Bribery vs. extortion:
allowing the lesser of two evils

August 3, 2009

FAHAD KHALIL† JACQUES LAWARRÉE‡ SUNGHO YUN*

Abstract

Both bribery and extortion weaken the power of incentives, but there is a tradeoff in fighting the two since rewards to prevent supervisors from accepting bribes create incentives for extortion. Which is the worse evil? A fear of inducing extortion may make it optimal to tolerate bribery, but extortion is never allowed. Extortion discourages "good behavior" because the agent suffers from it even though he has done the right thing, while a bribe acts as a penalty for "bad behavior". Our analysis provides lessons to fight corruption and explanations why developed countries may have an advantage in dealing with extortion.

JEL Classification: D82, L23
Key words: Monitoring, Corruption, Collusion, Bribery, Extortion, Framing, soft information.

† Department of Economics, University of Washington, Seattle, WA 98195, khalil@u.washington.edu, and ERG, Dhaka.
‡ Department of Economics, University of Washington, Seattle, WA 98195 and ECARES, Brussels lawarree@u.washington.edu
* Department of Economics, Hanyang University, Ansan, Korea, uwyunsh@hanyang.ac.kr

We would like to thank Jin Cao, Anke Kessler, Ivan Png, Troy Scott, Steve Shavell, seminar participants at various places, the editor, Jennifer Reinganum, two anonymous referees, and especially Juan Carrillo and Jean Tirole, for their help and suggestions. Devrim Yavus provided very useful research assistance.

1. Introduction

In the optimal design of organizations, the fight against bribery by enforcement officers relies on strong incentives to detect and report violations by agents. For instance, a driver under the influence of alcohol may attempt to bribe a police officer to let him off the hook for a DUI conviction, but a corrupt officer will find it less profitable to accept a bribe if he can collect a reward when turning in the drunk driver.[1] However, such a reward raises the specter of extortion since rewards to deter bribery may act as inducements to engage in extortion. Consider the case of an officer catching drivers who run red lights. Again, a reward would lower his incentive to accept a bribe from a driver caught running the light, but the same reward may invite a corrupt officer to claim that the driver ran the light when he did not. Incentive to deter bribery may lead a corrupt officer to extort innocent drivers.

The goal of this paper is to analyze the distinction between bribery and extortion, both in their impact on incentives for agents to obey rules, and on corrupt parties to engage in illegal transactions. Clarifying the distinction allow us to rank the two evils and characterize optimal mechanisms to fight them. In our model, bribery and extortion differ by their effect on an agent when corrupt enforcers manipulate evidence to extract money from the agent. Enforcers can manipulate evidence in two different ways: (a) make a favorable report about the agent — this will be called bribery in this paper; (b) make an unfavorable report about the agent — this will be called extortion.[2] We also use the generic term of corruption to describe bribery and extortion. We explain the differences in the impact on agents' incentives below, but note first that bribery is cooperative, with the parties pursuing a common objective, while extortion is antagonistic, with one party benefiting at the expense of the other. This difference in objectives of the parties engaging in a corrupt transaction can be exploited in the design

[1] The reward can be non-monetary such as good reputation, promotion, etc. Similarly, bribes and extortion payments can take the form of favors to members in an organization.

[2] In the legal literature, there is a debate on the definitions of extortion and bribery based on who initiates the corrupt transaction. For example, Ayres (1997) argues that in an environment where corruption is endemic, an individual initiating a side-payment to an enforcement agent could well be the victim of extortion rather than someone attempting to engage in bribery. See also Lindgren (1993). We are able to abstract from this debate by focusing on whether the corrupt behavior helps or hurts the agent as we are mainly interested in optimal incentives for the agent.

of anti-corruption measures. Yet, too often, the popular debate does not distinguish between the two, treating them together as merely illegal or immoral payments to enforcers, obfuscating fundamental issues.

By combining this difference in objectives and the role of teamwork in forging evidence, we present a model of extortion where attempts to deter bribery lead to a threat of extortion. We show that the supervisor remains useful even when there is no external honest enforcement available. We derive two main results: (i) extortion should always be deterred but allowing bribery may be optimal; (ii) bribery is deterred when information is hard but may be allowed when information is soft. There is an extensive literature in economics dealing with bribery, but our result that the threat of extortion makes bribery optimal is new.[3] We also find that the principal is better off if the agent has less bargaining power when negotiating a bribe, and that higher outside opportunities for the agent makes extortion less relevant.

The intuition for result (i) depends on the fact that there is a critical difference in the cost of providing incentives to the agent in the presence of bribery as compared to extortion. Even though both increase incentive cost, extortion discourages "good behavior" because the agent is subject to it even though he has done the right thing. Bribery, however, helps somewhat in providing incentive because it occurs when the agent is seen as violating the rules. The bribery payment acts as a penalty for "bad behavior". This is in line with the less formal literature that suggests that bribes may have some positive role to play but extortion does not (see Bardhan (1997)). Bribery can help "grease the wheels" in badly run organizations but, as Klitgaard (1988) noted, "Extortion is a particularly debilitating form of corruption."… "It leads not only to inefficiencies but the alienation of citizens from their government."

The above suggests extortion is worse than bribery, but it does not say why both should not be deterred. Indeed, in result (ii), we find that even if it is feasible to deter

[3] See the surveys by Tirole (1992) and Bardhan (1997), and references in Khalil and Lawarree (2006), or Silva et al. (2007) for recent contributions. The related literature is discussed in more detail at the end of this section.

both, it is optimal to allow bribery when information is soft. [4] Most of the existing literature, which relies on hard supervisory information, finds that deterring bribery is optimal.[5] Suppose, as in Tirole (1986) or Laffont and Tirole (1993), that the supervisor can either find hard evidence (positive or negative) or no conclusive evidence. With hard evidence, the supervisor can conceal information and pretend she has found no conclusive evidence but she cannot forge evidence. For example, in the case of DUI, a policeman may ignore a tainted blood sample but cannot create one. It turns out that in this information structure, the supervisor is only rewarded for reporting negative evidence, and a threat of extortion is not credible.[6]

In our model, we assume that information is soft for the agent-supervisor coalition, i.e., the coalition can forge evidence. This implies that the principal has to pay the supervisor a new reward to deter the forging of evidence, and this new reward makes extortion credible. A tradeoff between bribery and extortion appears when information is soft. Note that even though information is soft for the coalition, the principal is not powerless. It can exploit the difference in objectives between the agent and supervisor, which plays a key role in fighting extortion in the model. Recall that bribery is collusive and turns out to be more difficult to fight; we find that bribery occurs in equilibrium.

Our results are also consistent with the fact that extortion is mainly a problem in less developed countries relying mostly on soft evidence, while in developed countries hard evidence is more common and it is mainly bribery that makes the news. In the financial world for instance, making information hard can take various forms and be represented by the use of institutions like lawyers, CPAs, auditors, bankruptcy courts, independent directors and legal actions by the shareholders (see the survey paper by La Porta et al. (2000)).

---

[4] We will explain in Section 4 that Tirole's (1986) well-known 'collusion-proofness principle' does not apply due to interlinked coalition incentive constraints. A direct reward from the principal to the supervisor invites extortion, so the principal finds it cheaper to let the agent reward the supervisor in the form of a bribe.

[5] Our focus is on the agency literature that followed the pioneering work by Tirole (1986, 1992) as opposed to the non-agency literature (as reviewed in Bardhan (1997)). See also Mishra (2005) for a recent survey.

[6] The argument is explained in detail later. If the supervisor has no conclusive evidence, she has no discretion and no bribery or extortion can occur. Consequently, if she has positive evidence about the agent and wants to threaten to extort by concealing it, her threat is not credible. This is because she will not be rewarded if she reports no conclusive evidence.

There is a growing literature that examines relationships between Information Technology (IT) adoption and incentives in organizations. Most of this literature has found evidence that harder information provided by IT allows the principal to provide stronger incentives.[7] This is consistent with our model as hard-information based contracts are likely to be less susceptible to extortion and the agent's incentives can be made stronger and more efficient. An interesting extension would be to find out if police corruption has declined after the introduction of technologies such as video camera in patrol cars or red-light cameras. A recent study, lending support to this view, shows that red-light cameras help prevent traffic stops triggered by extortion motives based on racial considerations (Colb (2001)).

We consider extensions of the model and derive further results. Extortion is a less serious issue when the agent has less bargaining power, and if he has stronger outside opportunities. A lesser bargaining power hurts the agent as the supervisor can extract a larger bribe. This makes the bribe a more effective penalty, and it is less costly to allow bribery. Better outside opportunities also make extortion less of an issue as they increase the agent's reservation utility and help protect the agent from the supervisor's extortion attempts. A higher reservation utility forces the principal to increase the risk-averse agent's wage while making it less dependent on the supervisor's report. We show that with strong enough outside opportunities extortion is no longer a threat for the agent. Again, this seems consistent with evidence that extortion is mainly a problem in less developed countries where agents have weaker outside opportunities.

Related Literature

There has been surprisingly little attention given to corruption in the economic theory of law enforcement. Shavell's (2004) authoritative textbook on law and economics has no references on corruption. This was also noted by Polinksy and Shavell (2000) in their comprehensive survey, in which the first item on the agenda for future research is the study of incentives for enforcement agents and the fight against corruption.

[7] For instance, Baker and Hubbard (2003) show how the adoption of on-board computers in trucks allowed trucking companies to offer better incentive contracts.

Corruption has received much more attention in the literature on hierarchies that followed Tirole (1986), but the issue of extortion was largely ignored. The intuition that rewards to enforcement agents may also encourage extortion has not played much of a role in this literature. Tirole (1986) showed that a corruptible supervisor can still be useful, but his model relies on hard information, and therefore extortion was not a credible threat. Much of the literature following Tirole focused on the problem of bribery in models where extortion is not relevant. For instance, extortion is not relevant in Kessler (2000) since information is hard. Baliga (1999) analyzes the case of soft information but extortion does not increase the implementation costs because the mechanism of the game allows the agent to quit when faced with the possibility of extortion.[8] In Kofman and Lawarree (1993) the information structure allows forging of evidence but rules out extortion by assumption. Like Tirole, the above papers find that it is optimal to deter bribery. We contribute to this literature by pointing out that if information is soft, the threat of extortion may make it optimal to allow bribery.[9]

Besides the non-agency literature reviewed in Bardhan (1997), there have been a few recent models of extortion in agency settings. These papers feature extortion in different settings and with a different focus than ours. Polinsky and Shavell (2001) study an optimal law enforcement problem. Andrianova and Melissas (2008) consider the case where the principal facing a threat of corruption has the option to legalize a socially undesirable activity. Acemoglu and Verdier (2000) study the choice between government intervention to address a market failure and the resulting bureaucratic corruption. Mookherjee (1997) and Hindriks et al. (1999) consider a tax-evasion model. Mookherjee focuses on reforms in public bureaucracies and Hindriks et al. on the redistributive properties of the tax scheme. To deter corruption, all five papers rely on the availability of incorruptible external enforcement agents and the penalties they can

[8] See also Faure-Grimaud et al. (2003) for a model of soft information with asymmetric information between the supervisor and the agent.

