cost of the 'time tax' imposed on managers by bureaucrats. Surveys of managers show that the time tax (as a share of senior managers' time) is usually higher than the 'bribe tax' (bribes as a share of firms' revenues).

tape model contains new insights on how these two kinds of corruption interact in equilibrium. One of the most striking results is that although the principal can provide the bureaucrat with such incentives that ex post collusion does not occur, the very threat of ex post collusion constrains the principal's ability to implement the social optimum.[5]

This result goes against a common claim that corruption may be 'good for growth' since it relaxes rigidity of bureaucracy (Huntington (1969)).[6] The major problem with this claim is that it does not discuss where the rigidity comes from. Either rigidity is socially wasteful and then it is not clear who has introduced it and why it is not abolished by legal means. Or rigidity is socially desirable and corruption is therefore only good for the colluding parties but harmful for the whole society. Our model shows that once rigidity (red tape) is endogenous this 'grease-in-the-wheel' arguments may fall apart. Potential corruption may result in overproduction of red tape. It is important to mention that principal can cope with either ex post or ex ante corruption but not with both. It is only in the presence of ex ante corruption, the threat of ex post corruption breeds red tape.[7]

The model therefore predicts a positive correlation between corruption and red tape: threat of corruption leads to excessive red tape. It is hard to test the implications of the model: by its very nature, corruption is not easy to measure and so is red tape. The few datasets on corruption include ones discussed in Bardhan (1997) and Kaufmann et al. (1999). The common approach to measuring corruption and red tape is to conduct a survey of experts or international businesses who can compare levels of corruption and red tape in different countries.[8] These data show (see Kaufman et al. (1999) and LaPorta et al. (1999)) that red tape and corruption are highly correlated.[9]

[5]See Tirole (1992) for a discussion of implications of potential collusion for the mechanism design.

[6]Notice that there is no convincing empirical evidence supporting this claim. Moreover, the country-level evidence shows that corruption is *bad* for growth (Mauro (1995)).

[7]Since red tape and corruption are jointly determined in our model, the model helps to reconcile two common views: 'regulation breeds corruption' (Bardhan (1997)) and 'when rules can be used to extract bribes more rules will be created' (Tanzi (1998)). Notice that we study the equilibrium level of red tape rather than the prohibitively high level that is announced to extort bribes but then negotiated down in exchange of illicit payments.

[8]There is also an alternative approach to measuring red tape. Parsons (1991) looks at the data on targeted public transfers. He estimates the impact of complexity of screening procedures in the U.S. welfare system on individuals' willingness to apply for benefits and therefore indirectly measures the cost of red tape.

[9]The data used in La Porta et al. (1999) show 86 per cent correlation between "corruption' and 'bureaucratic delays' (significant at 1% level). Kaufmann et al. (1999) aggregate indicators from all available sources (incl. the one used by LaPorta et al. (1999)). Their measures of 'government effectiveness' and 'graft' have a 92% correlation (significant at 0.1% level). On the other hand as both La Porta et al. (1999) and Hellman et al. (1999) argue there may be an identification problem: both variables are determined by something else (e.g. legal origin or state capture in the country).

Our work follows Banerjee (1997) and Dewatripont and Tirole (1999) in an attempt to understand what is so special about government bureaucracy that makes it so notoriously inefficient. Dewatripont and Tirole (1999) point at multiple objectives and fuzzy missions of government agencies. Banerjee (1997) studies allocation of a scarce good with cash constraints. We look at the case with externalities where market mechanisms (such as auctions) do not work and a welfare-maximizing government has to use red tape which may result in inefficiencies in the presence of ex ante and ex post corruption.

The structure of the paper is as follows. In Section 2 we build the basic model of a three-tier hierarchy with informative red tape and ex ante and ex post corruption. We describe agents' payoffs and timing, and make assumptions about parameter values that rule out trivial cases. In Section 3 we study how the threat of ex ante and ex post corruption influences the equilibrium level of red tape. We start with the benchmark case of the socially optimal level of red tape (Subsection 3.1). Then, in Subsection 3.2, we show that principal can implement the social optimum if there is no threat of ex post corruption. In Subsection 3.3 we introduce a possibility of ex post corruption and solve for equilibrium. Subsection 3.4 compares the outcomes with and without each type of corruption. Section 4 discusses robustness of the results and potential extensions of the model. In Subsection 4.1 we discuss how to endogenize propensity for corruption. In Subsection 4.2 we study what happens if the principal can use multiple contracts to screen types. Subsection 4.3 discusses robustness of the results to changes in distribution of bargaining power. Section 5 concludes.

# 2 A model of informative red tape

## 2.1 The setting

We will consider a hierarchy of a principal P (government), an agent A (customer), and a bureaucrat B (government official) who supervises the agent and reports to the principal. There is also another agent T ('taxpayers' or 'treasury'). In the model, T does not make any decisions but P takes the taxpayers' interest into account along with those of other agents. B and A are selfish while P maximizes social welfare.

Principal can supply a unit of a good (or a service) to the agent. The agent's valuation of the good is $\theta > 0$. The cost of provision of the good (borne by the taxpayers) is $c$. The net social benefit of provision is therefore $v = \theta - c$. The agent can be either of a 'good' type characterized by private value $\theta^g$ and cost of provision $c^g < \theta^g$ or of a 'bad' one with $\theta^b$ and $c^b > \theta^b$, correspondingly. Since $v^b < 0 < v^g$, the first best is to provide the good to the good type and not to provide to the bad one.[10]

[10]The framework is quite general and covers surprisingly many governmental activities: pro-

The prior distribution of types is common knowledge: with probability $\pi \in (0,1)$ the agent is of the bad type and with probability $1-\pi$ she is of the good type. The principal cannot distinguish the types ex ante and may want to use red tape. We model the red tape as a questionnaire that consists of a number of tests. The greater the amount of red tape (the number of tests), the more it costs the agent, but the more likely the agent's type is revealed. The red tape is measured in terms of its cost to the agent $r \geq 0$. The outcome $\rho$ of the tests is either 'pass' $\rho = 1$ or 'fail' $\rho = 0$. The good type passes any number of tests with probability 1. The bad type passes $r$ tests with probability $1 - I(r)$. Here $I(r)$ is a measure of informativeness of the red tape. We will assume that $I(r)$ is an increasing concave twice differentiable function: $I'(r) > 0$, $I''(r) < 0$, $I(0) = 0$, and $I(\infty) \leq 1$. The share of the good types among those who pass $r > 0$ tests $\frac{1-\pi}{\pi(1-I(r))+(1-\pi)}$ is greater than the share among the whole population $1-\pi$. Moreover, the former share increases with $I(r)$.[11]

In addition to the red tape, the principal can use price $p \geq 0$ and application fee $t \geq 0$. The price is paid whenever the agent *gets* the good and the application fee is paid whenever she *applies* for the good.[12] Therefore the agent is offered a mechanism $(r, p, t, \sigma(\cdot))$ where $r$ is the amount of red tape, $p$ is the price, $t$ is the application fee, and $\sigma : \{0,1\} \to \{0,1\}$ is the provision rule contingent on the outcome of the tests $\rho$. There are $2 * 2 = 4$ possible provision rules $\sigma(\cdot)$ : (i) provide to everyone $\sigma(\rho) = 1$; (ii) provide to nobody $\sigma(\rho) = 0$; (iii) provide only to those who pass the test $\sigma(\rho) = \rho$, and (iv) provide only to those who fail the test $\sigma(\rho) = 1 - \rho$. As we will show below, it makes sense to use any non-trivial amount of red tape only if $\sigma(\rho) = \rho$.

P is not competent at administering the red tape and hires a bureaucrat B. P cannot observe the level of red tape. However, P can observe ex post the distribution of types who got the good or who got rejected. The technology for

---

vision of goods and services with externalities on the third parties, assignment of procurement contracts, admitting students to universities, recruitment of public servants, etc. Actually, it also describes the process of reviewing papers for publication in academic journals. The setting is basically equivalent to that of Tirole (1996) and Hauk and Saez Marti (1998). There the principal is better-off assigning task 1 to good type and task 2 to the bad type. In our setting, the task 1 is 'providing the good' while the task 2 is 'not providing the good'.

[11] In this paper, we allow only for the type II error. One can imagine a more general setting where both type I and type II errors may occur: the good type passes $r$ tests with probability $P^g(r)$ while the bad type passes $r$ tests with probability $P^b(r)$. The share of the bad type among those who have past $r$ tests is $\frac{\pi P^b(r)}{\pi P^b(r)+(1-\pi)P^g(r)}$ which is smaller than $\pi$ whenever $P^b(r) < P^g(r)$. The informativeness (i.e. the difference between $\frac{\pi P^b(r)}{\pi P^b(r)+(1-\pi)P^g(r)}$ and $\pi$) increases with $P^g(r)/P^b(r)$. In particular, $I(r) = 1 - P^b(r)/P^g(r)$ is a reasonable measure of informativeness.

[12] In practice, application fee is used rather rarely (except for the cases were application process is costly and the application fee is to cover the processing costs). The reason for that may be the 'presumed innocence' constraint. If the good type may fail to pass the test with some probability (see the previous footnote), it may happen that some good types end up paying but not receiving the good that they are supposed to get.

measuring the distribution ex post is as follows. Government costlessly checks a negligible number of agents and extrapolates the share of bad types that got the good. Also, the number of rejected applications is known ex post.[13] In this setting, bureaucrat's salary can be made contingent on ex post distribution. In the meanwhile, agents of the bad type who got through the red tape will keep the good with a probability very close to 1.[14] Therefore P can offer B a contract $(s^b, s^g, s^0)$ so that B is paid $s^b$ when the bad type gets the good, $s^g$ when the good type does and $s^0$ when nobody does. Apparently, only differential rewards $s^g - s^0$ (bonus for letting the good type get the good) and $s^0 - s^b$ (punishment for letting the bad type get the good) matter. The base salary $s^0$ is determined by B's reservation utility (for simplicity normalized to zero).[15] To shorten the proofs, we will only consider the monotonic rules $s^b \leq s^0 \leq s^g$.

Let us introduce some notation. Denote $R^g$ and $R^b$ maximum participation levels of red tape at $p = t = 0$ and $\sigma(\rho) = \rho$ for both types:

$$\theta^g = R^g, \;\; \theta^b(1 - I(R^b)) = R^b.$$

Introduce $\overline{r}$ to be a solution to $\theta^g - r = \theta^b(1 - I(r)) - r$ :

$$1 - I(\overline{r}) = \theta^g/\theta^b. \tag{1}$$

It is easy to check that $\overline{r} \geq R^b$ if and only if $R^b \geq R^g$. Thus it is either (i) $\overline{r} < R^b < R^g$ or (ii) $R^g \leq R^b \leq \overline{r}$. Let us also introduce

$$r^* = \arg\max_{r \in [0,\overline{r}]} \left\{ \pi |v^b| I(r) - r \right\}. \tag{2}$$

Denote $\mathtt{1}(x)$ the indicator function which takes value of 1 whenever statement $x$ is true and that of 0 otherwise. Also, $[y]_+ = \max\{y, 0\}$.

