UniS

University of Surrey

Discussion Papers in Economics

BENEFICIAL COLLUSION IN CORRUPTION CONTROL: THE CASE OF NONMONETARY PENALTIES

By

Mehmet Bac
(Sabanci University)
&
Parimal Kanti Bag
(University of Surrey)

DP 02/05

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk

Beneficial Collusion in Corruption Control: The Case of Nonmonetary Penalties

Mehmet Bac* Parimal Kanti Bag†

April 25, 2005

Abstract

We analyze a corruption model where a principal seeks to control an agent's corruption by supplementing a costless noncollusive outside detector such as the media with a collusive internal supervisor. The principal's objective is to minimize the overall costs, made up of enforcement costs and social costs of corruption. If the penalties on the corrupt agent and a failing supervisor are nonmonetary in nature and yet the two parties can engage in monetary side-transfers, the principal may stand to benefit by allowing supervisor-agent collusion. This benefit may even prompt the principal to actively *encourage* collusion by hiring a dishonest supervisor in strict preference over an honest supervisor. *JEL* Classification Numbers: K42, D73, D78.

Key Words: Corruption, monitoring, collusion, bounty hunter mechanism.

*Sabanci University, Faculty of Arts and Social Sciences, Tuzla, Istanbul 81474, Turkey. E-mail: bac@sabanciuniv.edu

†Department of Economics, University of Surrey, Guildford, Surrey GU2 7XH, U.K.; E-mail: p.bag@surrey.ac.uk

1 Introduction

Bribes are rarely entirely detected by internal control mechanisms of government offices. Often they remain hidden until, if ever, a whistle is blown by outside detectors such as the media. Detection of corruption mainly by outside sources suggests ineffectiveness of internal control mechanisms and/or potential presence of collusion within the system, that is, an agreement between internal supervisors and bribe-taking bureaucrats whereby the latter transfer part or all of their bribe collection to the former when detected to avoid being reported in return. In this paper we develop a three-layer hierarchy model of a public office to study the role of collusion in controlling corruption.[1] We identify an environment in which collusion can be beneficial for the principal.

The institutional set-up of our model is in the tradition of Tirole (1986, 1992) and closer to Kofman and Lawarrée (1993), who study a principal-supervisor-agent hierarchy where the supervisor can collude with the agent and underreport the agent's true productivity (or type) as a cover for bad performance due to low effort. They also introduce a second, incorruptible external auditor which in our context is costless and represents sources of detection such as the media. In Kofman and Lawarrée's optimal mechanism, which is based on a truth-telling equilibrium, collusion is prevented but the authors note that there could be nontruthful equilibrium that involves collusion and which is payoff equivalent to the truth-telling (collusion-free) equilibrium. In their set-up, as in Tirole (1986, 1992) and Laffont and Tirole (1991), collusion is *always harmful* to the principal. Kofman and Lawarrée (1996) show that the principal may prefer taking the risk of allowing collusion (to economize on corresponding deterrence costs) if the internal supervisor is more likely to be honest than dishonest. Kessler (2000) reformulates the Kofman-Lawarrée framework to allow monitoring of effort but suppresses auditing of the agent's pro-

[1] In some models corruption and collusion correspond to the same phenomenon. For instance, an excessively polluting firm may bribe the inspector (Mookherjee and Png, 1995) or a citizen who evades taxes similarly bribes the tax official (Hindriks et al., 1998), on detection of the wrongful activity. In ours they correspond to two distinct interactions at different layers of the hierarchical chain, with differing cost implications: *corruption* occurs when at the bottom layer of the hierarchy a public official knowingly grants licenses to undeserving applicants that inflict direct social costs, whereas *collusion* is agreement to a monetary side-transfer from the corrupt public official to his supervisor for covering up corruption. Collusion involves no direct social costs but affects corruption. A general survey of the literature on corruption is Bardhan (1997).

ductivity and shows, under specific assumptions, that the possibility of collusion between the supervisor and the agent imposes no cost on the principal.

In our framework, potential collusion between a government appointed supervisor and a license issuing public official is again an issue. The supervisor makes a binary effort choice in monitoring and may collude with the official hiding that the official accepted bribes and abused his power in issuing licenses. Our result departs from the literature in favoring supervisor-agent collusion as a means to deter corruption. We show that motivating the supervisor with the lure of a side-transfer from the agent, i.e. a *bounty hunter* mechanism, when feasible, would strictly dominate any direct rewards-based mechanism specifically designed to eliminate collusion. In fact, where collusion prevention is costly the principal prefers a dishonest, collusive supervisor to an honest supervisor who never colludes. Thus, in some situations the principal even actively *encourages* collusion.