[9] Others have also shown that it may be optimal to allow bribery, but unlike us they do not feature extortion. For instance, Kofman and Lawarree (1996), Acemoglu and Verdier (2000) and Auriol (2006) introduce uncertain auditor type; Che (1995) and Mookherjee and Png (1995) feature auditor moral hazard; Strausz (1997a), Olsen and Torsvik (1998), Lambert-Mogiliansky (1998), and Khalil and Lawarree (2006) consider renegotiation and no-commitment; Shin (2007) is another example of beneficial collusion. See also Cadot (1987) or Carrillo (2000a) on the perverse effects of anti-corruption measures.

impose.[10] Instead, we focus on internal mechanisms to deter bribery and extortion by developing an informational structure that makes a corrupt supervisor useful even though incorruptible external enforcers are absent. In a procurement setting without external enforcement, Auriol (2006) allows for the possibility of extortion and bribery. However, unlike us, the extortion payment is only a redistribution between the agent and the supervisor causing no allocative inefficiency. Thus she indentifies circumstances where extortion may be allowed since it is costly to deter, but bribery is only allowed when the corruptibility of the auditor is not known (as in Acemoglu and Verdier (2000), and Kofman and Lawarrée (1996)).

2. The Setup

We build on a standard moral hazard problem in a principal/supervisor/agent hierarchy. The principal (it) is the owner of a firm, the agent (he) is the productive unit in the firm, and the supervisor (she) collects information for the principal. The agent produces an output x which depends on his level of effort, $e \in \{0, 1\}$. If the agent works, that is, $e = 1$, he produces x_H with probability π and x_L with probability $1 - \pi$, where $x_H > x_L$, and $\pi \in (0, 1)$. If he shirks, that is, $e = 0$, he produces x_L with probability one.[11] While the level of output x is observed by all parties, the level of effort e is private information of the agent. The agent's disutility of effort is given by φe, where $\varphi > 0$. The output belongs to the principal, who pays a transfer w to the agent. We assume that the agent is risk averse with a separable utility function given by $U(w, e) = u(w) - \varphi e$, where u is concave, $u(0) = 0$, and satisfies the Inada conditions ($u'(0) = + \infty$ and $u'(+ \infty) = 0$). The principal, who is risk-neutral, offers a take-it-or-leave-it contract to the agent, who has a zero reservation utility.[12] We assume that x_H - x_L is large enough such that it is profitable to induce the agent to work, that is, exert $e = 1$. The principal's objective is to minimize its expected cost of inducing $e = 1$.

[10] Models of red tape (see, e.g., Banerjee (1997), and Guriev (2004)) also rely on a similar assumption.
[11] In Section 5, we show that our main results are robust to a more general production function.
[12] We consider the case of a strictly positive reservation utility in Section 5.

In the absence of a supervisor, the contract for the agent can only be based on x, and the wages would be w_L when x_L is produced and w_H when x_H is produced.[13] In this model, the optimal contract in the absence of a supervisor — the *second-best contract* — requires that $w_H^s = u^{-1}(\varphi / \pi)$ and $w_L^s = 0$.[14] In other words, the principal compensates the agent only when there is definitive evidence that the agent worked, i.e., when x_H is realized. The agent does not obtain any rent.

The supervisor's role is to collect information about the agent's effort level and to report it to the principal. Since x_H can be realized only with $e = 1$, there is no reason to use the supervisor following x_H, and the principal will send the supervisor only when it observes x_L. Following Tirole (1986), we assume that the supervisor observes the true level of effort with probability p or obtains no conclusive evidence with probability $1 - p$, where $p \in (0, 1)$. The supervisor's signal σ can take three values: $\sigma \in \{0, \varnothing, 1\}$, where $\varnothing$ denotes that the supervisor does not have conclusive evidence about effort. Therefore, the agent is given a wage w_H following x_H, and w_r, following x_L, where r is the supervisor's report with $r \in \{0, \varnothing, 1\}$. We assume that the supervisor is costless but the principal may want to pay her a wage s to deter corruption.[15] The supervisor is risk neutral, and it is common knowledge that the supervisor is corruptible.[16] Without loss of generality, the wage to the supervisor depends only on her own report and is denoted by s_r. We assume that the supervisor's reservation utility is zero. Both the agent and supervisor are protected by limited liability such that $w_r \geq 0$ and $s_r \geq 0$.[17]

Supervision Technology and Corruption

In this sub-section we will define bribery and extortion and introduce a key new feature in the supervision technology that makes extortion relevant. The supervisor is corrupt in

[13] We assume the principal itself does not have the expertise to monitor the agent. See Strausz (1997b) for a model of collusion comparing monitoring by the principal or delegation of monitoring to the supervisor.
[14] Technically, the second best requires $u(w_L) = 0$, but in our model this implies $w_L = 0$ since we have assumed $u(0) = 0$.
[15] We abstract from supervisor's moral hazard (costly effort) in order to focus on the tradeoff between extortion and bribery. Mookherjee and Png (1995) have shown that bribery may occur in equilibrium if the supervisor exerts an unobservable audit effort.
[16] In Section 5 we consider the case where the supervisor is honest with a positive probability.
[17] Without limited liability, the first best could be reached since $e = 0$ is off the equilibrium path. When the supervisor reports that $e = 0$, the principal can impose an infinite punishment on the agent, and also give a large reward to the supervisor to deter corruption.

the sense that she may not always report what she has observed to the principal. She will report the truth only if it is in her interest to do so. In this environment, we identify two types of corrupt behavior, which we define below. The key distinction is whether the corrupt behavior benefits the agent (the case of bribery) or hurts him (the case of extortion).

Definition 1. *Bribery* occurs when the supervisor accepts a payment from the agent in return for misreporting information to benefit the agent.

Definition 2. *Extortion* occurs when the supervisor obtains a payment from the agent by threatening to misreport information to hurt the agent. *Framing* would occur if the attempt at extortion fails and the supervisor misreports information that was favorable to the agent.

Bribery and extortion are accompanied by side-contracts between the supervisor and the agent whereas framing is not. With bribery, the supervisor and the agent forge information to maximize their joint surplus. With extortion (resp. framing), the supervisor acts alone by threatening to misreport (resp. actually misreporting) evidence since she is acting against the agent's interest. We require that extortion or framing be sequentially rational; the supervisor's threat is credible only if she receives a higher payoff by misreporting evidence than by revealing it truthfully.

We depart from the literature on monitoring with hard information, which relies on the idea that it is relatively easy to conceal but very costly to forge information. In reality, there is often an asymmetry in the cost of forging information if the supervisor tries to do it alone or if she has help from the agent. Consider the previous example of the blood test taken after a car accident. If the police officer or the lab worker colludes with the driver, they can easily substitute another untainted blood sample. This means that information can be more easily manipulated when several people collaborate.[18] In the spirit of the recent literature on communication (Dewatripont and Tirole (2005) or Caillaud and Tirole (2007)), we emphasize that forging evidence, like information transmission, is a team activity and its cost depends on the amount of help from team members. Dewatripont and Tirole argue that, working together as a team, the sender and

[18] In financial auditing for instance, the auditee can help the auditor draw "favorable samples."

the receiver can make soft information hard. Because our focus is on fraud, we look at the opposite issue: can hard information be made soft? The key feature of our model is that information that is hard for the supervisor can become soft for the supervisor-agent coalition.[19]

We incorporate the cost of forging evidence into our model in a simple way as follows. First, without the agent's cooperation, the supervisor cannot forge information (her cost is infinite) – she can only conceal it. Her information is hard. If $\sigma = e$, she can only report $r \in \{e, \varnothing\}$, and if $\sigma = \varnothing$, the only possible report is $r = \varnothing$. Second, with the agent's cooperation, the supervisor can forge evidence at zero cost and report that the agent has worked regardless of what she observed, i.e., it is possible to have $r \in \{0, \varnothing, 1\}$ regardless of σ. The information is soft for the coalition.[20]

It is worth noting that extortion by the supervisor is relevant when information is soft for the *coalition* but hard for the supervisor. Supervisory extortion would not be an issue if the information were either soft or hard. If the information were soft for the supervisor, the supervisor would be useless. Soft information makes the supervisor's signal irrelevant since she can report anything. Since corruption is possible, the supervisor can take advantage of any variation in the payment schemes offered as incentives to the agent. If the information were hard for both the supervisor and the coalition, extortion would not be relevant.[21] This is because a threat of extortion is credible only if the supervisor is able to collect a reward by suppressing information. Since evidence cannot be forged, the supervisor has no discretion when $\sigma = \varnothing$, and there is no need to reward the supervisor when $\sigma = \varnothing$. Therefore, the threat of extortion by suppressing evidence is vacuous in a model with hard information as it is the case in many prominent models like Tirole (1986, 1992) or Kessler (2000).[22] In our model, it is

---

[19] In our model, the principal (receiver) remains passive in the information transmission process, but the cost of forging depends on whether the supervisor (sender) acts alone or receives help from the agent.

[20] We thus assume that it costs the coalition the same whether the misreport is from $\varnothing$ to 1 or from 0 to 1, but this is not restrictive since misreporting after $\sigma = 0$ can be deterred without cost in equilibrium. Similarly, assuming that it is equally inexpensive to suppress evidence from 0 or from 1 is not restrictive.

[21] Thus the relevant type of corruption is bribery, which is deterred in equilibrium.

[22] There is a series of papers by Vafai (cited in Vafai (2005)) analyzing extortion under hard information. To make extortion credible Vafai relies on the "prohibitive psychological or emotional cost" of not carrying out a threat and he shows that bribery can be deterred without cost.

the reward to deter forging of information when $\sigma = \varnothing$ that makes the threat of extortion credible.