## 2.2 Timing

The timing is as in the Figure 1. First, A learns her type. Then principal offers a contract $(s^b, s^g, s^0)$ to B. P also chooses price $p$, application fee $t$ and provision rule $\sigma(\cdot)$. B decides whether to take the contract. If B takes the contract, he

[13]University admission may be a good example. Neither the admission committee (B) nor the trustees (P) can observe the quality of applicants (A) ex ante. On the other hand, the quality is observed ex post (for example, via placement of graduates) and P can formally or informally reward B for admitting better applicants.

[14]Principal can check a small representative sample but not the whole population. Therefore for every given agent, the probability of being caught is infinitesimal. If principal observed types of the whole population ex post and could take the good back from the bad type, red tape would not be needed.

[15]We assume that the payments are costless: a dollar paid to B costs T presicely one dollar. The contract $(s^b, s^g, s^0)$ may include monetary payments or fines, non-pecuniary benefits but does not include imprisonment.

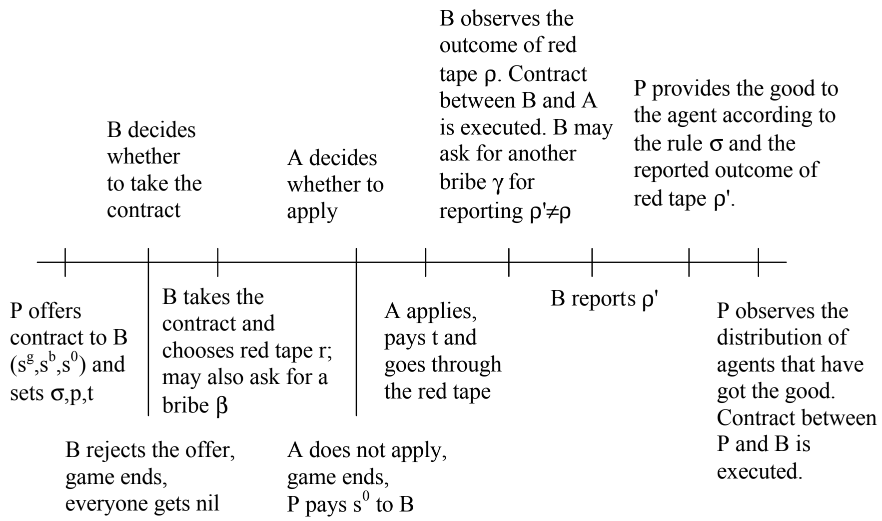


Figure 1: Timing

sets the level of red tape $r$. B may also ask A for a bribe $\beta$ for lowering the red tape: B may threaten A with a high level of red tape and offer a lower level for a bribe. Given $r, p, t$ and $\sigma(\cdot)$, agent decides whether to apply for the good. If she quits, the game ends and everyone gets zero. If she applies, she pays the application fee $t$ (to the taxpayers) and bribe $\beta$ to B. Then she undergoes the tests and bears the cost $r$. B learns the outcome $\rho$ of the tests and reports $\rho'$ to the principal. B may misreport the outcome ($\rho' \neq \rho$) in exchange for another bribe $\gamma$. B cannot fabricate the evidence of A's failure; she can only conceal the evidence, i.e. $\rho' \geq \rho$.[16] Thus if $\rho = 1$ then B can only report $\rho' = 1$. If $\rho = 0$, B may report $\rho' = 1$ or $\rho' = 0$.

Given the reported outcome of tests $\rho'$, P executes the contract. If $\sigma(\rho') = 1$, P provides the good to the agent, A gets the good and pays T the price $p$. If $\sigma(\rho') = 0$, A does not get the good. Then P observes the type of the agent who has got the good and T pays B his salary $s^i$, $i = 0, g, b$.

We will compare the outcomes with and without corruption. If there is no ex ante corruption then $\beta = 0$. If there is no ex post corruption then $\gamma = 0$. If corruption is allowed then the bribes $\beta$ and $\gamma$ are determined through bargaining between B and A. For the simplicity's sake, we assign all bargaining power to B: B makes A a take-it-or-leave-it offer. We assume that cost of transfer is negligible but given other things equal, smaller amount of gross transfers increases welfare.

The principal maximizes social welfare; everyone else is selfish. The agent

[16]This follows the framework in Tirole (1986). If B could both fabricate and conceal the evidence, the red tape would not make sence: reporting would be fully in B's discretion. For analysis of this case, see Polinsky and Shavell (2001).

maximizes her expected payoff net of cost of red tape. If the good type applies she gets

$$U^g = (\theta^g - p)\sigma(1) - t - r - \beta. \tag{3}$$

If the bad type applies she gets

$$U^b = (\theta^b - p)\left[\sigma(1)(1 - I(r)) + \sigma(\rho')I(r)\right] - t - r - \beta - \gamma I(r). \tag{4}$$

The bureaucrat maximizes her expected salary plus bribes taking into account that agents may decide not to apply for the good:

$$\begin{aligned} U^B &= s^0 + \Big((s^g - s^0)\sigma(1) + \beta\Big)(1-\pi)\mathbb{1}(U^g \geq 0) \\ &+ \Big((s^b - s^0)\left[\sigma(1)(1 - I(r)) + \sigma(\rho')I(r)\right] + \beta + \gamma I(r)\Big)\pi\mathbb{1}(U^b > 0) \end{aligned} \tag{5}$$

The first term represents B's base salary, the second term is B's bonus if A is of the good type and A applies for the good. The last term is B's bonus (or penalty) if A is of the bad type, A applies and gets the good.

The taxpayers' expected payoff is as follows:

$$\begin{aligned} U^T &= -s^0 + \Big\{t + \Big(p - c^g - (s^g - s^0)\Big)\sigma(1)\Big\}(1-\pi)\mathbb{1}(U^g \geq 0) + \\ &+ \Big\{t + \Big(p - c^b - (s^b - s^0)\Big)\left[\sigma(1)(1 - I(r)) + \sigma(\rho')I(r)\right]\Big\}\pi\mathbb{1}(U^b > 0) \end{aligned} \tag{6}$$

The principal maximizes the social welfare $W = U^T + U^B + (1-\pi)[U^g]_+ + \pi[U^b]_+$. Adding up (3)-(6) we obtain

$$W = \left(v^g\sigma(1) - r\right)(1-\pi)\mathbb{1}(U^g \geq 0) + \tag{7}$$

$$+ \Big(v^b\left[\sigma(1)(1 - I(r)) + \sigma(\rho')I(r)\right] - r\Big)\pi\mathbb{1}(U^b > 0) \tag{8}$$

The first term describes the social welfare if A is of the good type (this happens with probability $1 - \pi$). If A expects to get a negative surplus $U^g < 0$ she does not apply; the welfare is zero. If A's expected utility is positive then the agent applies and passes the test $\rho = 1$. Since the monetary transfers cancel out, the welfare equals $v^g\sigma(1) - r$. Similarly, the second term represents the welfare if the agent is of the bad type and chooses to apply.

It is important to emphasize the difference between ex ante and ex post corruption. Ex post corruption ('corruption with theft') is essentially collusion between the bad type and the bureaucrat, while ex ante corruption occurs when B has no information on the agent.[17] Ex ante, both good and bad types may give a bribe

[17] Policymakers have understood the distinction between ex post and ex ante corruption a long ago. As discussed in Bardhan (1997) and Elliott (1997), U.S. anti-corruption legislation has been stricter on prosecuting the ex post corruption when the bribe does influence the decision made by the bureaucrat rather than the ex ante one when the decision is not affected and bribe simply reduces the red tape.

to reduce the level of red tape. This kind of corruption seems innocent: unlike ex post corruption, there is no 'theft' from the public. On the other hand, ex ante corruption also involves changes in real terms: B and A choose the level of red tape best for their joint well-being which may well differ from the socially optimal one. Through setting B's incentives P may try to manipulate the outcome of ex ante collusion in order to achieve the level of red tape that is best for public.[18]

## 2.3 Assumptions

In order to concentrate on the most interesting case, we shall make the following assumptions.

*Assumption A1.* The pooling equilibrium without red tape yields positive social welfare

$$W^0 = (1-\pi)v^g + \pi v^b > 0. \tag{9}$$

This assumption implies that social welfare is greater when the good is provided to everyone rather than is not provided at all.

*Assumption A2.* Marginal benefit of a small amount of red tape is greater than its cost: $\pi|v^b|I'(0) > 1$.

This makes sure that the red tape is not trivial in the social optimum.

*Assumption A3.* Agents *cannot be* screened through prices only but *can be* screened via red tape : $\theta^g < \theta^b < \theta^g/(1 - I(\theta^g))$.

Assumption A3 implies that the bad type will always be willing to pay more for the good than the good type. Therefore if the good is sold at a given price $p$ or auctioned off, the first best cannot be achieved. On the other hand, red tape can separate the agents. A3 implies $\overline{r} < R^b < R^g$. Hence, if $r \in (R^b, R^g)$, and $p = t = 0$, $\sigma(\rho) = \rho$, then the good type will apply for the good while the bad type will not.

A3 rules out certain potential applications of the model. Indeed, if we consider the case of providing a good that incurs no externalities with its cost same for both types $c^g = c^b$, then $\theta^g - \theta^b = v^g - v^b > 0$. An example is a decision on allocation of a good that is in limited supply (e.g. spectrum band license). The cost of provision is the shadow price or Lagrange multiplier to the resource constraint $c^g = c^b = \theta^b$ which is the same for both types so that $v^b = 0 < v^g = \theta^g - \theta^b$ (see Banerjee (1997) for a comprehensive analysis of this case). If there are no cash constraints, the case $\theta^g > \theta^b$ is rather trivial: government should only announce that the good can be purchased by any agent at price $p \in (\theta^b, \theta^g)$ and no red tape is needed.

The applications with $\theta^g < \theta^b$ are usually related either to externalities or to differentiated cost of provision such issuing licenses, passports, visas etc. E.g. consider issuing licences to firms some of which want to engage in legal business

[18] We assume that P knows B's propensity for ex post and ex ante corruption. See Carrillo (2000) and Kofman and Lawarree (1996) for analysis of the case of asymmetric information.

and some are to engage in semi-criminal activities or ones endangering environment. The private return on the socially undesirable activities may be much higher. The same logic can be applied to issuing passports or visas, awarding certain government contracts. Another application is privatization of public property under imperfect capital markets. Suppose that government wants to allocate the property to the most efficient owner (the 'good' type) but the latter does not have enough cash and cost of borrowing is rather high. On the other hand it may occur that a less efficient owner (the 'bad' type) has enough cash to buy the property. In this case the good type's valuation is NPV of future revenues minus cost of borrowing which may be less than the bad type's NPV.

The second inequality in A3 makes sure that red tape can indeed separate the agents. Given $\theta^g < \theta^b$, the second inequality requires red tape to be sufficiently informative or private values to be high enough (for a given $I(\cdot)$, A3 becomes true if both $\theta^g$ and $\theta^b$ are multiplied by a sufficiently large scalar factor).