Our result favoring the bounty hunter scheme relies on one important assumption – the penalties for bribery and supervision oversight are *nonmonetary* in nature so that they do not directly appear in the principal's objective function. In contrast, Kofman and Lawarrée (1993) and other formulations of the incentive problems in three-layer hierarchies considered mainly monetary penalties that accrued to the principal.[2] While our assumption of nonmonetary penalties imposes a restriction on the principal's instruments in controlling corruption, the assumption captures a range of applications that Kofman and Lawarrée's (1993, 1996) model did not address, thus complementing their work. In corruption cases involving government departments, which is the focus of our analysis, it is somewhat unusual to suggest that the corrupt parties can "buy their way out" for the crime committed by filling in government's coffers with fines. Such an arrangement may even be considered too soft as a policy tool. Monetary penalties imposed in criminal trials to punish corrupt officials usually fall much short of the social cost of corruption. Thus, more often than not, nonpecuniary penalties are imposed exclusively or as a nonsubstitutable part of the punishment. At the same time there is nothing to prevent the corrupt parties to strike side-deals with monetary transfers.

We also characterize the optimal collusion-proof mechanism with notable implications for penalties, providing a contrast with the bounty hunter mechanism. For

[2]Kofman and Lawarrée also consider a case of nonpecuniary punishments, but their principal could replace such punishments with their monetary equivalent so long as the agent and the supervisor had sufficient wealth. See also footnote 9.

instance, a positive penalty may be imposed on the supervisor to dry up any potential surplus from collusion; on the other hand, when collusion prevention is not an issue (either because the optimal mechanism without collusion considerations happens to be collusion-proof, or a bounty hunter mechanism is used), the supervisor's penalty is always set at zero to avoid the deadweight loss associated with the nonmonetary nature of penalty and instead rely mainly on rewards for provision of monitoring incentives. Also, the agent's penalty may be set below the maximum permissible limit to prevent collusion, whereas in the bounty hunter mechanism the penalty may be maximal in order to minimize bribery and corruption.

The paper is organized as follows. The next section presents the model. Section 3 studies the collusion-proof mechanism, which is then compared in section 4 with the bounty hunter mechanism. Section 5 discusses alternative modifications in the basic model. Section 6 concludes. The proofs appear in an Appendix.

2 The model

A government officer, whom we refer to as the *agent*, is delegated the task of awarding a maximum number of n licenses among n applicants, with only one license per applicant. The licenses should be given only to *high*-quality applicants, denoted H; any *low*-quality applicant, of type L, should be turned down. A possible interpretation is that the agent has the authority to grant licenses to *all* high-quality applicants[3] and the maximum number of such applicants is n. We assume that granted licenses cannot be revoked.[4]

Giving a license to a low-quality applicant involves an *irrecoverable* social loss of $\delta > 0$. The social loss could arise, for instance, from the poor quality the applicant later offers the public or inefficiency and distortions in prospective investments in service provision. The probability of a random applicant to be of a particular type

[3]The licenses could be production permits and the authority wants skilled entrepreneurs, who are generally in short supply, to *always* receive such permits. Production organized by skilled entrepreneurs have obvious benefits, not the least of which is higher employment.

[4]A prime reason would be the state's inability to legally justify revoking an awarded license. In section 5 we discuss the case where low-quality applicants' licenses are revoked if detected. We also discuss the case where even some high-quality applicants may need to be turned down due to the limited number of licenses for distribution. Our basic results and their intuitions continue to hold in these cases. However, for ease of exposition, we prefer to work with the present version and assume no licenses can be revoked and no shortage of licenses (at most n licenses for n applicants).

is:

$$pr(L) = q, \;\; pr(H) = 1 - q, \;\; 0 < q < 1.$$

Each applicant knows his own type. We assume that given his expertise the screening agent can determine the true quality of the applicants easily, costlessly, and without failing. This assumption rules out the potential moral hazard problem in the agent's choice of screening effort and so keeps the analysis focused on another moral hazard problem, that of accepting bribes to award licenses to low-quality applicants.