Since bribery may occur in equilibrium, we need to be explicit about how side transfers are determined. Besides the standard assumption of enforceable side-contracts (see Tirole 1992), we also assume the side-transfers are determined according to the Nash bargaining solution.[23]

We summarize the model by presenting the timing of moves:

(1) The principal offers a contract specifying the transfers to the agent as a function of output and the supervisor's report; and the transfers to the supervisor as a function of her report.

(2) The agent and the supervisor accept/reject the contract.

(3) The agent decides whether to work ($e = 1$) or shirk ($e = 0$).

(4) Output x is realized. If the principal observes x_L, it sends the supervisor. If it observes x_H, the game moves to (8).

(5) The supervisor and the agent observe the signal σ.

(6) The supervisor and the agent choose whether or not to make a side-contract.

(7) The supervisor makes a report r.

(8) Transfers are realized.

3. Tradeoff between Bribery and Extortion

In this section we will argue that rewards to deter bribery will lead to extortion, but that it is feasible to deter both. In Section 4, we show that it is optimal to allow bribery but not extortion. First, as a benchmark, we briefly outline the *incorruptible-supervisor* contract and provide its details in Appendix A.

If the supervisor were incorruptible, the supervisor would not be paid any reward, and $s_r = 0$, for all r. The agent would only be rewarded when there is *definitive* evidence of effort, i.e., if x_H occurs or if x_L occurs but the supervisor finds evidence of work (r =

[23] Note that the supervisor and the agent negotiate the side-contract under symmetric information. Another strand of the literature considers collusion under asymmetric information (see Laffont and Martimort (1997, 2000), or recently Che and Kim (2006)).

1); the agent would be paid zero otherwise. The agent would not obtain any rent and be equally compensated after x_H and after x_L with $r = 1$, i.e., $w_H = w_1 > 0 = w_\varnothing = w_0$. The incorruptible supervisor provides insurance to the agent who can be paid a positive wage w_1 even after a low output. Therefore, his wage after x_H and the principal's expected wage payments are both smaller relative to the second-best contract.

This contract, however, is vulnerable to bribery. Since the supervisor is not being rewarded ($s_r = 0$), the agent will bribe the supervisor. When she finds no-evidence ($\sigma = \varnothing$) or evidence of shirking ($\sigma = 0$), the agent will help her fabricate evidence to give a report of work ($r = 1$) so that they can share the higher wage w_1 collected by the agent.

On first sight, this threat of bribery can be combated by introducing a reward for the supervisor when she reports shirking ($r = 0$) or no-evidence ($r = \varnothing$). If the reward is equal to w_1 (i.e., $s_0 = s_\varnothing = w_1$), there will be no incentive to bribe. The supervisor is turned into a bounty hunter as in, e.g., Tirole (1986) or Kofman and Lawarrée (1993). However, in our framework, this would introduce a new problem of extortion by the supervisor. To see this, note first that $s_1 = 0$ in the bounty-hunter scheme since there is no perceived threat of a bribe from the agent when $\sigma = 1$. Thus, when she has evidence of work, the supervisor will have an incentive to suppress this evidence to obtain the reward $s_\varnothing > 0$ rather than get $s_1 = 0$.[24] This is the tradeoff mentioned in the introduction: strong incentives to deter bribes create scope for a new kind of corruption, namely extortion. As noted above, this tradeoff does not appear when information is hard.

The least-cost-corruption-proof (LCCP) contract: no bribery or extortion

Next we present the contract where the principal deters both bribery and extortion. It is not clear a priori if it is optimal to deter all types of corruption particularly given the tradeoff discussed above. We show that the *LCCP* contract is not optimal in general because it restricts the principal too much in utilizing the information provided by the

[24] Anticipating extortion the agent will refuse to put in high effort (his incentive constraint will be violated). Note also that raising s_1 to $s_\varnothing$ is problematic since it would encourage the coalition to report $r = 1$ when $\sigma = \varnothing$.

supervisor.[25] However, the *LCCP* contract serves as a useful benchmark and is also a critical step when we derive the optimal contract in the next section.

Before presenting the principal's problem with the usual incentive and participation constraints, we first need to consider the stage of the game with bribery and extortion. To prevent bribery the principal will have to ensure that the contract satisfies the Coalition Incentive Compatibility (*CIC*) constraints. Defining the aggregate transfer $T = w + s$, we can present these constraints as:

$$T_\sigma \geq T_r\,, \qquad \text{where } T_\sigma = w_\sigma + s_\sigma\,, \quad T_r = w_r + s_r\,, \ \text{ for } \sigma, r \in \{0, \varnothing, 1\}. \qquad (CIC_{\sigma,r})$$

We have six (*CIC*) constraints and these can be satisfied only when $T_0 = T_\varnothing = T_1$, i.e., the aggregate transfers in every state following x_L must be the same.[26] This condition can also be written as:

$$w_0 + s_0 = w_1 + s_1, \qquad \Rightarrow \qquad s_0 = w_1 + s_1 - w_0\,, \qquad (1)$$

$$w_\varnothing + s_\varnothing = w_1 + s_1, \qquad \Rightarrow \qquad s_\varnothing = w_1 + s_1 - w_\varnothing. \qquad (2)$$

Since extortion/framing may occur only by suppressing evidence when $\sigma \in \{0, 1\}$, the principal will have to ensure that the contract satisfies two additional extortion/framing deterring (*EF*) constraints to prevent extortion/framing. These can be written as:[27]

$$s_1 \geq s_\varnothing, \qquad (EF_1)$$

$$s_0 \geq s_\varnothing. \qquad (EF_0)$$

Only the (EF_1) constraint is relevant for deterring extortion since it deters the suppression of positive evidence. The constraint (EF_0) deters the suppression of negative information, and bribery is the pertinent issue. Therefore, we will ignore the (EF_0) constraint and verify *ex post* that it is satisfied by our identified solutions in each case below.

---

[25] The *LCCP* contract can be optimal under specific conditions, e.g., if the agent had all the bargaining power when negotiating the side-contract, and if the agent's outside opportunity is high enough (see Section 5).

[26] We consider corruption cost in Section 5.

[27] The (*CIC*) and (*EF*) constraints are written as weak inequalities as it is standard in the literature. The implicit assumption is that, in the case equality, the principal can always break the tie with a small extra payment.

Given the (CIC) and (EF) constraints, corruption is deterred; the agent's (IR) and (IC) and the supervisor's (IR) are the same as those in the incorruptible supervisor case discussed above.

$$\pi u(w_H) + (1 - \pi)\,[pu(w_1) + (1 - p)\,u(w_\varnothing)] - \varphi \geq 0. \qquad (IR)$$

$$\pi u(w_H) + (1 - \pi)\,[pu(w_1) + (1 - p)\,u(w_\varnothing)] - \varphi \geq pu(w_0) + (1 - p)\,u(w_\varnothing) \qquad (IC)$$

Ignoring the (IR) as it is implied by the limited liability and the incentive constraints, we present the principal's program – denoted by P^o – which prevents both bribery and extortion/framing:

Problem P^o:

$$Min\ \pi(w_H) + (1 - \pi)\,[p(w_1 + s_1) + (1 - p)\,(w_\varnothing + s_\varnothing)]$$

s.t. (IC), (1), (2), (EF_1), (EF_0), $w_H \geq 0$, $w_r \geq 0$ and $s_r \geq 0$, where $r \in \{0, \varnothing, 1\}$.

The solution to this problem is the *LCCP contract* and it is characterized in the following lemma:

Lemma 1 *The least-cost-corruption-proof (LCCP) contract has the following features:*

(i) If the supervisor's signal is not very accurate ($p \leq \pi$), the contract is equivalent to the second-best or no-supervisor contract of Section 2.

(ii) If the supervisor's signal is accurate enough ($p > \pi$), it is optimal to use the supervisor, and the contract to the agent satisfies:

$$w_H^o > w_1^o = w_\varnothing^o > 0 = w_0^o\,,$$

where $\dfrac{u'(w_1^o)}{u'(w_H^o)} = \dfrac{1-\pi}{p-\pi}$, $\pi\, u(w_H^o) + (p-\pi)\, u(w_1^o) = \varphi$ *, i.e., the agent obtains an* ex ante *rent.*

- *The supervisor's contract involves*: $s_1^o = s_\varnothing^o = 0 < s_0^o = w_1^o$

  but the supervisor receives no ex ante *rent.*[28]
- *The principal's expected cost is* $C^o = \pi\, w_H^o + (1 - \pi)\, w_1^o$.

Proof: See Appendix B.

---

[28] Since the agent does not shirk in equilibrium, the signal $\sigma = 0$ is off the equilibrium path, and the supervisor's rent is zero even though $s_0 > 0$.

There are two main findings from this lemma: (a) the threat of extortion restricts the principal's ability to use the supervisor's information, and (b) the supervisor will be used only if she is accurate enough. We explain the intuition for these results below.

Take the case when the supervisor's signal is $\varnothing$ and the agent considers bribing her to report evidence of work (r = 1). There are two ways of fighting bribery in our model: (i) pay a high reward ($s_\varnothing = w_1$) to the supervisor, but that would violate (EF_1) as we have seen above; (ii) remove the agent's stake in bribery by paying him a high wage ($w_\varnothing = w_1$). Indeed the binding constraints, (EF_1) and (2), imply that $w_\varnothing = w_1$ to prevent both bribery and extortion. The *LCCP* contract makes use of the second option and rewards the agent even when there is no evidence of work: the agent who shirks without being caught is also treated as if he worked ($w_\varnothing = w_1$), and this increases incentive cost.

Since the agent gets a high wage w_1 (= $w_\varnothing$) with probability $1 - p$ even when he shirks, the supervisor may not be useful unless she is accurate enough. This is different from the incorruptible-supervisor case where she is useful for any $p > 0$. The net effect on the (*IC*) can be seen by setting $w_\varnothing = w_1$ and rearranging terms:

$$\pi u(w_H) + (p - \pi)u(w_1) = \varphi .$$

If $p \leq \pi$, the agent is more likely to receive the transfer w_1 when he shirks rather than when he works, in which case it would be optimal to set $w_1 = 0$. We would have $w_1 = w_\varnothing = w_0 = 0$, and the principal would not rely on the supervisor's report at all. We would also have $s_r = 0$ for all r, and thus, the contract would be equivalent to the second-best contract.

On the contrary, if $p > \pi$, paying a positive w_1 is useful in providing incentive to the agent since he is more likely to receive a positive transfer when he works. Since $w_\varnothing^o = w_1^o$, this is costly to the principal, and therefore it is optimal to set $w_1^o < w_H^o$. The expected cost for the principal is smaller than under the second best, but higher than the case with an incorruptible supervisor.