*Assumption A4.* The share of the bad type is sufficiently low $\pi < \widehat{\pi}$, where $\widehat{\pi}$ is such that

$$(1-\widehat{\pi})\overline{r} = \min_{r\in[0,\overline{r}]} \left\{ r + \widehat{\pi}|v^b|(1-I(r)) \right\}$$

The role of this assumption will become clear in the next section. We will show that implementing the social optimum is hard whenever A4 holds and is rather easy otherwise.

# 3 Solving the model

In this Section we will solve for equilibrium. First, we will find the social optimum, i.e. the outcome that is achieved when B maximizes social welfare. Then we will check whether this social optimum can be implemented if B is selfish, i.e. maximizes (5). Essentially, the model has two layers of agency problems. First, there is an adverse selection problem with the bad type pretending to be the good one. Second, there is a moral hazard problem with B choosing the red tape $r$ and reporting its outcome $\rho$. B may want to choose a wrong level of the red tape because he may not internalize the full social cost of red tape. B may also want to misreport $\rho$ since he can extort bribes from the bad type ex post. The main goal of the paper is to study what level of red tape can be implemented by the benevolent principal if she can only write contracts on ex post distribution rather than on level and outcome of red tape directly.

We will study how the presence of ex ante and ex post corruption influences P's ability to achieve the social optimum. We will solve the model for four scenarios (with and without ex ante and ex post corruption). Each scenario seems plausible: either type of corruption requires specific assumptions on enforceability of illicit contracts that may or may not hold in certain environments. Ex post collusion is rather straightforward: B agrees to destroy the evidence (or hand it over to A) in exchange for a bribe. In many cases, though, B cannot credibly destroy the evidence: e.g. there may exist photocopies that are enough to implicate A. Ex ante corruption also requires commitment: B sets the (illegally low) level of red tape for a payment from A. It may be the case that after getting the bribe, B still sets a high $r$. Since the initial threat was a high $r$ and there is no legal mechanism to enforce the illicit contract, A may be unable to punish B for withdrawing his promise.

## 3.1 Social optimum

In order to have a benchmark, we shall find the socially optimal level of red tape. Red tape is informative but it is also costly therefore the choice of red tape is not at all trivial. If the agent's type were common knowledge, no red tape would be needed $r = 0$; the good type would be given the good and the bad one would not. The first best level of social welfare would be $W^{FB} = (1-\pi)v^g$. In what follows, we will assume that agent's type is known only to the agent and will refer to the resulting *second best* as the social optimum. Essentially, we will check how well P could cope with the adverse selection problem if there were no moral hazard.

In this Subsection, we assume that P herself sets the level of red tape and observes the outcome of the tests.[19] The principal chooses the quadruple $(r, p, t, \sigma(\cdot))$ to maximize (8) subject to (3)-(4), $\beta = \gamma = 0$, $\rho' = \rho$.

**Proposition 1** *Under A1-A3, the social optimum is as follows.*

1. *If $\pi \geq \widehat{\pi}$ (A4 does not hold) then P sets $r = \overline{r}$, $p = 0$, $t = \theta^g - \overline{r}$ and $\sigma(\rho) = \rho$. The good type applies, while the bad type does not. The social welfare is $\overline{W} = (1-\pi)(v^g - \overline{r})$.*

2. *If $\pi < \widehat{\pi}$ (A4 holds), then P sets $r = r^*$ (see (2)), $p = 0, t = 0$ and $\sigma(\rho) = \rho$. Both types apply. The welfare is*

$$W^* = (1-\pi)v^g + \pi v^b(1 - I(r^*)) - r^*.$$

The Proposition explains the role of Assumption A4 (which is formally equivalent $W^* > \overline{W}$). If the bad type is sufficiently common, then P is better-off deterring the bad type even if this imposes high cost on the good type (it takes at least $r = \overline{r}$ to make sure that the bad type does not apply). If the bad type is rather rare, then P prefers to allow both types to apply since the social loss from providing the good to some of the bad types is relatively low.

Thus, under A1-A4, the social optimum has the following features. First, price and application fee do not help to distinguish agents. Second, the agents are not

[19] Alternatively, one can assume that B never takes bribes $\beta \equiv \gamma \equiv 0$ and acts in the interest of P i.e. always sets the level red tape that the principal wants to implement and truthfully reports the evidence $\rho' = \rho$.

fully discriminated ex ante i.e. the fact of applying does not tell the agent's type for sure. The types are partially screened: the good type gets the good with a higher probability than the bad one. The share of the good type among those who get the good is $(1-\pi)/(1-\pi I(r^*)) > 1-\pi$. Third, the good type receives lower rent $U^g < U^b$. Indeed, $r^* < \overline{r}$ implies $\theta^g - p - t - r^* \leq (\theta^b - p)(1 - I(r^*)) - t - r^*$ for any $p, t \geq 0$.[20]

## 3.2 Implementing the social optimum with ex ante corruption

In order to implement the social optimum, P needs to overcome two moral hazard problems. First, B should have incentives to choose the right level of red tape. Second, B should truthfully report the outcome of the tests. In this Subsection we will assume that B cannot engage in ex post corruption so only the first problem remains.

The social optimum can be easily implemented if there is no corruption. P chooses $s^b, s^g, s^0, p, t$ and $\sigma(\cdot)$ to maximize (8), predicting B's and A's best responses given $\beta = \gamma = 0$. Under given $s^b, s^g, s^0, p, t, \sigma(\cdot)$ and $\beta = \gamma = 0$ B chooses $r$ and $\rho'$ expecting A to maximize $U^i$, $i = g, b$. P should set $s^b = s^g = s^0$, $p = t = 0$ and $\sigma(\rho) = \rho$. Since B is indifferent which level of red tape to choose, she chooses the social optimal $r$ and truthfully reports the outcome $\rho' = \rho$.[21]

Let us turn to the case where B can extort bribes ex ante but cannot engage in the ex post collusion (i.e. $\gamma = 0$). The bureaucrat threatens the agent with a high level of red tape $r^t$ but also offers to lower the red tape down to $r$ in exchange for a bribe $\beta$ to maximize his surplus which is legal rewards plus bribes (5). Since there is no control from the principal, B can threaten a prohibitively high level of the red tape $r^t = \max\{R^g, R^b\}$ at which both types get zero rent (by taking the outside option). The bureaucrat may choose whether (i) to deter one type and extract all the rent from the other one or (ii) let both types in. If both types apply then B extracts all the rent from the type with lower rent leaving certain informational rent to the other type.

Under given $s^b, s^g, s^0, p, t, \sigma(\cdot)$ B chooses $\beta$ and $r$ to maximize (5) where $U^g$ and $U^b$ are as in (3) and (4), and $\gamma = 0, \rho' = \rho$.

**Proposition 2** *Assume that A1-A3 hold. If there is ex ante corruption but no threat of ex post corruption then P can implement the social optimum, i.e. there exist such $\sigma(\cdot), p, t$ and $s^b, s^g, s^0$ that B chooses the socially optimal $r$.*

[20] This is an implication of A4. If the good type got a higher rent, P could sort the bad type out by raising the application fee. In that case, the welfare would not exceed $\overline{W}$ which, according to A4, is below $W^*$.

[21] The flat contract implements the social optimum because the red tape incurs no cost to B. If the red tape were costly to B, P would have to come up with an incentive contract such that B's private marginal cost of the red tape were equal to his private marginal benefit $(s^0 - s^b)\pi I'(r^*)$ in the social optimum.

1. *If A4 does not hold then the social optimum is implemented through setting a sufficiently large punishment* $s^0 - s^b$ *for letting the bad types get the good. P sets* $\sigma(\rho) = \rho$ *and* $p = t = 0$. *Then B chooses* $r = \overline{r}$ *and* $\beta = \theta^g - \overline{r}$.

2. *If A4 holds then P implements the social optimum via the following mechanism. P sets* $\sigma(\rho) = \rho$, $p = t = 0$, $s^g - s^0 = \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$ *and* $s^0 - s^b = |v^b|$. *Then B chooses* $r = r^*$ *and* $\beta = \theta^g - r^*$.

The first part of the Proposition is quite intuitive: if P wants to exclude the bad type, she can simply introduce a high punishment for letting the bad type in (or, alternatively, a high reward $s^0 - s^b$ for excluding the bad type). B deters B from applying by setting $r \geq \overline{r}$. Given that the bad type does not apply, B has all incentives to set the red tape as low as possible since the lower the red tape, the greater the good type's rent that B can expropriate through bribes. Therefore B sets the minimum level of red tape that is sufficient to deter the bad type $r = \overline{r}$.

The second part is more tricky. P wants to implement an equilibrium in which B encourages both types to apply. B gets

$$s^0 + (1 - \pi)(s^g - s^0) + \pi(s^b - s^0)(1 - I(r)) + \min\{\theta^g - r, \theta^b(1 - I(r)) - r\}.$$

The last term represents the bribe $\beta$. Since $r^* < \overline{r}$, the good type's surplus is below the bad type's. Therefore it is the bad type who receives a positive informational rent after paying a bribe. The bribe equals $\beta = \theta^g - r$ . The bureaucrat fully internalizes the cost of red tape: if A's cost of red tape increases by one dollar, B can extort one dollar less in bribes. Therefore if P gives B the incentives aligned with the social value of provision $s^0 - s^b = -v^b$, B chooses $r$ to maximize the social welfare.

If A1-A4 hold, the role of prices and application fees is purely redistributional: it is a non-distortive tax on a bureaucrat paid to the treasury. If P wanted to reduce the bribe $\beta$ by a dollar, P might simply raise the application fee by a dollar hence redistributing income from the bureaucrat to the treasury without any changes in real terms.[22]

## 3.3 The role of ex post corruption

In this Subsection we study the implementation problem if ex post corruption is allowed. Ex post corruption may occur once A fails the test. If the evidence of failure is concealed and $\sigma(\rho) = \rho$, A gets the good and gains $\theta^b$. On the other hand, the bureaucrat gets a lower salary ($s^b$ instead of $s^0$). If collusion increases the joint surplus $\theta^b + s^b > s^0$ then the bribe $\gamma \in [s^0 - s^b, \theta^b]$ redistributes the

[22]Actually, by increasing the application fee up to $\theta^g - r$, P can make sure that the bribe is reduced to zero. Moreover, if P sets $t$ higher than $\theta^g - r$ then $\beta$ becomes negative: B has to pay part of the application fee himself. Notice that allowing for negative bribes does not give B a contractual advantage over P: the constraint $t \geq 0$ is not binding.

gain from A to B so that both B and A benefit from collusion. Since B has all bargaining power, he extorts the bribe of $\gamma = \theta^b$.

The principal can prevent collusion by setting $s^0 - s^b \geq \theta^b$. Indeed, in this case the bureaucrat earns more by reporting the bad type than by accepting the bribe and letting her go.