Note that it is not the supervisor but the agent who benefits from the supervisor's ability to misreport information under the corruption-proof contract. The supervisor cannot affect the agent's payoff by misreporting that $r = \varnothing$ when $\sigma = 1$, and therefore she cannot command any rent. The agent who is the potential victim, on the contrary, obtains a higher utility than his reservation level. Technically, Lemma 1 establishes that $\pi\, u(w_H^o) + (p-\pi)\, u(w_1^o) - \varphi = 0$ and therefore the rent $\pi\, u(w_H^o) + (1-\pi)\, u(w_1^o) - \varphi$ must be positive.

4. The Optimal Contract: Bribery in Equilibrium

In this section we characterize the optimal contract when the supervisor can engage in both types of corruption. From the *LCCP* contract, we know that the supervisor is useful when she is accurate enough, i.e., when $p > \pi$. Therefore, the interesting question is whether it is possible to improve upon the *LCCP* contract by allowing some type of corruption when $p > \pi$, which is the case we analyze in the rest of the paper.[29]

When the agent and supervisor engage in a side contract, their payoffs are determined by the Nash bargaining solution. We define $w_{r\sigma}$ and $s_{r\sigma}$ as the agent's and the supervisor's respective payoffs from Nash bargaining when the signal is σ and the supervisor reports r.

For example, if the agent bribes the supervisor to report work ($r = 1$) when there is no evidence ($\sigma = \varnothing$), the coalition will obtain $s_1 + w_1$ which they will share. This implies that the agent's payoff when $\sigma = \varnothing$ and $r = 1$ is not w_1, the direct payment from the principal, but rather $w_{1\varnothing}$, which is a fraction of $(s_1 + w_1)$. The Nash bargaining problem that determines $w_{1\varnothing}$ and $s_{1\varnothing}$ is given by the following:[30]

$$\max_{w,s} \left(u(w) - u(w_\varnothing)\right)^\alpha \left(s - s_\varnothing\right)^{1-\alpha}$$
$$s.t. \quad w + s = w_1 + s_1$$

where $\alpha \in (0, 1)$ is the agent's bargaining power.

[29] Note that if it is possible to improve on the *LCCP* contract, it will be optimal to use the supervisor even when $p < \pi$, but for high enough p.

[30] In Appendix D we consider the Nash bargaining problems for all values of r and σ.

All the computations, and particularly the agent's (*IC*) constraint, must now be based on the relevant Nash bargaining payoffs instead of the direct transfers from the principal. They are presented in detail in the appendix and we only outline the main intuition here in the text. We first prove that extortion will never be allowed, i.e., the optimal contract must satisfy (EF_1): [31]

Lemma 2: *Any contract that induces e = 1 but violates* (EF_1) *is strictly dominated by the least-cost-corruption-proof (LCCP) contract.*

Proof: See Appendix C.

If the contract violates (EF_1), the threat to extort, e.g., report ∅ when $\sigma = 1$, is credible. Lemma 2 shows that the principal would rather prevent both kinds of corruption by offering the *LCCP* contract than allow extortion to occur. The intuition is that extortion appears as a penalty on the agent after he has done the "right thing" (exerted effort). Since it discourages good behavior, extortion increases the cost of providing incentive.[32]

Given that extortion will not occur in equilibrium, we now argue that the principal can improve on the *LCCP* contract by allowing bribery. We present our main result as Proposition 1 showing that allowing some bribery is indeed optimal, but allowing extortion is not, which is a novel result in the literature.

[31] Polinsky and Shavell (2001) find that, depending on parameter values, it may be optimal to allow extortion/framing and deter bribery. Their model is very different from ours and relies on incorruptible external enforcers to detect corruption. More specifically, the principal can choose different probabilities of detecting bribery, framing, and extortion, and also choose different levels of sanctions for each offence. They also introduce another parameter θ that determines how likely an innocent agent will be in a position to be framed. The relative values of these parameters may make it optimal to deter bribery and allow extortion/framing. For instance, if the parameter θ is very small, then allowing extortion/faming is not very costly, and the principal should focus on deterring bribery.

[32] Technically (see Appendix C) we show that the agent gets the same payoff from Nash bargaining whether σ is ∅ or 1. Therefore, the supervisor's report is not useful in distinguishing between these states and the agent has less incentive to provide effort. The *LCCP* contract dominates such a contract since it also does not distinguish between ∅ and 1 but the supervisor is not rewarded ($s_1 = s_\emptyset = 0$).

Proposition 1: *It is optimal to use the supervisor if p > π. If the agent does not have all the bargaining power (α < 1), bribery occurs when the signal σ = ∅, but the supervisor cannot benefit from extortion and framing. The optimal contract will have the following features:*

- $w_H^* > w_1^* > 0 = w_\varnothing^* = w_0^*$; and $s_1^* = s_\varnothing^* = 0 < s_0^* = w_1^*$.
- *When $\sigma = \varnothing$, the agent's payoff from Nash bargaining is $w_{1\varnothing}^* < w_1^*$, while the supervisor's payoff is $s_{1\varnothing}^*$, with $w_{1\varnothing}^* + s_{1\varnothing}^* = w_1^*$.*
- *The principal's expected cost, denoted by C^*, is given by*

$$C^* = \pi(w_H^*) + (1 - \pi)\, w_1^*.$$

Proof: See Appendix D.

We first explain why bribery may help in providing incentives. One weakness of the *LCCP* contract is that it deters bribery by removing the stake of bribery, setting $w_\varnothing = w_1$. This implies that a shirking agent will obtain a high compensation when the signal is inconclusive about the true effort. To improve incentives, the principal needs to create a gap in the agent's payoffs when $\sigma = \varnothing$ and $\sigma = 1$, rewarding the agent when the signal corroborates a high effort. Setting $w_1 > w_\varnothing$ will create such a gap but also induce bribery.[33] While bribery will reduce the gap, it will not completely remove it (unless the agent has all the bargaining power). In the Nash bargaining solution, the agent and the supervisor share the benefit of misreporting the signal. Since the agent only captures a fraction of the higher w_1, there is a gap between the agent's payoffs when $\sigma = \varnothing$ and 1. Stated differently, even though the agent collects w_1 from the principal whether $\sigma = \varnothing$ or 1, his payoff is lower when $\sigma = \varnothing$ since he has to pay a bribe to the supervisor.

The bribe acts as a penalty on the agent after bad behavior or shirking, and it improves his incentives. This reduces the principal's cost compared to the *LCCP* contract.[34] Bribery provides an indirect way to create a variation in the agent's payoffs,

[33] Lowering $w_\varnothing$ decreases $w_{1\varnothing}$, which helps incentives (see the (*IC*)). Therefore, the principal sets $w_\varnothing = 0$ in equilibrium.

[34] Note that even though bribery improves incentives compared to *LCCP*, it is costly since the principal has to pay w_1 with a positive probability even when the agent shirks.

and reduce incentive cost when direct attempts by the principal would induce extortion. In other words, the principal could not have mimicked the outcome with a corruption-proof contract, which means that Tirole's collusion-proof principle does not apply in our context. Let us explain.

Tirole (1986 and 1992) shows that, under some circumstances, there is no loss of generality to derive an optimal contract that is bribery-proof. Since the principal can anticipate the side contracts between the agent and supervisor, it can provide incentives to deter bribery by replicating the payoffs associated with the side contracts. However, in our context, this intuition does not apply. If the principal tried to replicate the payoffs derived in Proposition 1, it would need to change the transfers to match the payoffs in the Nash bargaining solution and set $w_{\varnothing} = w^*_{1\varnothing}$ and $s_{\varnothing} = s^*_{1\varnothing}$. However, this would introduce extortion when $\sigma = 1$ since $s^*_1 = 0$, and the equilibrium payoffs would differ from those stipulated in the contract.

As noted by Tirole himself (Section 2.5, Tirole (1992)), non-separabilities in the constraints that deter corruption can cause bribery to occur in equilibrium. When these constraints are interlinked, satisfying one constraint raises the cost of satisfying another one and it may be too costly to satisfy them all. In our case it is the interaction between the (*CIC*) and (*EF*) constraints that causes the collusion-proofness principle to fail. With such interlinked-constraints, we show that it is cheaper to allow bribery than to fight it.

This captures well an intuition often mentioned in the applied literature, that allowing bribery can create markets that improve incentives (Bardhan (1997)). Here, the principal relies on the supervisor to extract a bribe from the agent, and lower the agent's payoff in state $\varnothing$, when it cannot directly do so for fear of encouraging extortion. The latter is also consistent with the widely held belief that extortion is always counter productive. Extortion punishes the agent when he has done the right thing, while bribery increases the cost of shirking or violating rules.

It is important to keep in mind, however, that in our model the principal has a limited set of tools to fight corruption, namely, the wage contract. In reality, the principal may have more instruments to deter corruption, e.g., using expensive but more honest enforcers. Our analysis highlights that corruption is multifaceted, and significant tradeoffs ought to be kept in mind when designing strategies to fight the various forms of corruption.

5. Extensions

Agent's bargaining power hurts the principal

When bribery is deterred, the bargaining power of the coalition members does not matter. The principal competes with the agent for the supervisor's report, and the reward given to the supervisor must exceed any viable offer from the agent. In our model the bargaining power is relevant since the principal lets bribery occur in equilibrium. Although the principal cannot affect the relative bargaining power by its choice of contract, the contract influences the bargaining outcomes since it determines the threat points and the pie to be shared.

In this section, we show that the principal is better off when the supervisor has relatively more bargaining power.[35] The reason is that the supervisor can extract a larger bribe from the agent, which makes the bribe a more effective penalty and allows the principal to improve incentives. The agent earns a positive return $w_{1\varnothing}^*$ ($< w_1^*$) from Nash bargaining when $\sigma = \varnothing$, but his return falls as his bargaining power goes down. This implies that the (*IC*) becomes slack, which allows the principal to increase its payoff by adjusting the transfers.

If the agent's bargaining power is reduced down to zero, extortion would not impose any cost on the principal. The principal's payoff is identical to what it would be in the hypothetical case where extortion could be deterred at zero cost.[36] When the agent's bargaining power is zero, his share of w_1 is also zero when $\sigma = \varnothing$. The entire w_1 is taken by the supervisor as a bribe.[37] In the hypothetical case where extortion could be deterred at zero cost, the principal does not have to worry about extortion by assumption

[35] We assume that the agent's bargaining power is common knowledge.
[36] This would be the case if, for example, an agent threatened with extortion has free access to an accurate appeals process.
[37] The agent's payoff is zero when $\sigma = \varnothing$ since $w_\varnothing = 0$.

and can deter bribery by paying a reward $s_{\varnothing} = w_1 > s_1$. There would be no difference between the optimal contract where the agent has zero bargaining power and the optimal contract when extortion could be deterred at zero cost. Thus we conclude that the threat of extortion introduces additional cost on the principal only if the agent has bargaining power.