**Definition 1** *The reward schedule* $(s^b, s^g, s^0)$ *is said to be collusion-proof if* $s^0 - s^b \geq \theta^b$.

If B's salary is not collusion-proof $s^0 - s^b < \theta^b$, B will always report that A passes the test $\rho' = 1$ while the true outcome $\rho$ may be either success or failure. Therefore, $\sigma(0)$ becomes irrelevant, and the provision rule $\sigma(\rho) = \rho$ ('provide the good only to those who pass the test') performs as badly as the rule $\sigma(\rho) = 1$ ('provide to everyone'). The latter rule does not take advantage of the red tape's informativeness and never achieves welfare greater than $W^0$ (9). If P wants to do better than that, she should offer a collusion-proof contract.[23] The problem is that collusion-proofness may constrain the set of mechanisms which P can use. If the bad type is sufficiently common (A4 does not hold), P wants to deter the bad type from applying. Therefore P may set the punishment $s^0 - s^b$ as high as possible; collusion-proofness is not binding. On the other hand, under A1-A4, P would like to set $s^0 - s^b = |v^b|$. If $|v^b|$ is below $\theta^b$, collusion-proofness becomes a binding constraint for the mechanism design problem; the social optimum cannot be implemented.

**Proposition 3** *Suppose that assumptions A1-A4 hold and both ex ante and ex post corruption is allowed. P can achieve the social optimum if and only if* $\theta^b \leq |v^b|$. *If A1-A3 hold but A4 does not, then P implements the social optimum via setting* $s^0 - s^b$ *high enough.*

The threat of collusion limits P's ability to implement the social optimum. P wants to reduce the excessive red tape; the cost of red tape is high while benefit is low: bad types are relatively rare. To provide B with incentives for lowering the red tape, P does not severely punish B for letting the bad type in. But the reward schedule with low $s^0 - s^b$ may be not collusion-proof.[24]

[23] This is a well-known argument for increasing public servants' incentives in order to prevent corruption (Becker and Stigler (1974)). Most governments rather tend to set the wages quite low in the public sector but try to provide high-powered incentives via back-logged wages and generous pensions. The anecdotal evidence is for higher compensations (Klitgaard (1988)) while the empirical evidence shows that the effect is present but weak (Rijchenghen and Weder (1997)). Modern contract theory may also suggest that high wages can eliminate corruption but may create other problems if P wants B to do multiple tasks. In our simple model, the principal is able to eliminate collusion by setting $s^0 - s^b$ high enough and chooses to do so.

[24] In some sense, this result may be explained by the multi-tasking problem: P has only one tool $s^0 - s^b$ for providing incentives for two activities: choosing the right level of red tape $r$ and reporting the outcome of red tape $\rho$. If A4 does not hold, then there is no conflict between the two activities: very high $s^0 - s^b$ rewards both.

If $\theta^b > |v^b|$, the maximum welfare that P can achieve, belongs to the interval $[\max\{\overline{W}, W^0\}, W^*)$. Indeed, P can easily implement $\overline{W}$ via raising $s^0 - s^b$. On the other hand, $W^0$ can be implemented via setting $\sigma(\rho) = 1$ (this is virtually equivalent to firing the bureaucrat).

It is interesting that the potential ex post collusion only hurts the principal's ability to implement the social optimum in the presence of ex ante corruption. Indeed, if there is no ex ante corruption, B does not internalize the cost of red tape and tends to overproduce the red tape. However, P can make sure that this does not happen via raising the application fee. If there is no ex ante corruption, high application fee decreases B's willingness to generate too much red tape: if there is too much red tape, the good type does not apply and B's salary never exceeds $s^0$. This mechanism does not work in the presence of ex ante corruption: by colluding with A, B can make up for a higher application fee with a lower bribe.

**Proposition 4** *Assume A1-A3. Suppose that ex post corruption is allowed but ex ante corruption is not.*

1. *If A4 does not hold, then P implements the social optimum via setting sufficiently high* $s^0 - s^b$.
2. *If A4 holds, the principal implements the social optimum via the following mechanism:* $\sigma(\rho) = \rho$, $p = 0$, $t = \theta^g - \overline{r}$, $s^g - s^0 > 0$, $s^0 - s^b \geq \theta^b$. *Then the bureaucrat chooses* $r = r^*$ *and truthfully reports the outcome of the tests* $\rho' = \rho$.

Again, there is no ex post collusion in equilibrium. The threat of collusion does not prevent the principal from implementing the social optimum. Principal is able to confine B's willingness to overproduce red tape through setting high application fee.

## 3.4 Discussion

Summing up the Propositions 1-4 we see that ex ante corruption and ex post corruption may limit the principal's ability to implement the socially optimal level of the red tape. If the bad type is common $\pi > \widehat{\pi}$ then P can always implement the social optimum (which is to exclude the bad type) whenever B is honest or corrupt. If the bad type is rather rare $\pi < \widehat{\pi}$ then P can only implement the social optimum if there is no threat of ex post corruption or there is no threat of ex ante corruption.

It is interesting that P can implement the social optimum if there is only potential ex ante corruption or only potential ex post corruption. However, when both types of corruption are present, P may have a problem: the tools that P would use to contain the ex post corruption would give wrong incentives to a bureaucrat who takes bribes ex ante. Vice versa, the mechanisms that implement the social optimum with the bureaucrat who may be involved in ex ante but not ex post corruption perform rather badly when ex post corruption is introduced. The most striking part is that P can actually eliminate corruption both ex post and ex ante but the very threat of corruption prevents the principal from achieving the social optimum.

Another important result is that corruption leads to overproduction of the red tape rather than underproduction of it. Indeed, if $\theta^b > |v^b|$, then offering a collusion-proof contract always gives the bureaucrat too much incentives to catch the bad type so that the red tape is greater than the socially optimal level.[25] This runs against a commonly held view best stated by Huntington (1968):

> "... the only thing worse than a society with a rigid, overcentralized, dishonest bureaucracy is one with a rigid, overcentralized, honest bureaucracy".

Why do we get the opposite result? First, we assume that the bureaucrat's incentives are set by a rational benevolent principal who is also rather powerful, i.e. she has a number of tools to make sure that an honest (or partially honest) bureaucrat gets as close to the social optimum as possible. In the real life, B's incentives may be set by yet another self-interested bureaucrat or a politician captured by an interest group. Second, and more important distinction is that we endogenize the rigidity. It is obvious that if the official level of red tape is too high, bribery reduces it to a more reasonable level. Our model claims that if there were no threat of corruption, the official level would be reasonable in the first place.

The intuition for the overproduction of red tape by a (potentially) corrupt bureaucrat is as follows: the potentially corrupt bureaucrat is ex ante interested in more red tape than socially needed since he wants to extort bribes ex post. P can make sure that ex post collusion does not occur but only through offering high incentives to report the bad type ex post which unfortunately distorts the ex ante incentives for setting the red tape.

# 4 Extensions

The models of collusion in a three-tier hierarchy are generally very complex. In order to build a tractable model, we deliberately introduced a few shortcuts. In this Section, we will check whether the results are robust to replacing them with more realistic assumptions.

[25] An important exception is of course the case $\overline{W} < W^0$. In this case, if $\theta^b$ is much higher than $|v^b|$ and the inefficiency is too large, P prefers to forget about red tape and gives the good to everybody $\sigma(\rho) = 1$ i.e. effectively fires the bureaucrat. In this case there is no red tape at all $r = 0$.

## 4.1 Endogenizing propensity for corruption

An important drawback of the analysis in Section 3 is that we study equilibria under exogenous propensity for ex post and ex ante corruption. For a given propensity for corruption, we solve for red tape and the level of corruption, but the question remains what determines the propensity. In this Subsection we will concentrate on the following issue: if B benefits from colluding with the bad type (i.e. $\theta^b > s^0 - s^b$), what may prevent him from collusion? The analysis of propensity for ex ante corruption would be just the same.

The most important factor of propensity for corruption is enforceability of illicit contracts. B and A cannot take the collusive contract to the court of law. Thus A cannot be always sure that B reports what A has paid for. For example, if A pays B *before* the report $\rho'$ is filed, then the subgame perfect behavior would be to report the true state $\rho' = 0$ and get the legal reward $s^0 - s^b$ on top of the bribe $\gamma$. On the other hand, if A pays to B *after* the report is filed, then A does not need to pay the bribe. The collusion may therefore occur only if there is some private enforcement mechanism (such as repeated interaction, reputation, organized crime, etc.) or the nature of evidence is such that it can be handed over by B to A in exchange for the bribe (and cannot be duplicated).

The other explanation of propensity for corruption is B's personal non-pecuniary cost of being involved in corrupt behavior. Let us denote this cost $D$. If $D > \theta^b$, then there is no danger of ex post collusion and it is easy to achieve the social optimum. The cost of being corrupt may include personal costs of committing an immoral act or violating a social norm (i.e. acting in a non-conformist way). The cost would also increase if the principal sets up a supervisory unit that would be checking all potentially corrupt bureaucrats. Then $D$ would include expected disutility of being caught (cost of potential punishment times the probability of being caught).

To endogenize the cost $D$, one should build a model with several bureaucrats. There is a number of theories where corruption emerges due to the numerical externality, either of static or of dynamic nature. The moral cost is explained by a culture prevailing in the society which in turn depends on educational choices of parents. As Hauk and Saez Marti (1998) show, the greater number of parents teach their children to behave honestly, the higher is the individual return to honest behavior. Therefore there may be multiple equilibria: one without corruption (high $D$) and one with high propensity for corruption (low $D$). The same story works for social norms and a preference for conformity: the more bureaucrats are corrupt, the less costly it is for each of them to lull their conscience. Another dynamic externality is due to imperfect observability of individual track record. Tirole (1996) builds a model where only collective reputation can be observed. There can also be multiple equilibria with and without corruption. Indeed, if many group members are corrupt it does not pay off to be honest: the outside world will still treat any member of the group as corrupt.

The numerical externality of static nature exists in the models with endogenous probability of being caught (see Sah (1991), Clague (1993), Banerjee (1997)). If P has only a limited number of skilled and honest supervisors, the probability of being caught for each bureaucrat decreases with the number of potential perpetrators. Therefore there can be at least two stable equilibria (at least for some parameter values). First, there is an equilibrium where nobody is corrupt; the probability of being caught is very high so $D$ is very high. Second, there is another equilibrium where everyone wants to be corrupt and the probability of being caught is negligible ($D = 0$).

Thus one can easily come up with a model with several bureaucrats where numerical externalities result in multiplicity of equilibria. In low-corruption equilibria there will be no threat of corruption while in the high-corruption equilibria the corruption will be possible. Our results suggest that in the second equilibrium P will prevent ex post corruption via a collusion-proof contract, but her ability to implement the social optimum will be undermined. Therefore in the second equilibrium, red tape may be overproduced.

## 4.2 Separating contracts

In the Section 2 we have assumed that A is offered only one contract $(r, p, t, \sigma(\cdot))$. This is quite a restrictive assumption. The modern incentive theory suggests that when P wants both types to apply, she should offer A a menu of two contracts so that the bad type takes one contract and the good type takes the other one. In this Subsection we will first discuss why a single contract may still be the case and then will show that introduction of contract menus does not change the results qualitatively.