At the other extreme, if the agent has all the bargaining power, deterring all forms of corruption is optimal. Indeed, allowing bribery in equilibrium has no deterrent effect since the agent gets the entire w_1 when the supervisor misreports. Therefore, the bribe does not create any variation in the agent's payoffs, the *raison d'être* of allowing bribery in the first place. If the agent has all the bargaining power, the principal's payoff is identical to its payoff under the *LCCP* contract where $w_1 = w_{\varnothing}$. Our findings are summarized in Proposition 2:

Proposition 2: *(i) The principal's payoff increases with the supervisor's bargaining power. (ii) At the limit, if the supervisor has all the bargaining power, the principal's payoff is identical to the case where extortion could be deterred at zero cost. (iii) At the other limit, if the agent has all the bargaining power, the principal's payoff is identical to the payoff under the LCCP contract.*

Proof: See Appendix E in the Technical Appendix.

Better outside opportunities make extortion less relevant

Previously we suggested that more developed countries can rely more intensively on hard evidence and therefore suffer less from extortion. In this section, we provide another possible explanation why extortion is less of a problem in more developed countries. We show that if the agent has better outside opportunities, he is less likely to be the target of extortion. The reason is that the wage of an agent with better outside opportunities has to be raised to satisfy the higher reservation utility. With a risk-averse agent, the most efficient way to increase his expected utility is by reducing the variation in the wages on the equilibrium path and relying on a low wage off the equilibrium path to provide incentives. This implies that the agent's wage when the supervisor has no evidence ($w_{\varnothing}$)

increases relatively more than the wages in the other states. Intuitively, a risk averse agent with better outside opportunities is less likely to accept a contract in which he may be punished even though he has worked hard.

For a high enough reservation utility, we show that the agent's wage is made independent of the supervisor's report as long as the report does not reveal shirking ($r = 0$). If the supervisor reveals shirking, the agent is punished with a zero wage. This sanction is relatively more severe when the outside opportunities are high. This could be an explanation for why developing countries with weaker outside opportunities for their workers may suffer more from extortion. Our result is also consistent with the argument that economic agents such as bureaucrats with high salaries are less susceptible to corruption. Often such a claim relies on the decreasing marginal utility of income or an efficiency-wage argument. Our argument is different. In our model, as outside opportunities grow, the agent's wage increases but his rent does not. The supervisor's report can be used to reduce the agent's exposure to risk, provided he works, and extortion becomes less of an issue at the same time. We summarize our result in the proposition below.

Proposition 3: *If the agent's reservation utility is high enough, extortion is not a relevant issue for the principal.*

Proof: See Appendix F in the Technical Appendix.

Technically, we show in Appendix F that the optimal contract derived by only deterring bribery also deters extortion when the reservation utility is high enough. The reason is that an increase in the agent's reservation utility forces an increase in $w_\varnothing$ in order satisfy the (*IR*) constraint. However, such an increase would violate the (*IC*) unless w_H and w_1 are increased as well. The (*CIC*s) require the same total payments in each state so the principal gains by not increasing w_1 at the same rate as $w_\varnothing$ because, by doing so, it can decrease the reward $s_\varnothing$. For a high enough reservation utility, we obtain $w_\varnothing = w_1$, which implies that $s_\varnothing = 0 = s_1$ and extortion ceases to be a relevant threat. The optimal contract is therefore identical to the *LCCP* contract.

Of course, if the reservation utility is increased further, the wages $w_\varnothing = w_1$ are increased to the point where $w_\varnothing = w_1 = w_H$ and the first best is reached. The threat of a large penalty ($w_0 = 0$) if the agent is found shirking is enough to provide the agent an incentive to work.

Generalizing the production technology: possibility of success after low effort

One simplifying assumption of our model was that low effort always yielded a low output. In this section we consider the more general case where low effort can also yield a high output, which corresponds to a situation where the agent can get lucky, and we show that our main results generalize. The main findings are that extortion remains a threat after low output, but it is not relevant after high output. When output is low, bribery is allowed and extortion is deterred, but when output is high, both bribery and extortion are deterred.

We outline the extended model and the intuition and present the technical details in Appendix G in the Technical Appendix. Suppose the likelihood of producing the high output is π_1 when $e = 1$, and π_0 when $e = 0$, where $\Delta\pi = \pi_1 - \pi_0 > 0$. The payments to the agent and supervisor will depend on the output and the supervisor's report, and they are denoted by w_r^j, and s_r^j, where $j \in \{L, H\}$, for the two output levels, and $r \in \{0, \varnothing, 1\}$ are the supervisor's reports.

To grasp the intuition, recall first that, so far, a high output was an absolute guarantee of high effort, but now a high output could result from a low effort by a lucky agent. Therefore, the principal will want to send the supervisor even after high output. Still, the high output is more likely after a high effort than a low effort. Therefore, given a signal $\varnothing$, it is more likely that a high effort was exerted when the output is high compared to when the output is low.[38] Consequently, raising the wage $w_\varnothing^H$ (after high output and report $\varnothing$) helps incentives, whereas raising the wage $w_\varnothing^L$ (after low output and report $\varnothing$) hurts incentives. Thus, when facing the threat of bribery, the principal deters bribery by raising $w_\varnothing^H$ all the way to w_1^H and removes the stake of bribery. This

[38] We assume that the signal $\varnothing$ is equally likely after a high output or low output.

way of fighting bribery does not induce a threat of extortion unlike providing a reward to the supervisor. However, after low output, the principal cannot increase $w_{\varnothing}^{L}$ as it would have a negative incentive effect. The alternative method of fighting bribery, a reward to the supervisor, would introduce a threat of extortion as in our main model. Thus, the principal finds it optimal to allow bribery after low output, and we find that our main result generalizes – a fear of inducing extortion can make bribery optimal.

Honest supervision

As noted earlier, several authors have entertained the possibility that the supervisor is honest, and we briefly discuss the implications of this possibility in our model. Suppose, as in Kofman and Lawarree (1996),[39] that the supervisor is corruptible with a probability less than one, denoted by $c \in [0, 1]$. That is, the supervisor can be one of two types, honest or corruptible. An honest type does not engage in any form of corruption, but a corruptible type would engage in both bribery and extortion if it increases her payoff. In the main text we assumed $c = 1$.

Introducing the possibility of honest supervision increases the effectiveness of monitoring, so it would seem that the principal may no longer need to allow bribery in the optimal contract. This intuition is misleading as it forgets that an honest supervisor is not costly when bribery is allowed. The honest supervisor refuses bribes and reports her true signal. So it is actually less costly to allow bribery when the supervisor is potentially honest. If allowing bribery is already optimal when $c = 1$, it must also be optimal with $c < 1$ since bribery will occur even less frequently. The cost of allowing bribery goes down as the probability of having an honest supervisor goes up.

This is different from the effect of introducing honest types in standard models (Tirole (1986), (1992)) where bribery is deterred in equilibrium. Indeed, Kofman and Lawarree (1996) show that if the likelihood of honest types is large enough, it is better to tolerate bribery than to deter corruption by distorting contracts. Since the threat of extortion is sufficient to make tolerating bribery optimal even when $c = 1$, our main

[39] See also Acemoglu and Verdier (2000) and Auriol (2006). For a dynamic model where the supervisor privately knows her propensity for corruption, see Carrillo (2000b).

result, Proposition 1, is robust to the case where $c < 1$, i.e., it may be optimal to allow bribery even if there are honest supervisors in the population.[40]

Corruption cost

Suppose that the agent and the supervisor suffer a cost when entering into a corrupt agreement. This cost can be, for instance, the psychological burden of participating in an illegal activity (see Tirole (1992)), transaction costs (Faure-Grimaud et al. (2002)) or the fear of being detected in the future (Khalil and Lawarree (2006)). Introducing a *corruption cost* for the agent and the supervisor makes it easier for the principal to deter both forms of corruption. However, our optimal contract is robust to the presence of small corruption cost.

We capture corruption cost with a parameter $\gamma \in (0, 1)$, such that the agent and the supervisor only share γw_1 when reporting $r = 1$ after $\sigma = \varnothing$.[41] On the one hand, allowing bribery is now more effective because the cost of corruption creates an even larger gap between the agent's payoffs when $\sigma = 1$ or $\varnothing$. On the other hand, allowing bribery remains costly because the principal pays w_1 to the agent regardless of whether $\sigma = 1$ or $\varnothing$. This suggests that, when γ is small, it is possible that the principal could do better by deterring bribery. Let us explain. Even though the principal pays w_1, the supervisor and the agent now bargain over the smaller pie γw_1. To deter bribery the principal must pay $w_\varnothing = \gamma w_1$ when $\sigma = \varnothing$, which satisfies ($CIC_{\varnothing 1}$). For γ very small, the value of $w_\varnothing$ that deters bribery is also very small, and the principal is better off deterring bribery. The optimal contract deters corruption and is similar to the honest supervisor contract for a high enough corruption cost.

---

[40] As a technical aside, note that the proof of Proposition 1 would have to be significantly modified. When $c = 1$, we are able to use the *LCCP* contract to show that allowing extortion is not optimal. With $c < 1$, the principal immediately benefits from the presence of an honest type, and we cannot construct a corruption-proof contract that would mimic the cost of a contract that violates (EF_1). However, we can use the optimal contract to show that it dominates any contract that induces extortion.

[41] Recall that $s_1 = 0$ in Proposition 1.

6. Conclusion

In this paper, we have highlighted key differences between bribery and extortion. One difference is that extortion has a worse implication on the agent's incentives. Another difference is that bribery involves cooperation between corrupt parties, while extortion is antagonistic. This suggests that extortion should be easier to fight by exploiting the difference in objectives. Both distinctions point to making extortion the prime focus in a fight against corruption.

The fight against extortion could take the form of an efficient and transparent appeals process if an innocent agent has a credible recourse against an extortion attempt. However, as noted commentators have pointed out, detection of extortion can be difficult, for example, because extortion reports may be seen as malevolent.[42] Without an appeals process, the principal may have to induce bribery to create the desired variation in the agent's payoffs as in our model.