First, one may argue that offering one contract rather than two involves lower transactions costs. When there are only two types, the difference is certainly not very striking, while if there are many types (e.g. a continuum), the costs can become quite substantial. The second and the more important argument for a single contract is related to the role of the application fee. As it will be seen below, a separating contract does better than a pooling contract only if P uses both prices and application fees: the good type gets a contract with a positive application fee and lower red tape, the bad type pays no application fee but a positive price and goes through a greater amount of red tape. If P were not allowed to use application fees (e.g. due to the presumed innocence constraint), then the only menu that separates types would give a higher level of red tape to the good type. This is certainly inferior to the pooling contract: charging extra red tape to the good type is a pure cost for society.

Another issue with separating contracts is commitment. If the principal offers contracts 1 and 2 so that in equilibrium the good type takes contract 1 and the bad type takes contract 2, then the principal can infer the agent's type by the contract chosen. Ex post, after the agent has chosen contract 2, the principal knows the agent is of the bad type. Therefore it is optimal ex post not to provide her with the good. Expecting this, the bad type will never apply for contract 2 and separation will not happen.[26] It is well known, though (Salanie (1997)) that one can solve this problem if mixed strategies are allowed. For example, suppose that the bad type applies for the contract 2, while good type is indifferent between contracts 1 and 2; she applies for the contract 1 with probability $1-\xi$ and for the contract 2 with probability $\xi$. Then if P observes that A has taken contract 2 and got through the red tape $r_2$, P infers that A may be either of the good type (with probability $\frac{(1-\pi)\xi}{(1-\pi)\xi+\pi(1-I(r_2))}$) or of the bad type (with the probability $\frac{\pi(1-I(r_2))}{(1-\pi)\xi+\pi(1-I(r_2))}$). If $\xi$ is sufficiently high $(1-\pi)\xi v^g + \pi(1-I(r_2))v^b \geq 0$, then P will prefer to provide the good to the agents who have taken the second contract.

We will first describe the social optimum.

**Proposition 5** *Under A1-A3, the social optimum with partial separation is as follows. There exists such $\tilde{\pi} > \widehat{\pi}$ that:*

(a) *If $\pi < \tilde{\pi}$, then the optimal menu of contracts is as follows. The principal offers two contracts $(r_1, p_1, t_1, \sigma_1(\cdot)), (r_2, p_2, t_2, \sigma_2(\cdot))$ where $\sigma_1(\rho) = \sigma_2(\rho) = \rho$, $t_2 = p_1 = 0$, $t_1 = p_2 + r_2 - r_1$, $p_2 = \theta^b [1 - I(r_1^*)/I(r_2^*)]$, and $r_1 = r_1^*, r_2 = r_2^*$. Here $r_1^*, r_2^*$ maximize*

$$W^s = (1-\pi)(v^g - r_1) - \pi|v^b| (1 - I(r_2)) \left(1 + \frac{r_2 - r_1}{v^g}\right) - \pi r_2. \quad (10)$$

*subject to $\theta^b I(r_1) = (r_2 + \theta^b - \theta^g) I(r_2)$, $r_1 \leq \overline{r}$ and*

$$\xi = \pi|v^b|(1 - I(r_2)/[(1-\pi)v^g]. \quad (11)$$

*Both types apply. The bad type chooses the second contract. The good type chooses the first contract with probability $1-\xi$ and the second contract with probability $\xi$.*

(b) *If $\pi \geq \tilde{\pi}$ then it is optimal to offer one contract $r = \overline{r}$, $p = 0$, $t = \theta^g - \overline{r}$. The good type applies, while the bad type does not. The social welfare is $\overline{W} = (1-\pi)(v^g - \overline{r})$.*

Let us now turn to the case where B is self-interested. Formally, principal sets $(p_1, t_1, \sigma_1(\cdot)), (p_2, t_2, \sigma_2(\cdot))$ and offers $(s^b, s^g, s^0)$ to B. Then B completes the menu of contracts by $(r_1, \beta_1)$ and $(r_2, \beta_2)$. After this, A decides which contract $i = 1, 2$ to take and undergoes the red tape, B reports the outcome of the tests.

[26]Certainly, in a rule-of-law economy , a benevolent government should be able to commit to fulfilling its promises even if they are detrimental to the social welfare ex post. We have also tried to study the model with commitment and full separation and the results turned out to be very similar.

If $\pi \geq \tilde{\pi}$, the social optimum is the excluding contract and P can easily implement it by setting very high $s^0 - s^b$ and positive $s^g - s^0$. So let us turn to the case $\pi \geq \tilde{\pi}$ where P wants to implement a separating optimum. Apparently, P can implement the social optimum if there is no threat of corruption. If B cannot take bribes neither ex ante nor ex post, P should give B a flat contract $s^0 = s^g = s^b$. If both ex ante and/or ex post corruption may occur, implementation is again problematic.

**Proposition 6** *Assume A1-A3.*

1. *If $\pi \geq \tilde{\pi}$, then P implements the social optimum.*

2. *If $\pi < \tilde{\pi}$ and B can take bribes ex ante but not ex post then P implements the social optimum $(r_1^*, r_2^*)$ via setting $\sigma_1(\rho) = \sigma_2(\rho) = \rho$, $t_1 = t_2 = 0$, $p_1 \geq 0$, $p_2 = \theta^b - (\theta^b - p_1)I(r_1^*)/I(r_2^*)$, $s^0 - s^b = |v^b|\,[1 - (p_2 - p_1)/v^g]$ and $s^g - s^0 \geq \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$.*

3. *If $\pi < \tilde{\pi}$ and B can take bribes both ex ante and ex post then P implements the social optimum if and only if $\theta^b \leq |v^b|$.*

4. *If $\pi < \tilde{\pi}$ and B can take bribes then ex post but not ex ante P implements the social optimum via setting $t_1 = p_2 + r_2^* - r_1^*$, $p_2 = \theta^b(1 - I(r_1^*)/I(r_2^*))$, $s^g - s^0 \geq \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$ and $s^0 - s^b = \theta^b$.*

The logic of the Proof is similar to the proofs of Propositions 2-4. The first claim is trivial. The second claim is a little more complicated. As well as in the case of a pooling contract (Proposition 2), B would internalize the cost of red tape through bribes. The difference however is that since the share of the bad type who pass the test depends on $r_2$, the probability $\xi$ of the good type to choose the contract 2 also depends on $r_2$. Therefore B's incentives $s^0 - s^b$ should be slightly different from $|v^b|$ to adjust for this effect. The last statement is very similar to the Proposition 4: in order to implement the social optimum P sets the application fee high enough so that overproduction of the red tape would discourage the good type from applying.

## 4.3 The role of bargaining power

In the model above (Section 2) we have assigned all the bargaining power to B. Let us check what happens if the agent also has some bargaining power. Suppose that the distribution of bargaining power between B and A is $\lambda : 1-\lambda$, i.e. B gets only a fraction $\lambda \in [0, 1]$ of the {B,A} coalition's joint surplus. Going through proofs of Propositions 1-4, we can easily see that nothing changes in real terms unless there is a threat of ex post corruption.

The Propositions 1, 2 and 4 remain intact. The level of red tape is the same, only the distribution of rents between B and A changes in favor of A. The Propositions 3 that describes the second-best in the presence of potential ex post collusion, should be re-written. The matter is that the collusion-proofness constraint $s^0 - s^b \geq \theta^b$ becomes $s^0 - s^b \geq \lambda\theta^b$; it is less restrictive now. Instead of $\theta^b \leq |v^b|$ one should put in $\lambda\theta^b \leq |v^b|$. In other words, the less is the bargaining power of the bureaucrat, the more likely P is able to implement the social optimum.

This analysis suggests that the distribution of bargaining power determines the strength of the relationship between the threat of ex post corruption and the excessive red tape. Can the government increase the agent's bargaining power? A simple approach is to introduce competition among bureaucrats. Suppose that there are several offices that screen agents, and the agents can choose which one to go through. Then the bureaucrat's bargaining power decreases.[27] Unfortunately, competition is not always possible due to indivisibility of the public good or costs of information processing (Bardhan (1997)).

## 4.4 Other extensions

We have built a model of red tape that is informative but costly for the agent. One may argue that the red tape is also costly for the bureaucrat. The reason why we have focused on the red tape being costly to the agent rather than to the bureaucrat is because it is the cost for the agent that distinguishes the red tape from pure monitoring. A pure monitoring technology is costly for the monitor to operate but is free for the agent to be monitored. In the case of pure monitoring, the monitor's actions impose no externality on the agent and it is easier to provide incentives for the socially optimal level of monitoring if the ex post distribution is contractible (it is a standard moral hazard model). If the monitoring technology includes elements of pure red tape and pure monitoring (i.e. costly for both A and B), then the problem of internalizing the agent's cost is still present.[28]

We have also assumed that both price and application fee are non-negative. As innocent (and realistic) as it may look, this assumption is crucial for the whole analysis. If it were possible to charge a negative price, the socially optimal contract would be as follows: a negligible amount of red tape, a very large application fee, and a very large negative price. Namely, agents would have to make

[27]One should distinguish competition between several bureaucrats who produce substitutable services from coordination between bureaucrats who produce complementary services. The latter, however, can also be an important determinant of each bureaucrat's share in the joint surplus. See Rasmusen and Ramseyer (1994) for the explanation of the 'Tullock paradox' (Bardhan (1997)) along these lines.

[28]The cost of the red tape to the bureaucrat may also influence the principal's ability to eliminate the threat of ex post corruption. The cost may be too high for the principal (or an external auditor/supervisor) to check whether $\rho' = \rho$.

a large deposit, then pass a test that would reveal the bad type with a small but positive probability. Those who got through the test would receive the deposit back plus the good, while others would not get anything. Thus even an infinitesimal amount of red tape would be enough to deter the bad type from applying. However this contract would require very large transfers and therefore would not be efficient if we introduce the cost of transfer. Also, one would need to take into account agent's cost of borrowing explicitly rather than simply incorporate it into the private valuation of the good. (In our simple model the transfers are finite and therefore total cost of borrowing may be taken as given but this is not adequate if the transfers grow very large). In some sense, the assumption that prices and application fees are non-negative is the sacrifice we have made in order to be able to neglect the cost of transfers. Also, the presumed innocence constraint would rule out the use of application fees in a more general model where the red tape produces both type I and type II errors. If the use of application fees were not allowed, then the constraint that prices are non-negative is not important anymore. Notice that deterring the bad type from applying would still be possible with a combination of a positive price and high red tape. The only difference would be that since high prices hurt the bad type relatively less than high application fees, the amount of red tape required to exclude the bad type would be higher. Hence, the excluding contract would result in lower welfare.

An important issue is the distribution of control rights. We have assumed that P chooses the provision rule $\sigma(\cdot)$, price $p$ and application fee $t$ while B sets $r$. This seems to be a natural assumption: P cannot observe red tape but does observe whatever is paid to the government budget. On the other hand, one can imagine a situation where P is too busy to control $p$ and $t$ and delegates the choice of those to the bureaucrat. It turns out that if B controlled prices and application fees some results would change dramatically. In particular, P would not be able to prevent overproduction of the red tape by setting a high application fee. Thus if B could take bribes ex post but not ex ante, the collusion-proof contract would provide B with incentives to increase the red tape up to $r = \overline{r}$.[29] The control over $\sigma(\cdot)$ is not as important. Actually, in our model only one provision rule $\sigma(\rho) = \rho$ makes sense unless P wants to fire B and abandon the red tape altogether. Therefore a model where $\sigma(\cdot)$ is chosen by B but where P can decide not to hire B, would provide the same results.