We show that there is a tradeoff between fighting bribery and extortion, but our analysis provides several strategies for reducing the threat of extortion. It is important to underline that the tradeoff only appears if information is soft. If information is hard, there is no such tradeoff and bribery is deterred in equilibrium. Our results suggest that organizations that must rely on soft information may also need to allow bribery.[43] By making its information "harder" an organization will suffer less from corruption, but making information harder can be costly. For instance, speeding tickets should rely on

[42] Furnivall (1956) studying bribery and extortion in Burma noted, "Those who gained their ends by bribery naturally made no complaint, and complaints from those who suffered were suspect as malicious. Such evidence as was available mostly came from people who had given bribes and, as accomplices, their evidence, even if admissible, was doubtful. It was difficult and dangerous for any private individual to set the law in motion, and in practice this was hardly possible except by some local or departmental superior of the man suspected of corruption." Klitgaard (1988) discussing tax-assessor extortion noted that the appeals process is not straightforward: "In one of the most notorious versions [of extortion] a tax assessor would slap an unrealistically high assessment on the taxpayer. The taxpayer could appeal, but that would take time and effort; furthermore, the taxpayer might not be sure what the 'correct' tax really was."

[43] While there are many reported examples of explicit bribery in the media, an interesting example of allowing collusion/bribery in organizations is a leniency bias in job performance appraisal. Our result provides one rationale for why many organizations which use job performance appraisal as an incentive device may allow a leniency bias. See Bretz et. al (1992) for a survey on studies related to this issue, and Johnson and Liebcap (1989) for an example of leniency in the federal government.

sophisticated cameras or shareholders ought to be able to appeal auditing reports to reliable and incorruptible experts. Developing countries with less resources and technological abilities, and weak legal environment also have less capability to make information hard and, therefore, we should expect that bribery to be a more pervasive problem. The fight against corruption should therefore emphasize the need to rely on hard evidence. Similarly, extortion can be thwarted and both types of corruption be deterred if the agent's outside option is high, or if the agent's bargaining strength is high. Neither is likely to be true in a poor country with a weak appeals process.

Tirole's collusion-proofness principle does not apply in our model because rewards to deter bribery increase the cost of deterring extortion – the bribery and extortion constraints are inter-linked. One implication of bribery occurring in equilibrium is to validate in a model the popular notion that bribery can be useful to "grease the wheels" in inefficient organizations. However, this is only a second-best result – bribery is optimal in our model because it allows the principal to cause a variation in the agent's payoffs when direct payments would only have resulted in introducing extortion, which is a worse problem. Extortion penalizes an agent after "good" behavior, while bribery at least imposes some penalty for "bad" behavior.

Appendices:

Proofs of Case III of Proposition 1, and Appendices E, F and G can be found in a Technical Appendix at *www.rje.org/sup-matl.html*.

Appendix A Incorruptible Supervisor

Suppose the supervisor always reports truthfully what she has observed. The agent's participation and incentive constraints are

$$\pi u(w_H) + (1-\pi)\,[pu(w_1) + (1-p)\,u(w_\varnothing) - \varphi \geq 0, \qquad (IR)$$

$$\pi u(w_H) + (1-\pi)\,[pu(w_1) + (1-p)\,u(w_\varnothing) - \varphi \geq\; pu(w_0) + (1-p)\,u(w_\varnothing). \qquad (IC)$$

The incentive constraint (IC) can be rewritten as $\pi u(w_H) + (1-\pi)\,pu(w_1) - \pi(1-p)\,u(w_\varnothing) - pu(w_0) \geq \varphi$. Given limited liability and the (IC), we can ignore both the agent's and the supervisor's participation constraints in each case we consider.

The principal's program when the supervisor is truthful can be written as follows:

$Min \quad \pi(w_H) + (1-\pi)\,[p(w_1 + s_1) + (1-p)\,(w_\varnothing + s_\varnothing)]$

s.t. $\quad$ (IC), $w_H \geq 0$, $w_r \geq 0$ and $s_r \geq 0$, where $r \in \{0, \varnothing, 1\}$.

The optimal levels of transfers are obtained using standard techniques.

Appendix B Proof of Lemma 1

In the Problem P^o of Section 3, we will first ignore the constraint (EF_0) and verify later that it is satisfied by the optimal contract. Using (2) to replace $s_\varnothing$ everywhere, we can rewrite (EF_1) as $w_\phi \geq w_1$. The principal's problem is:

$Min \quad \pi w_H + (1-\pi)\,(w_1 + s_1)$

s.t. $\quad$ (IC), (1), $w_\phi \geq w_1$ and the non-negativity constraints.

Note that once we ignore (EF_0), the variable s_0 does not appear anywhere else in the problem except in (1). Therefore, we are free to choose s_0 to satisfy this constraint (1) as long as $s_0 \geq 0$. Again using standard techniques, we can derive the optimal levels of transfers reported in Lemma 1. ■

Appendix C Proof of Lemma 2

Consider Problem P^o of Section 3, but without imposing the (*CIC*), i.e., constraints (1) and (2). We also ignore (EF_0) for now and verify later that it is indeed satisfied.

We proceed in steps. In Step (i) we show that the agent receives the same *payoff* from Nash bargaining for $\sigma \in \{\emptyset, 1\}$ if the constraint (EF_1) is violated, but the supervisor earns an *ex ante* rent. In Step (ii) we show that the principal can achieve the same cost with a corruption-proof contract. This (constructed) corruption-proof contract is more costly than the *LCCP* contract. This proves the claim.[44]

Define T_k: $T_k = w_k + s_k$ for $k = \{0, \emptyset, 1\}$. The agent-supervisor coalition will choose the report to maximize their joint payoff, and we define m by $T_m = \max\{T_0, T_\emptyset, T_1\}$. Then define $w_{r\sigma}$ and $s_{r\sigma}$ as the agent's and the supervisor's respective payoffs (from Nash bargaining where relevant) when the signal is σ and the supervisor reports r.

<u>Step (i)</u> If (EF_1) is violated, i.e., $s_1 < s_\emptyset$, then the agent gets identical payoffs for $\sigma = \emptyset$ or $\sigma = 1$; the same is true for the supervisor:

---

[44] Although we rely on the axiomatic approach, our bargaining outcome can be related to the outcome of a strategic alternative-offers-bargaining model with a risk of breakdown, where time between offers are very small (see e.g., Binmore et al. (1986) or Osborne and Rubinstein (1990)). In our model, the bargaining is about negotiating an illegal side contract which is fraught with uncertainty, e.g., opportunities to interact may disappear abruptly or the principal may require an early report. Therefore, the exogenous risk of a breakdown in negotiations will be the dominant force that drives the parties to an agreement. Moreover, since the principal will set a short deadline for a report from the supervisor, discount rates play a minor role compared to the fear of a breakdown. In such a case, it is appropriate to choose the breakdown point as the disagreement or threat point in the Nash bargaining solution as we have done in the paper (See Binmore et al. (1986) p.183 or Osborne and Rubinstein (1990) p.88). Finally note that the outside option, which acts as a constraint on the bargaining set, is not binding in this model. The supervisor will have to make a report even when the parties fail to agree and, therefore, the outside option and the break down point are identical.

(a) If $T_m = T_\varnothing$: Given $s_1 < s_\varnothing$, the supervisor will report $r = \varnothing$ when $\sigma = \{\varnothing, 1\}$, and the agent will not find it profitable to bribe the supervisor into announcing $r = 1$. Therefore, payoffs will be: $w_{m1} = w_{m\varnothing} = w_\varnothing$; $s_{m1} = s_{m\varnothing} = s_\varnothing$.

(b) If $T_m > T_\varnothing$: For $\sigma = \{\varnothing, 1\}$, the supervisor reports $r = m \in \{0, 1\}$ since bribery may occur, and the coalition receives T_m. Their payoffs are given by Nash bargaining. Given that only the supervisor reports, the threat point is $r = \varnothing$ for $\sigma \in \{\varnothing, 1\}$ since $s_1 < s_\varnothing$. The Nash bargaining problem is given by

$$\max_{w,s} \left(u(w) - u(w_\varnothing)\right)^\alpha \left(s - s_\varnothing\right)^{1-\alpha}$$
$$s.t. \quad w + s = T_m,$$

where $\alpha \in (0, 1)$ is the agent's bargaining power.[45]

The solution is denoted by $w_{m\sigma}$ and $s_{m\sigma}$ for $\sigma \in \{\varnothing, 1\}$. Whether $\sigma = \varnothing$ or 1, the threat point and T_m remain unchanged. Therefore, the bargaining problem remains unchanged, and the Nash bargaining payoffs for the agent and the supervisor must also be identical whether $\sigma = \varnothing$ or 1. They are: $w_{m1} = w_{m\varnothing}$; $s_{m1} = s_{m\varnothing} > 0$ since $s_\varnothing > s_1 \geq 0$.

Therefore, from (a) and (b), we have proved that $w_{m1} = w_{m\varnothing}$ regardless of m.

<u>Step (ii)</u>: Next we compute the expected cost of any contract that violates (EF_1). Consider the contract denoted by $\{\hat{w}_H, \hat{w}_r, \hat{s}_r\}$ that induces $e = 1$, but violates (EF_1), i.e., $\hat{s}_\varnothing > \hat{s}_1$. Then the expected cost is: $\pi(\hat{w}_H) + (1 - \pi)(\hat{T}_m)$ where $\hat{T}_m = max\{\hat{T}_0, \hat{T}_\varnothing, \hat{T}_1\}$,and $\{\hat{w}_H, \hat{w}_r, \hat{s}_r\}$ satisfy the (IC) constraint:

$$\pi\, u(\hat{w}_H) + (1 - \pi)\{p\, u(\hat{w}_{m1}) + (1 - p)\, u(\hat{w}_{m\varnothing})\} - \varphi \geq p\, u(\hat{w}_{m0}) + (1 - p)\, u(\hat{w}_{m\varnothing}). \qquad (IC)$$

[45] In the strategic alternative offers bargaining model, values of $\alpha \neq ½$ can still be justified by asymmetry in the bargaining procedure or in the beliefs about the likelihood of breakdown (Binmore et al. (1986), p. 187).