Another important work to be done is a model with continuous types. In such a model the second best equilibrium may let some bad types in and keep

[29]Ironically, in this case, the bureaucrat who takes bribes ex ante and (potentially) ex post would produce less red tape in equilbrium than a bureaucrat who does not take bribes ex ante. So Huntington's claim ('the dishonest rigid bureacracy is better than honest rigid bureaucracy') would hold. However, there is a catch: ex ante corruption reduces red tape and increases welfare only in the presence of the threat of ex post corruption. The 'honest' bureaucrat does not take bribes as long as he is paid for being honest. If there were no threat of ex post corruption, ex ante corruption would not increase welfare.

some bad types out, or even exclude some good types. Also, the principal may choose to offer a contract that is *partially* collusion-proof: i.e. the bureaucrat has incentives to collude with some agents but not to collude with others. Therefore the corruption patterns and efficiency analysis can be very rich. Unfortunately, such a model would be too complicated: in order to distinguish between social and private values, we will need to consider a two-dimensional distribution of agents.

# 5 Conclusions

We have built a theory of informative red tape in a principal-bureaucrat-agent hierarchy, where the bureaucrat can take bribes ex ante and ex post. The red tape is modelled as series of tests that are costly to the agent but also reveal some information about the agent. In our model, both red tape and bribes are endogenously determined given the bureaucrat's propensity for each type of corruption. Our model helps to understand how red tape and different kinds of corruption interact in equilibrium. The model reconciles a number of common sense insights about red tape: (i) bureaucracy is slow and inefficient and there is too much red tape because bureaucrats do not care about the cost of red tape for the customers; (ii) still, rules and regulations can help to provide certain benefits to people who would not get them otherwise; (iii) corruption may undermine effect of rules and regulations but helps to reduce the excessive red tape; (iv) corrupt bureaucrats introduce more rules to extort bribes; (v) if bureaucrats are paid well, corruption is less common.

In our model, a benevolent government provides a good to different types of agents. Provision involves externalities so that the agents who are eligible for the good are willing to pay less than the agents whom the government wants to exclude. Because of the externalities, the agent's type cannot be perfectly screened by market mechanisms. This is why the government may want to use informative red tape. On the other hand, red tape is costly, so that the social optimum may involve partial rather than perfect screening. Government could deter bad types from applying but exclusion of the bad types may require too much red tape and therefore be too costly for the good type. Thus the benevolent government may prefer to allow the bad type to apply and receive the good with a positive probability. Unfortunately, even this allocation is hard to implement if the bureaucrat who operates the red tape is corrupt.

The government faces a complex web of challenges. On top of the adverse selection problem (agent's type is her private information), there is a two-dimensional moral hazard problem: the bureaucrat may set an inefficient level of the red tape ex ante and misreport the agent's type ex post. In the equilibrium the principal manages to solve the latter problem but this undermines her ability to solve the former. It turns out that although the threat of ex post corruption is not realized in equilibrium, it still results in overproduction of red tape, even after the bureaucrat reduces the red tape in exchange for bribes.

The contribution of the paper is threefold. First, it shows that even if principal has a few instruments for providing incentives such as prices, application fees and payments contingent on the ex post distribution, she fails to implement the social optimum if bureaucrat is corrupt. Second, analyzing ex ante and ex post corruption separately may be misleading since it is only the presence of both that makes the social optimum impossible to achieve. Third, the resulting allocation has too much (rather than too little) red tape compared with the social optimum.

Another interesting result is that excessive red tape occurs because of the threat of corruption even if there is no corruption in equilibrium. The principal can eliminate ex post corruption via offering a collusion-proof contract and reduce ex ante bribes via raising application fees. But the need for altering the bureaucrat's contract adds a constraint to the mechanism design problem and reduces the principal's ability to achieve efficient outcome.

The main empirical prediction of the model is that the threat of corruption leads to higher rather than lower red tape. This is consistent with country-level data although more empirical analysis should be carried to solve the identification problems and distinguish between potential and observed corruption. On one hand our model predicts a positive relationship between red tape and *potential* corruption. On the other hand, as argued in Subsection 4.1, the numerical externality may make potential corruption be correlated with observed corruption. The more corruption is observed in the country, the less costly it is for any single bureaucrat to take bribes. This, in turn, leads to a greater threat of corruption and excessive red tape. The policy implications are therefore straightforward: the effect of corruption on red tape is not at all innocent. Fighting corruption and especially potential corruption may indeed reduce red tape and increase welfare.

# Proofs

Proof of Proposition 1. The structure of the proof is as follows. We will divide all contracts into the four subsets: those under which both types apply, only the good type applies, only the bad type applies, and nobody applies. For each of the subsets we will find the one that achieves the maximum welfare. Then we will compare the four potential candidates for the social optimum. According to A3, all the four subsets are not empty: there exist contracts that allow to encourage only one type to apply, both types to apply and nobody to apply.

Let us first discuss the choice of the provision rule $\sigma(\rho)$. There are four possible rules: provide to everyone, provide to none, provide only to those who pass, and provide only to those who fail. If P wants to use any non-trivial amount of red tape $r > 0$ then only the rule $\sigma(\rho) = \rho$ (provide to those who pass) makes sense. Indeed, $\sigma(\rho) = \rho$ dominates the $\sigma(\rho) = 1 - \rho$. On the other hand, if P chooses one of the non-discriminating rules $\sigma(\rho) = 0$ or $\sigma(\rho) = 1$, then the red tape does not matter and should be abandoned. Thus P should compare the performance of three contracts: (i) provide to nobody, charge no red tape; (ii) provide to everybody, charge no red tape and (iii) charge positive red tape and provide to those who pass the tests. According to A1, (ii) is better than (i). According to A2, (iii) is better than (ii). Thus P chooses $\sigma(\rho) = \rho$. If the good type applies she gets the good with probability 1, while if the bad type applies she gets the good with probability $1 - I(r^b)$.

If the contract is such that neither type applies (e.g. application fee is very high $t > \theta^b > \theta^g$), P gets zero $W = 0$. If the contract is such that the bad type applies and the good type does not (e.g. $t = r = 0$, $p \in (\theta^g, \theta^b)$), then the welfare is negative $W = \pi v^b(1 - I(r^b)) - r^b < 0$.

Let us turn to the case where the good type applies and the bad type does not. P chooses $r, p, t$ to maximize the welfare

$$(1 - \pi)(v^g - r)$$

subject to the individual rationality constraints $U^b \leq 0 \leq U^g$ :

$$(\theta^b - p)(1 - I(r)) - r - t \leq 0 \leq \theta^g - p - r - t.$$

Since $p \geq 0$, these constraints imply $I(r) \geq 1-(\theta^g-p)/(\theta^b-p) \geq 1-\theta^g/\theta^b = I(\overline{r})$. Thus, $r \geq \overline{r}$ and $W = (1 - \pi)(v^g - r) \leq \overline{W}$. There exists a contract that achieves $W = \overline{W}$: take $r = \overline{r}$, $p = 0$ and $t = \theta^g - \overline{r}$. We will refer to this contract as the 'excluding contract'.

Now let us turn to the case where both agents apply for the good. P should offer $r, p, t$ that satisfy both types' participation constraints. The optimal contract maximizes

$$W = (1 - \pi)(v^g - r) + \pi((1 - I(r))v^b - r)$$

subject to

$$\begin{aligned} \theta^g - p - t - r &\geq 0, & (12)\\ (\theta^b - p)(1 - I(r)) - t - r &\geq 0. & (13) \end{aligned}$$

Apparently, we are interested only in contracts with $r \leq \overline{r}$. The contracts with $r > \overline{r}$ are dominated by the excluding contract. Given $r \leq \overline{r}$, the constraint (12) implies (13): $(\theta^b - p)(1 - I(r)) - t - r = \theta^g - p - t - r + pI(r) + \theta^b - \theta^g - \theta^b I(r) \geq \theta^b(I(\overline{r}) - I(r)) \geq 0$. Therefore, any $r \leq \overline{r}$ can be implemented: taking $p, t$ such that $t + p \leq \theta^g - \overline{r}$ would satisfy both participation constraints. Such $p, t$ exist: $p = t = 0 \leq \theta^g - \overline{r}$. The optimal pooling contract is $(r^*, p, t)$ where $r^*$ is given by (2), $t + p \leq \theta^g - r^*$. According to our assumption that P minimizes gross transfers, $t = p = 0$.

The welfare equals $W^* = (1 - \pi)v^g - \pi|v^b| + \pi|v^b|I(r^*) - r^*$. According to A2, $r^* > 0$ and $W^* > (1 - \pi)v^g + \pi v^b$. Therefore, the optimal pooling contract is more efficient than giving away the good to all types without any screening. This in turn implies, that the optimal pooling contract dominates contracts that exclude the good type and we only need to compare the pooling contract and the excluding contract.

It turns out that $W^* > \overline{W}$ if and only if $\pi$ is sufficiently small. Let us compare $(1 - \pi)v^g - \overline{W}$ and $(1 - \pi)v^g - W^*$. The former equals $(1 - \pi)\overline{r}$ which is a linear decreasing function of $\pi$ ($\overline{r}$ does not depend on $\pi$). It takes the value of $\overline{r}$ at $\pi = 0$ and 0 at $\pi = 1$. On the other hand, $(1 - \pi)v^g - W^* = r^* + \pi|v^b|(1 - I(r^*))$ is an increasing function of $\pi$ (by the envelope theorem). It takes the value of $r^* \leq \overline{r}$ at $\pi = 0$ and is still positive at $\pi = 1$. Therefore there exists such $\widehat{\pi} \in (0, 1)$ that $W^* > \overline{W}$ if and only if $\pi < \widehat{\pi}$. This $\widehat{\pi}$ is precisely the one introduced in A4.

■

Proof of Proposition 2. Let us study B's behavior under given $s^g \geq s^0 \geq s^b$, $p, t$ and $\sigma(\rho) = \rho$ (as argued in the Proof of Proposition 1, social optimum involves non-trivial amount of red tape which in turn requires $\sigma(\rho) = \rho$). By setting $\beta$ and $r$, B can deter each type from applying. Therefore, B should choose the set of applying types out of four possible combinations $\emptyset, \{g\}, \{b\}, \{g, b\}$.

Let us first prove that P can easily implement the outcome where the bad type is excluded, and B sets $r = \overline{r}$ so that P achieves the welfare $\overline{W}$. To do so, P must set a high punishment $s^0 - s^b$ for giving the good to the bad type and a sufficiently high bonus $s^g - s^0$ for giving it to the good type. Let us calculate how large the punishment and the bonus should be.