Given step (i), we can define $\hat{W}_m = \hat{w}_{m1} = \hat{w}_{m\varnothing}$, $\hat{S}_m = \hat{s}_{m1} = \hat{s}_{m\varnothing}$ and simplify the (*IC*):[46]

$$\pi\, u(\hat{w}_H) + (p - \pi)\, u(\widehat{W}_m) - \varphi \geq p\, u(\hat{w}_{m0}) \qquad (IC)$$

Note that $\hat{S}_m > 0$ since the supervisor receives at least $\hat{s}_\varnothing$ from Nash bargaining and $\hat{s}_\varnothing > \hat{s}_1 \geq 0$.

(iii) We construct a corruption-proof contract $\{w'_H, w'_r, s'_r\}$ that has the same expected cost as $\{\hat{w}_H, \hat{w}_r, \hat{s}_r\}$ and also implements $e = 1$. Construct $\{w'_H, w'_r, s'_r\}$ by defining: $w'_H = \hat{w}_H$, $w'_1 = w'_\varnothing = \widehat{W}_m$, $w'_0 = 0$, $s'_1 = s'_\varnothing = \hat{S}_m$, and $s'_0 = \hat{T}_m$.

Next, we check that $\{w'_H, w'_r, s'_r\}$ is indeed corruption-proof and implements $e = 1$:

(*CIC*) is satisfied since $w'_k + s'_k = \hat{T}_m$, $k \in \{0, \varnothing, 1\}$; ($EF_k$) is satisfied since $s'_k \geq s'_\varnothing$, $k \in \{0, 1\}$; and (*IC*) is satisfied by construction of w'_k. To see this, recall that $\hat{w}_k$ satisfies (*IC*), where $k \in \{H, m0, m\varnothing, m1\}$, and that $w'_0 \leq \hat{w}_{m0}$ from the Nash bargaining solution. Finally, note that $\{w'_H, w'_r, s'_r\}$ is different from the *LCCP* contract since $\hat{S}_m > 0$, whereas in *LCCP* contract $s^0_1 = s^0_\varnothing = 0$. Since the *LCCP* contract is the unique optimum among corruption-proof contracts, it strictly dominates both $\{w'_H, w'_r, s'_r\}$ and $\{\hat{w}_H, \hat{w}_r, \hat{s}_r\}$. ■

Appendix D Proof of the Proposition 1

Refer to Problem P^o in Section 3, but without imposing the (*CIC*), i.e., constraints (1) and (2). The agent-supervisor coalition will choose the report to maximize their joint payoff, which will be T_m. Since bribery may occur, the objective function becomes $\pi\, w_H + (1 - \pi)\, T_m$. We consider three cases depending on whether $m = 1, \varnothing$, or 0 respectively. We show that cases I and II are identical involving bribery, while case II yields the *LCCP*

[46] Note that $\hat{s}_0$ could be larger or smaller than $\hat{s}_\varnothing$ – both cases are captured in $\hat{w}_{m0}$.

contract. We finally prove that allowing bribery strictly dominates the *LCCP* contract when the agent's bargaining power is less than 1.

The (*IC*) constraint is: $\pi\, u(w_H) + (1-\pi)\, p\, u(w_{m1}) - \pi(1-p)\, u(w_{m\varnothing}) - p\, u(w_{m0}) - \varphi \geq 0$.

We ignore the constraint (EF_0) for now and verify later that it is indeed satisfied by the optimal contract.

Case I: $T_m = T_1$ $\qquad$ *Min* $\pi w_H + (1-\pi)\, T_1$ s.t.

$$\pi\, u(w_H) + (1-\pi)\, p\, u(w_{11}) - \pi(1-p)\, u(w_{1\varnothing}) - p\, u(w_{10}) - \varphi \geq 0 \qquad (IC)$$

$$s_1 \geq s_\varnothing, \qquad (EF_1)$$

and the non-negativity constraints on the transfers.

We begin with observations that characterize part of the solution.

(a) $w_{11} = w_1$ and $s_{11} = s_1$: Since $s_1 \geq s_\varnothing$, the supervisor's threat point is to report 1 when $\sigma = 1$. Given that $T_m = T_1$, the Nash Bargaining Solution (NBS) implies that $s_{11} = s_1$, and $w_{11} = w_1$.

(b) $T_0 = T_1$, and $w_{10} = w_0 = 0$: To see this, note that w_0 and s_0 only appear in (*IC*) through w_{10}. By setting $s_0 = T_1$ and $w_0 = 0$ the principal makes the agent's payoff in the NBS $w_{10} = 0$, and this does not cost the principal anything since s_0 does not appear in the objective function. Given that $s_0 = T_1$ and $w_0 = 0$, we have $T_0 = T_1$.

Since $s_0 = T_1$, we have $s_0 \geq s_\varnothing$, and (EF_0) is satisfied.

(c) $w_\varnothing = 0$: To see this, note that $w_\varnothing$ does not appear in the objective function and enters only the (IC) through $w_{1\varnothing}$ via the threat-point payoff of the agent in the Nash bargaining problem. The Nash bargaining problem that determines $w_{1\varnothing}$ and $s_{1\varnothing}$ is given by

$$\max_{w,s} \left(u(w) - u(w_\varnothing)\right)^\alpha \left(s - s_\varnothing\right)^{1-\alpha}$$
$$s.t. \quad w + s = w_1 + s_1$$

Lowering $w_{\varnothing}$ decreases $w_{1\varnothing}$, which helps incentives (see the (*IC*)). Therefore, the principal will set $w_{\varnothing} = 0$ in equilibrium.

(d) (EF_1) is binding ($s_1 = s_{\varnothing}$): Given (a), $w_{11} = w_1$, and therefore s_1 enters directly in the objective function and indirectly in the (*IC*) through $w_{1\varnothing}$. Since a decrease in s_1 reduces $w_{1\varnothing}$, for a given $s_{\varnothing}$, the principal will lower s_1 until (EF_1) binds, and thus $s_1 = s_{\varnothing}$.

(e) $s_1 = s_{\varnothing} = 0$: In the Nash bargaining problem, $s = s_1 + w_1 - w$. Since $s_{\varnothing} = s_1$ and $w_{\varnothing} = 0$, the bargaining problem becomes $max\ (u(w))^{\alpha}\,(w_1 - w)^{1-\alpha}$, and its solution $w_{1\varnothing}$ is independent of s_1. Therefore, s_1 ($= s_{\varnothing}$) can be reduced to zero to minimize the objective function without affecting the incentive constraint. Thus, $s_1 = s_{\varnothing} = 0$.

Given (a) – (e) and the binding (*IC*) constraint, we can write the Lagrangian as follows:

$$L = \pi w_H + (1-\pi)\, w_1 - \lambda\,[\,\pi u(w_H) + (1-\pi)\, p\, u(w_1) - \pi (1-p)\, u(w_{1\varnothing}) - \varphi]$$

$$\partial L/\partial w_H = \pi - \lambda\, \pi u'(w_H) = 0 \qquad \text{(d1)}$$

$$\partial L/\partial w_1 = (1-\pi) - \lambda[(1-\pi)\, p\, u'(w_1) - \pi\,(1-p)\, u'(w_{1\varnothing})\, dw_{1\varnothing}/dw_1\,] = 0 \qquad \text{(d2)}$$

From (d1), $u'(w_H) = 1/\lambda$; from (d2), $u'(w_1) = 1/(\lambda p) + [\pi(1-p)]/[(1-\pi)p]\ u'(w_{1\varnothing})$ $[\,dw_{1\varnothing}/dw_1\,]$.

Since the bargaining set becomes larger as w_1 increases, we have $dw_{1\varnothing}/dw_1 > 0$, and therefore $u'(w_H) < u'(w_1)$, which implies $w_H > w_1$. The solution is such that $w_H > w_1 > 0 = s_1 = s_{\varnothing} = w_{\varnothing} = w_0$ and $s_0 = w_1 = T_1$. Note that the (*CIC*) is violated when $\sigma = \varnothing$ – the coalition is strictly better off by reporting $r = 1$ or $r = 0$.

<u>Case II: $T_m = T_{\varnothing}$</u> $\qquad$ $Min\ \pi w_H + (1-\pi)\, T_{\varnothing}$ s.t.

$$\pi u(w_H) + (1-\pi)\, p\, u(w_{\varnothing 1}) - \pi (1-p)\, u(w_{\varnothing\varnothing}) - p\, u(w_{\varnothing 0}) - \varphi \geq 0 \quad (IC)$$

$s_1 \geq s_\varnothing,$ (EF_1)

and the non-negativity constraints on the transfers.

We begin with observations that characterize part of the solution.

(a) $w_{\varnothing\varnothing} = w_\varnothing$ and $s_{\varnothing\varnothing} = s_\varnothing$: The supervisor's threat point is to report $\varnothing$ when $\sigma = \varnothing$. Given that $T_m = T_\varnothing$, the NBS implies that $s_{\varnothing\varnothing} = s_\varnothing$, and $w_{\varnothing\varnothing} = w_\varnothing$.

(b) $w_\varnothing \geq w_1$: To see this, note that $T_\varnothing \geq T_1$ and $s_1 \geq s_\varnothing$.

(c) $s_0 = T_\varnothing$, $w_0 = 0$, and $w_{\varnothing 0} = 0$: To see this, note that s_0 and w_0 only appear in (IC) through $w_{\varnothing 0}$. By setting $s_0 = T_\varnothing$ and $w_0 = 0$, the principal makes the agent's payoff in the NBS $w_{\varnothing 0} = w_0 = 0$ and this does not cost the principal anything since s_0 does not appear in the objective function. Given $s_0 = T_\varnothing$ and $w_0 = 0$, we have $T_0 = T_\varnothing$. Note also that (EF_0) is satisfied since $s_0 = T_\varnothing \geq s_\varnothing$.

(d) $w_1 = w_\varnothing$: To see this, note that w_1 only appears in the (IC) through $w_{\varnothing 1}$ via the threat point payoff of the agent. Therefore the principal can increase $w_{\varnothing 1}$ and relax the (IC) by increasing w_1. Since $w_\varnothing \geq w_1$ from (b), w_1 will be increased until $w_1 = w_\varnothing$.

(e) EF_1 is binding ($s_1 = s_\varnothing$): To see this, note that s_1 only enters the (IC) through $w_{\varnothing 1}$. The principal can increase $w_{\varnothing 1}$ by reducing s_1 since s_1 is the threat-point payoff of the supervisor. A decrease in s_1 increases $w_{\varnothing 1}$, which helps incentives. Therefore, the principal reduces s_1 until (EF_1) binds and thus $s_1 = s_\varnothing$.

(f) $w_{\varnothing 1} = w_\varnothing = w_1$: To see this, note that $s_1 = s_\varnothing$, $w_1 = w_\varnothing$ and $T_1 = T_\varnothing$.