If B chooses to deter both types (e.g. by setting $r > \max\{R^g, R^b\}$), she gets $U^B = s^0$. If B wants only the bad type to apply, she maximizes

$$U^B = s^0 + \pi\left((s^b - s^0)(1 - I(r)) + \beta\right)$$

subject to $\theta^g - p - t - r - \beta \leq 0 \leq (\theta^b - p)(1 - I(r)) - t - r - \beta$. If the punishment giving the good to the bad type is very high $s^0 - s^b > \theta^b$ then $U^B < s^0$.

If B wants to exclude only the bad type, she maximizes

$$U^B = s^0 + (1 - \pi)((s^g - s^0) + \beta)$$

subject to $(\theta^b - p)(1 - I(r)) - t - r - \beta \leq 0 \leq \theta^g - p - t - r - \beta$. The two inequalities imply $I(r) \geq \frac{\theta^b - \theta^g}{\theta^b - p} \geq 1 - \theta^g/\theta^b = I(\bar{r})$, i.e. $r \geq \bar{r}$. If P gives B a positive bonus for letting the good type in $s^g - s^0 > 0$ then $U^B > s^0$.

If B allows both types in, then he sets $\beta = \min\{(\theta^b - p)(1 - I(r)) - t - r, \theta^g - p - t - r\}$ and chooses $r$ to maximize

$$s^0 + (1 - \pi)(s^g - s^0) + \pi(s^b - s^0)(1 - I(r)) + \min\{\theta^g - p, (\theta^b - p)(1 - I(r))\} - t - r. \quad (14)$$

Therefore, P can implement the excluding contract by setting $p = t = 0$, $s^g - s^0 > 0$ and $s^0 - s^b \geq \max\left\{\frac{\theta^b}{\pi}, \theta^b + \bar{r}\frac{\theta^b}{\theta^g}\frac{(1-\pi)}{\pi}\right\} > \theta^b$. The latter condition makes sure that if both types applied, B's payoff (14) would be less than $U^B = s^0 + (1 - \pi)((s^g - s^0) + \theta^g - \bar{r}$ which B gets by excluding the bad type.

Now let us assume that A4 holds and P wants to implement $r^*$ and sets $p = t = 0$, $s^g - s^0 = \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$ and $s^0 - s^b = |v^b|$. B has the following options: (i) allow both types to apply and set $r < \bar{r}$; (ii) allow both types to apply and set $r \geq \bar{r}$, (iii) exclude the bad type, (iv) exclude both types, (v) exclude the good type. Apparently, B prefers (iii) to (ii) . Let us compare (i) and (iii). When $r < \bar{r}$, the good type gets lower rent, $\beta = \theta^g - r$, so that B chooses $r$ to maximize

$$U^B = s^0 + (1 - \pi)(s^g - s^0) - \pi|v^b|(1 - I(r)) + \theta^g - r.$$

The solution is $r = r^*$ (2). If B wants to exclude the bad type he sets $\beta = \theta^g - \bar{r}$, $r = \bar{r}$ and gets

$$U^B = s^0 + (1 - \pi)(s^g - s^0) + (1 - \pi)(\theta^g - \bar{r}).$$

A4 implies that B strictly prefers (i) to (iii). A1 and A2 imply that B prefers (i) to (iv). Now we only need to compare (i) to (v). Surprisingly, excluding the good type is quite an attractive option for B: if the good type is excluded, B extorts all the rent for the bad type which may be quite high. By excluding the good type, B gets $U^B = s^0 + \pi(\theta^b - |v^b|)(1 - I(r)) - r$. If $\theta^b < |v^b|$, then $U^B < s^0$. On the other hand, if $\theta^b > |v^b|$ then B chooses $r = 0$ and gets $U^B = s^0 + \pi(\theta^b - |v^b|)$. If B chooses the pooling contract $r = r^*$, he gets

$$U^B = s^0 + (1 - \pi)(s^g - s^0) + \theta^g - \pi|v^b| + \pi|v^b|I(r^*) - r^*,$$

which is greater since $s^g - s^0 \geq \frac{\theta^b - \theta^g}{1-\pi}$. Hence, the social optimum $r = r^*$ is implemented.

*Comment.* Notice that giving B social incentives $s^g - s^0 = v^g$ and $s^0 - s^b = |v^b|$ does not always implement the social optimum $r = r^*$. Indeed, if $\pi\theta^b > (1-\pi)v^g + \theta^g + \pi|v^b|I(r^*) - r^*$ then B will prefer to exclude the good type and set $r = 0$. This is not important in this model, since $s^g - s^0 = \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$ does the job, but might matter a great deal in a more general setting where both type I and type II errors may occur with a non-trivial probability.

■

Proof of Proposition 3. The 'if' part is simple. One needs to refer to the Proof of Proposition 2. Since $s^0 - s^b = |v^b| \geq \theta^b$, P implements the social optimum with a collusion-proof contract.

We will prove the 'only if' part by contradiction. Suppose that $|v^b| < \theta^b$. If P wants to achieve $W^*$ she needs to offer a collusion-proof contract $s^0 - s^b \geq \theta^b$. Otherwise B always reports $\rho' = 1$ and the welfare is $(1-\pi)v^g - \pi|v^b| < W^*$.

Suppose that P manages to provide incentives for B to implement an outcome where $r < \overline{r}$, both types apply and $\sigma(\rho) = \rho$ (otherwise the welfare is always below $W^*$). Apparently, B $\beta = \theta^g - r - p - t$ and then chooses $r$ to maximize

$$U^B = s^0 + (1-\pi)(s^g - s^0) + \pi(s^b - s^0)(1 - I(r)) + \theta^g - r - p - t.$$

Therefore $I'(r) = \frac{1}{\pi(s^0 - s^b)} \leq \frac{1}{\pi\theta^b} < \frac{1}{\pi|v^b|} = I'(r^*)$. Hence, under a collusion-proof contract B strictly overproduces the red tape $r > r^*$ and the social optimum cannot be implemented.

■

Proof of Proposition 4. If A4 does not hold then it follows the Proof of Proposition 2 (the optimal excluding contract is collusion-proof). If A4 holds, then P achieves the social optimum through the following contract. P sets $\sigma(\rho) = \rho$, $p = 0$, $t = \theta^g - r^*$, and $s^0 - s^b = \max\{|v^b|, \theta^b\}$, $s^g - s^0 = \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$. This contract is collusion-proof. B chooses $r$ to maximize

$$U^B = s^0 + (1-\pi)\left(s^g - s^0\right)\mathbb{1}(U^g \geq 0) + \pi(s^b - s^0)(1 - I(r))\mathbb{1}(U^b > 0). \quad (15)$$

Apparently, B needs to make sure that the good type applies (excluding the good type gives $U^B < s^0$ while excluding both types gives $s^0$ only which is less than B gets if the bad type is excluded). Therefore B can only set $r \leq r^*$. Hence, $U^b > U^g \geq 0$ and the bad type applies as well. Since the bad type applies, B wants to set $r$ as high as possible to increase the probability to catch the bad type ($(s^b - s^0)(1 - I(r))$ increases with $r$). Hence B choose the socially optimal level of red tape $r = r^*$.

■

Proof of Proposition 5. Apparently, the (partially) separating contract weakly outperforms the pooling contract described in Proposition 1. Therefore the optimal menu of contracts will be either the excluding contract or a (partially) separating one. Under the former the bad type does not apply while the good type applies and gets the good. Under the latter both types apply.

We already know that the optimal excluding contract is $r = \overline{r}$, $p = 0$, $t = \theta^g - \overline{r}$, $\sigma(\rho) = \rho$. The welfare is $\overline{W} = (1-\pi)(v^g - \overline{r})$. Let us now look at the separating contract. P offers two contracts $(r_1, p_1, t_1, \sigma_1(\cdot))$ and $(r_2, p_2, t_2, \sigma_2(\cdot))$ so that the good type is indifferent between the two contracts and gets a non-negative surplus, while the bad type is better-off if she applies for the second contract that gives her a non-negative surplus, too. One can easily check that $\sigma_1(\rho) = \sigma_2(\rho) = \rho$ (see the Proof of Proposition 1). Then P maximizes

$$W = (1-\pi)v^g + \pi(1 - I(r_2))v^b - (1-\pi)(1-\xi)r_1 - ((1-\pi)\xi + \pi)r_2. \quad (16)$$

subject to the individual rationality (IR) and incentive compatibility (IC) constraints for the good type (17) and for the bad type (18), respectively:

$$\begin{aligned} 0 &\leq \theta^g - p_1 - t_1 - r_1 = \theta^g - p_2 - t_2 - r_2, && (17)\\ 0 &\leq (\theta^b - p_2)(1 - I(r_2)) - t_2 - r_2 \geq (\theta^b - p_1)(1 - I(r_1)) - t_1 - r_1. && (18) \end{aligned}$$

P is interested only in contracts with $r_1 \leq r_2$. Indeed, suppose that P offers a couple of contracts with $r_1 > r_2$ (that meet the constraints (17)-(18)). Then by offering just one contract $(r_2, 0, 0)$, P achieves the welfare $W = (1-\pi)(v^g - r_2) + \pi((1 - I(r_2))v^b - r_2)$ which is better than (16). Also, we may neglect the contracts with $r_1 > \overline{r}$ since those are dominated by the excluding contract. Since $r_1 \leq \overline{r}$, the first contract gives the bad type a higher rent. Therefore, the good type's IR and the bad type's IC constraints jointly imply the bad type's IR constraint. Hence the latter can be omitted.

Under given $r_2$, P chooses $p_1, t_1, p_2, t_2, r_1$ to minimize $r_1$. The only lower bound for $r_1$ is the bad type's IC constraint (the right one in (18)), so it must be binding:

$$(\theta^b - p_1)I(r_1) = (\theta^b - p_2)I(r_2). \quad (19)$$

Let us prove that $p_1 = 0$. Indeed, every menu of contracts with $(r_1, p_1, t_1)$,$(r_2, p_2, t_2)$, where $p_1 > 0$ is dominated by a menu $(r_1 - \varepsilon, 0, t_1 + p_1)$ and $(r_2, p_2, 0)$, where $\varepsilon > 0$ is sufficiently small — decreasing $p_1$ while keeping $t_1 + p_1$ strictly relaxes constraint (18) and allows to decrease $r_1$.

Similarly, $t_2 = 0$. First, (17) implies that $t_1 - t_2 = r_2 - r_1 + p_2 - p_1 \geq 0$. Second, take an arbitrary pair of contracts $(r_1, p_1, t_1)$ and $(r_2, p_2, t_2)$ that meet the constraints (17)-(18) and $r_1 \leq r_2$, $p_1 \leq p_2$, $t_1 \geq t_2 > 0$, and replace it with a pair $(r_1, p_1 + t_2, t_1 - t_2)$ and $(r_2, p_2 + t_2, 0)$. The new menu satisfies all the constraints, moreover it relaxes both constraints (18). Since (18) are not binding, we can strictly decrease $r_1$ and get strictly higher welfare. Thus every menu of contracts $(r_1, p_1, t_1)$ and $(r_2, p_2, t_2)$ is dominated by a menu with $t_2 = 0$.

Hence (19) becomes

$$\theta^b I(r_1) = (r_2 + \theta^b - \theta^g) I(r_2). \tag{20}$$

Therefore P chooses $r_2$ to maximize (16) subject to (20) and (11).