(g) $s_\varnothing = 0$: given that $w_{\varnothing 0} = 0$, $s_\varnothing$ only appears in the objective function and therefore will be reduced to zero.

Given (a) – (g), we can rewrite the minimization problem as

$Min\ \pi w_H + (1-\pi) w_1$ s.t. $(IC)\quad \pi u(w_H) + (p-\pi) u(w_1) - \varphi \geq 0$

And the Lagrangian is: $L = \pi w_H + (1-\pi) w_1 - \lambda [\, \pi u(w_H) + (p-\pi) u(w_1) - \varphi]$.

The FOCs give the optimal w_H and w_1 for case II:

$$\partial L/\partial w_H = \pi - \lambda\, \pi u'(w_H) = 0 \qquad \text{(d3)}$$

$$\partial L/\partial w_1 = (1-\pi) - \lambda (p-\pi) u'(w_1) = 0 \qquad \text{(d4)}$$

The conditions (d3) and (d4) characterize the optimal w_H and w_1 in case II, and all the transfers are identical to those reported in Lemma 1. Therefore, we have shown that the optimal contract under case II is the *LCCP* contract.

The optimal contract: In the Technical Appendix on the website, we show that the solution to Case III, where $T_m = T_0$, is identical to case I. In other words, the coalition as well as the principal are indifferent between the reports $r = 1$ or $r = 0$ when $\sigma = \varnothing$. In the paper, we focus on the case where the report is $r = 1$ to provide all our intuition.

Thus, it is sufficient to compare cases I and II to determine the optimal contract. Note that the expected cost in each case is $\pi w_H + (1-\pi) w_1$ but the (*IC*) constraints differ:

(*IC*) in Case I $\quad \pi u(w_H) + (1-\pi)\, p\, u(w_1) - \pi (1-p)\, u(w_{1\varnothing}) - \varphi = 0$

(*IC*) in Case II $\quad \pi u(w_H) + (p-\pi) u(w_1) - \varphi = 0$.

Since Nash bargaining implies $w_{1\varnothing} < w_1$ for $\alpha < 1$, the lowest expected cost under case II can be achieved under case I with a slack (*IC*). Therefore, the optimal contract under case I results in a smaller expected cost than case II, and we have proved that bribery will occur when $\sigma = \varnothing$. ■

REFERENCES

Andrianova, S. and Melissas N. "Corruption, Extortion, and the Boundaries of the Law." *Journal of Law, Economics, and Organization,* (forthcoming).

Acemoglu, D. and Verdier T. "The Choice between Market Failures and Corruption." *American Economic Review*, Vol. 90 (2000), pp. 194-211.

Auriol, E. "Corruption in Procurement and Public Purchase." *International Journal of Industrial Organization*, Vol. 24 (2006), pp. 867-885

Ayres, I. "The Twin Faces of Judicial Corruption: Extortion and Bribery." *Denver University Law Review*, Vol. 74 (1997), pp. 1231-1254.

Baker, G. and Hubbard, T. "Make versus Buy in Trucking: Asset Ownership, Job Design, and Information." *American Economic Review*, Vol. 93 (2003), pp. 551-572.

Baliga, S. "Monitoring and Collusion with Soft Information." *Journal of Law, Economics and Organization*, Vol. 15 (1999), pp. 434-440.

Banerjee, A., "A Theory of Misgovernance", *Quarterly Journal of Economics*, Vol. 112(4) (1997), pp. 1289-1332.

Bardhan, P. "Corruption and Development." *Journal of Economic Literature*, Vol. 35 (1997), pp. 1320-1346.

Binmore, K.G., Rubibstein, A. and Wolinsky, A. "The Nash Bargaining Solution in Economic Modelling.", *Rand Journal of Economics*, Vol. 17(1986), pp.176-188.

Bretz, R.D., Milkovich, G. T. and Read, W. "The Current State of Performance Appraisal Research and Practice: Concerns, Directions, and Implications." *Journal of Management*, Vol. 18 (1992), pp. 321–332.

Cadot, O. "Corruption as a Gamble." *Journal of Public Economics*, Vol. 33 (1987), pp. 22 - 44.

Caillaud, B. and Tirole J. "Consensus Building: How to Persuade a Group." *American Economic Review*, Vol. 97 (2007), pp. 1877-1900.

Carrillo, J. "Corruption in Hierarchies." *Annales d'Economie et de Statistique*, Vol. 59 (2000a), pp. 37-62.

Carrillo, J. "Graft, Bribes, and the Practice of Corruption." *Journal of Economics and Management Strategy*, Vol. 9 (2000b), pp. 257-286.

Che, Y.-K. “Revolving Doors and the Optimal Tolerance for Agency Collusion.” *RAND Journal of Economics*, Vol. 26 (1995), pp. 378-397.

Che, Y.-K. and Kim, J. “Robustly Collusion-proof Implementation.” *Econometrica*, Vol. 74 (2006), pp. 1063-1107.

Colb, S. “Stopping a Moving Target.” *Rutgers Race & the Law Review*, Vol. 3 (2001), (Symposium Article).

Dewatripont, M. and Tirole J. “Modes of Communication.” *Journal of Political Economy*, Vol. 113 (2005), pp. 1217–1238.

Faure Grimaud, A., Laffont, J.-J. and Martimort, D. "Risk Averse Supervisors and the Efficiency of Collusion." *Contributions to Theoretical Economics*, Vol. 2 (2002), pp. 1055-1055.

Faure Grimaud, A., Laffont, J.-J. and Martimort, D. “Collusion, Delegation and Supervision with Soft Information.” *Review of Economic Studies*, Vol. 70 (2003), pp. 253-279.

Furnivall, J.S. *Colonial Policy and Practice; A Comparative Study of Burma and Netherlands India*. New York: New York University Press, 1956.

Guriev, S., “Red Tape and Corruption”, *Journal of Development Economics*, Vol. 73 (2004), pp. 489-504.

Hindriks, J., Keen, M. and Muthoo, A. “Corruption, Extortion and Evasion.” *Journal of Public Economics*, Vol. 74 (1999), pp. 395–430.

Johnson R. and Libecap G. “Bureaucratic Rules, Supervisor Behavior, and the Effect on Salaries in the Federal Government.” *Journal of Law, Economics, and Organization*, Vol. 5(1989), pp. 53-82.

Kessler, A.S. “On Monitoring and Collusion in Hierarchies.” *Journal of Economic Theory*, Vol. 91 (2000), pp. 280-291.

Khalil, F. and Lawarrée, J. “Incentives for Corruptible Auditors in the Absence of Commitment.” *Journal of Industrial Economics*, Vo. 54 (2006), pp. 269–291.

Klitgaard, R. *Controlling Corruption*, Berkeley: University of California Press, 1988.

Kofman, F. and Lawarrée, J. “Collusion in Hierarchical Agency.” *Econometrica*, Vol. 61 (1993), pp. 629-656.

Kofman, F. and Lawarrée, J. "On the Optimality of Allowing Collusion." *Journal of Public Economics*, Vol. 61 (1996), pp. 383-407.

Laffont, J.-J., and Martimort, D. "Collusion under Asymmetric Information." *Econometrica*, Vol. 65 (1997), 875–911.

Laffont, J.-J., and Martimort, D. "Mechanism Design with Collusion and Correlation." *Econometrica*, Vol. 68 (2000), 309–342.

Laffont, J-J. and Tirole J. *A Theory of Incentive in Procurement and Regulation*, Cambridge, Mass.: MIT Press, 1993.

Lambert-Mogiliansky, A. "On Optimality of Illegal Collusion in Contracts." *Review of Economic Design*, Vol. 3 (1998), pp 303-328.

La Porta, R., Lopez-de-Silanes, F., Shleifer A. and Vishny, R. "Investor Protection and Corporate Governance." *Journal of Financial Economics*, Vol. 58 (2000), pp.3-27.

Lindgren, J. "The Theory, History and Practice of the Bribery-Extortion Distinction," *University of Pennsylvania Law Review*, Vol. 141 (1993), pp. 1695-1740.

Mishra, A. *The Economics of Corruption,* New York: Oxford University Press, 2005.

Mookherjee, D. "Incentive Reforms in Developing Country Bureaucracies: Lessons from Tax Administration." In Pleskovic, B. and Joseph Stiglitz, J. eds. *Proceedings of the Annual Bank Conference On Development Economics*, Washington D.C.: World Bank, 1997.

Mookherjee, D. and Png, I. "Corruptible Law Enforcers: How Should They Be Compensated?" *Economic Journal*, Vol. 105 (1995), pp. 145-59.

Olsen, T.E. and Torsvik, G. "Collusion and Renegotiation in Hierarchies: A Case of Beneficial Corruption." *International Economic Review*, Vol. 39 (1998), pp. 413-438.

Osborne, K.J. and Rubinstein, A. *Bargaining and Markets*. San Diego: Academic Press, 1990.

Polinsky, A.M. and Shavell, S. "The Economic Theory of Public Enforcement of Law." *Journal of Economic Literature*, Vol. 38(2000), pp. 45-76.

Polinsky, A.M. and Shavell, S. "Corruption and Optimal Law Enforcement." *Journal of Public Economics*, Vol. 81 (2001), pp. 1-24.

Shavell, S. *Foundations of Economic Analysis of Law*. Cambridge, Mass: Harvard University Press, 2004.

Shin, D. "Contracts under Wage Compression: A Case of Beneficial Collusion." *Southern Economic Journal*, Vol. 74 (2007), pp 143–157.

Silva, E., Kahn, C. and Zhu, Z. "Crime and Punishment and Corruption: Who Needs 'Untouchables'?", *Journal of Public Economic Theory*, Vol. 9 (2007), pp. 69-87.

Strausz, R. "Collusion and Renegotiation in A Principal-Supervisor-Agent Relationship." *The Scandinavian Journal of Economics*, Vol. 99 (1997a), pp. 497-518.

Strausz, R. "Delegation of Monitoring in a Principal Agent Relationship." *Review of Economic Studies*, Vol. 64 (1997b), pp. 337-357.

Tirole, J. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics, and Organization*, Vol. 2 (1986), pp. 181-214.

Tirole, J. "Collusion and the Theory of Organizations." In Laffont, J.-J., eds., *Advances in Economic Theory*, *Sixth World Congress*, Vol.2. New York: Cambridge University Press, 1992.

Vafai, K. "Collusion and Organization Design." *Economica*, Vol. 72 (2005), pp.17-37.