Notice that under the optimal separating contract the good type gets zero rent and the bad type gets a positive rent. Both types are indifferent between contracts 1 and 2.

Denote the maximum value of (16) $W^{s*}$. By construction, $W^{s*} \geq W^*$. Similarly to the proof of Proposition 1 we can show that $W^{s*} > \overline{W}$ if and only if $\pi$ is sufficiently small, hence there exists such $\tilde{\pi} \in (0,1)$ that $W^{s*} > \overline{W}$ if and only if $\pi < \tilde{\pi}$. Since $W^{s*} \geq W^*$, it must be the case that $\tilde{\pi} \geq \bar{\pi}$.

■

PROOF OF PROPOSITION 6. The Proof is similar to the Proofs of Propositions 2-4. Proving the first statement is straightforward: P can easily implement the excluding contract by setting a very high punishment $s^0 - s^b$ for giving the good to the bad type.

The proof of the second statement is harder than that of Proposition 2. Suppose that the principal sets $\sigma_i(\rho) = \rho$, $t_i, p_i, i = 1, 2$. The bureaucrat then sets $\beta_i, r_i, i = 1, 2$. The agent chooses whether to apply and which contract to apply for.

Suppose that the social optimum is achieved $r_1 = r_1^*, r_2 = r_2^*$. Then both types apply and the good type gets the same rent from both contracts. The bureaucrat maximizes

$$U^B = s^0 + (1-\pi)(s^g - s^0) - \pi(s^0 - s^b)(1 - I(r_2)) + (1-\pi)(1-\xi)\beta_1 + ((1-\pi)\xi + \pi)\beta_2.$$

The incentive compatibility and individual rationality constraints are as follows:

$$\begin{aligned} 0 &\leq \theta^g - p_1 - t_1 - r_1 - \beta_1 = \theta^g - p_2 - t_2 - r_2 - \beta_2, \qquad (21) \\ 0 &\leq (\theta^b - p_2)(1 - I(r_2)) - t_2 - r_2 - \beta_2 \geq (\theta^b - p_1)(1 - I(r_1)) - t_1 - r_1 - \beta_1 \end{aligned}$$

Since $r_1^* < \bar{r}$ the contract 1 would give a higher rent to the bad type than to the good type. Therefore, the bad type's IR constraint is not binding. Under given $r_1$ and $r_2$, B can raise $\beta_1$ and $\beta_2$ by the same amount up to the level which makes the good type's IR constraint binding: $\beta_1 = \theta^g - p_1 - t_1 - r_1$. Hence, $\beta_2 = \theta^g - p_2 - t_2 - r_2$.

Substituting those into the bad type's IC constraint we get

$$I(r_1) \geq (\theta^b - p_2) I(r_2) / (\theta^b - p_1). \tag{22}$$

while the bureaucrat's objective function becomes

$$\begin{aligned} U^B &= s^0 + (1-\pi)(s^g - s^0) - \pi(s^0 - s^b)(1 - I(r_2)) + \qquad (23) \\ &\quad + (1-\pi)(1-\xi)(\theta^g - p_1 - t_1 - r_1) + ((1-\pi)\xi + \pi)(\theta^g - p_2 - t_2 - r_2). \end{aligned}$$

For a given $r_2$, B wants to minimize $r_1$ ($\xi$ depends on $r_2$ only (11)). The only remaining constraint for $r_1$ is the bad type's IC (22), therefore it should be binding: $I(r_1) = (\theta^b - p_2)I(r_2)/(\theta^b - p_1)$. Comparing it to the relationship between $r_1^*$ and $r_2^*$, we obtain the condition for the optimal $p_1, p_2$ :

$$(\theta^b - p_2) = (\theta^b - p_1)(\theta^b - \theta^g + r_2^*)/\theta^b. \tag{24}$$

Substituting (11) into (23) and using (10) we obtain:

$$\begin{aligned} U^B &= W^s + s^0 + (1-\pi)(s^g - s^0 - v^g) + \theta^g - (1-\pi)p_1 - \pi p_2 - \\ &\quad -\pi(s^0 - s^b - |v^b|)(1 - I(r_2)) + (p_1 - p_2)\pi|v^b|(1 - I(r_2))/v^g. \end{aligned}$$

Therefore B chooses the socially optimal $r_2$ whenever

$$s^0 - s^b = |v^b| - (p_2 - p_1)|v^b|/v^g. \tag{25}$$

Notice that $p_2 \geq p_1$ implies $s^0 - s^b \leq |v^b|$.

We also need to check that B prefers to encourage both types to apply rather than deter one of them or both, and does not want to set $r_1 > \overline{r}$. This can be easily done as in the Proof of Proposition 2 (the condition $s^g - s^0 \geq \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$ is sufficient).

The proof of the third statement is similar to the Proof of Proposition 3. The social optimum requires collusion-proofness $s^0 - s^b \geq \theta^b$. On the other hand, the necessary condition for implementing the socially optimal $r_2$ (25) implies $s^0 - s^b \leq |v^b|$. Hence, the social optimum can only be achieved if $\theta^b \leq |v^b|$. The condition $\theta^b \leq |v^b|$ is also sufficient: the social optimum can indeed be achieved whenever it holds. Let us take $p_1 = p_2 = \theta^b$. The condition (24) is satisfied, and (25) becomes $s^0 - s^b = |v^b| \geq \theta^b$. The bureaucrat's contract is collusion-proof.

The proof of the last statement is very simple. Take the contract described in the Proposition (5): $p_1 = t_2 = 0$, $t_1 = \theta^g - r_1^*, t_2 = \theta^g - r_2^*$ and offer B a collusion-proof contract $s^g - s^0 \geq \max\{v^g, \frac{\theta^b - \theta^g}{1-\pi}\}$, $s^0 - s^b \geq \theta^b$. B maximizes

$$U^B = s^0 + (1-\pi)(s^g - s^0) - \pi(s^0 - s^b)(1 - I(r_2)). \tag{26}$$

subject to IC and IR constraints for both types (17)-(18). Since the contract is collusion-proof, B want to increase $r_2$ as much as the good type's participation constraint allows. Since the good type takes contract 2 with a non-trivial probability, $r_2 = \theta^g - p_2 - t_2 = r_2^*$. On the other hand, to make sure that the good type also takes contract 1, B sets $r_1 = \theta^g - p_1 - t_1 = r_1^*$. Notice that under these $p_i$ and $t_i$ B cannot exclude the bad type without excluding the good type. The sufficiently high bonus $s^g - s^0$ makes sure that B prefers both types to apply.

# REFERENCES


BAC, MEHMET AND PARIMAL K. BAG (1998). "Corruption, collusion and implementation: A hierarchical design." *mimeo*, University of Liverpool.

BANERJEE, ABHIJIT (1997). "A theory of misgovernance." *Quarterly Journal of Economics,* 112(4), pp.1289-1332.

BARDHAN, PRANAB. (1997). "Corruption and development: a review of issues" *Journal of Economic Literature*, 35(3), pp.1320-1347.

BECKER, GARY S. AND GEORGE J. STIGLER (1974). "Law enforcement, malfeasance and the compensation of enforcers." *Journal of Legal Studies,* 3, pp.1-19.

CARRILLO, JUAN (2000). "Corruption in hierarchies." *Annales d'Economie et de Statistique* (forthcoming).

CLAGUE, CHRISTOPHER (1993). "Rule obedience, organizational loyalty and economic development" *Journal of Institutional and Theoretical Economics*, 149, pp. 393-414.

DEWATRIPONT, MATHIAS, IAN JEWITT AND JEAN TIROLE (1997). "The economics of career concerns, part II: application to missions and accountability of government agencies." *Review of Economic Studies*, 66(1), pp.199-217.

EBRD (1999). *Transition report: Ten years of transition*, London.

ELLIOTT, KIMBERLY ANN (ED.) (1997). *Corruption and the global economy.* Institute for International Economics, Washington, D.C.

HAUK, ESTHER AND MARIA SAEZ MARTI (1998). "On the cultural transmission of corruption." *Mimeo, Stockholm School of Economics.*

HELLMAN, JOEL S., GERAINT JONES AND DANIEL KAUFMANN (2000). "Seize the state, seize the day: An empirical analysis of state capture and corruption in transition" *Paper presented at ABCDE Conference*, Washington, D.C.

HUNTINGTON, SAMUEL (1968). *Political order in changing societies.* New Haven, CT: Yale University Press, pp. 498-99.

KAUFMANN, DANIEL, AART KRAAY AND PABLO ZOIDO-LOBATON (1999). "Aggregating governance performance." *World Bank Policy Research Paper 2195*, Washington, D.C.

Klitgaard, Robert (1988). *Controlling corruption.* University of California Press.

Kofman, Fred and Jacques Lawarree (1993). "Collusion in hierarchical agency." *Econometrica*, 61(3), pp. 629-56.

Kofman, Fred and Jacques Lawarree (1996). "On the optimality of allowing collusion." *Journal of Public Economics*, 61, pp. 383-407.

La Porta, Rafael, Florencio Lopes-de-Silanes and Andrei Shleifer (1999). "The Quality of Government." *Journal of Law, Economics and Organization*, 15, pp.222-279.

Mauro, Paulo (1995). "Corruption and growth." *Quarterly Journal of Economics,* 110, pp. 681-712.

Parsons, Donald O. (1991) "Self-screening in targeted public transfer programs." *Journal of Political Economy,* 99, pp. 859-76.

Polinsky, A. Mitchell and Steven Shavell (2001) "Corruption and optimal law enforcement." forthcoming in *Journal of Public Economics.*

Rasmusen, Eric and J. Mark Ramseyer (1994). "Cheap bribes and the corruption ban: A coordination game among rational legislators." *Public Choice,* 78 (3-4), pp.305-327.

Rijckenghen, Caroline van and Beatrice Weder (1997). Corruption and the rate of temptation: do low wages in the civil service cause corruption. *IMF Working paper No. 97-73.*

Rose-Ackerman, Susan (1978). *Corruption: A Study in Political Economy.* Academic Press.

Sah, Raaj (1991). "Social osmosis and patterns of crime." *Journal of Political Economy,* 99, pp.1272-95.

Salanie, Bernard (1997). Economics of Contracts: A Primer. MIT Press.

Shleifer, Andrei and Robert Vishny (1993). "Corruption." *Quarterly Journal of Economics,* 108, pp.599-618.

Spence, Michael, (1973) "Job market signalling," *Quarterly Journal of Economics*, 87, pp.355-74.

Tanzi, Vito (1998). 'Corruption around the world: Causes, consequences, scope and cures'. *IMF Staff Papers,* 45(4), pp. 559-94.

TIROLE, JEAN (1992). "Collusion and the theory of organizations." In J.-J. Laffont (ed.) *Advances in Economic Theory.* Sixth World Congress. Cambridge University Press.

TIROLE, JEAN (1996). "A theory of collective reputation (with applications to the persistence of corruption and firm quality)." *Review of Economic Studies*, 63, pp. 1-22.

WILSON, JAMES Q. (1989). *Bureaucracy: what government agencies do and why they do it.* NY: Basic Books.