On obtaining a license a low-quality applicant derives a personal benefit of z dollars in excess of his outside option. This benefit is the main source of bribery. On the other hand, a high-quality applicant's surplus from obtaining a license is equal to his competitive surplus elsewhere, normalized to zero. Thus a high-quality applicant will never offer a bribe, which implies bribery is proof that the agent deliberately issued the license to an undeserving, low-quality applicant. We assume that bribes cannot be confiscated.[5]

There are two potential sources of detection of bribery. First, the principal can hire a supervisor (an internal auditor) to check on the agent's potential involvement in bribery. The supervisor either exerts a fixed effort normalized to $e = 1$ that costs him in disutility $\eta > 0$, or shirks so that $e = 0$ and costs zero.[6] The supervisor is unable to detect any bribery if he shirks. With $e = 1$, however, if k licenses are given to low-quality applicants, the supervisor detects bribery with probability $r(k)$. The function $r(k)$ is increasing and strictly convex, with $r(k) \in [0, 1)$ in the relevant range $k \in [0, n]$, and $r(0) = 0$, $r'(0) = 0$. Thus, a larger number of license awards to low-quality applicants makes detection an increasingly likely event. We treat the agent's choice of k as a continuous variable to facilitate derivations and provide intuitions, but the qualitative results will be the same for the alternative (and more accurate) discrete variable interpretation.

[5]The assumption would be reasonable if the agent consumes or diverts the bribe money beyond the authority's reach. Moreover, it is possible that bribery gets uncovered long after the event, and even if it is detected promptly there could be legal/practical difficulties prohibiting its recovery. In any case relaxing the assumption is not going to change the qualitative result much, as we discuss in section 5.

[6]The supervisor's binary monitoring choice could be replaced by a continuous monitoring effort choice, but this would introduce unnecessary complication: the probability of detection becomes a function of both the number of bribes accepted and the level of monitoring effort.

The second potential source of detection is "external" and detects bribery, if there is any, with a fixed probability $0 < \alpha < 1$. The external detection source is assumed incorruptible, i.e., never colludes with the agent or the supervisor, and completely free of charge; we simply refer to it as the *media*.[7] To keep the analysis simple we also assume that α does not depend on the number k of license awards to low-quality applicants.[8]

All parties are risk neutral and their outside option payoffs are normalized to zero. The various parameters of the model such as z, α etc. and the monitoring technology, $r(\cdot)$, are common knowledge.

The incentive scheme includes a flat wage w_S and a reward p_S to the supervisor for reporting bribery; if the supervisor fails to report bribery uncovered by the media, then he is penalized F_S. The agent's wage is denoted w_A. If bribery is uncovered and reported either by the supervisor or the media, the agent is penalized F_A. The penalty is the same whether the agent took bribe from one or more than one applicant. There is an upper bound $\overline{F} > 0$ on the penalties for the crime in question, which we assume is common to F_A and F_S; $\overline{F}$ is determined outside the model by the jurisdiction, the constitution or the executive power with considerations much broader than those dictated by the objective of our principal. We interpret penalties F_A and F_S as nonmonetary sanctions so that they do not explicitly appear as negative items in the principal's cost minimization objective.[9]

[7]The external source may represent an elite (group of) employee(s) of a government agency such as *Independent Commission Against Corruption* in the case of Hong Kong (see Klitgaard, 1988), or *Central Bureau of Investigation* and *Central Vigilance Commission* in the case of India. Alternatively, it can be thought of as journalists/watchdogs who hunt for corruption news.

[8]A justification for this assumption is that the media's intensity of monitoring is set exogenously with the aim of uncovering corruption and various other news in a whole range of public life. The chance of the media catching wrongdoing in a particular government department is fairly low and largely insensitive to the *scale* of a particular event. In section 5 we briefly discuss the impact of allowing α to vary with the number of bribes.

[9]In our context penalties often take the form of blacklisting, sacking or demotion, and even imprisonment. While imprisonment is costly, these costs are usually the responsibility of the state's law and order department and not of a specific government department. The assumption of *only* nonmonetary penalties in our setup is different from *both* pecuniary and nonpecuniary punishments in Kofman–Lawarrée (1993). In their model, so long as the monetary equivalent of nonpecuniary punishment does not exceed the agent's wealth, the principal can reward the supervisor by replacing nonpecuniary punishment with its monetary equivalent. They thus include monetary penalties as positive items in the principal's expected payoff maximization objective. In contrast, we do not consider the penalties as *direct* components of the principal's objective function

The corruption game we analyze is comprised of four stages.

1. *Contract stage.* The principal designs the contracts and offers one for the agent and one for the supervisor. Acceptance leads to the second stage.
2. *Corruption-inspection stage.* The agent meets applicants and executes his bribe solicitation strategy without knowing whether the supervisor actually engaged (or is going to engage) in monitoring (or equivalently, auditing). The supervisor determines his monitoring strategy under incomplete information about realization of the agent's strategy.
3. *Collusion stage.* The outcome of the corruption-inspection stage is realized. If the supervisor detects bribery, the two parties may collude.
4. *Execution of contracts.* The principal receives a report from the supervisor and possibly also from the "media", then executes the contracts.

The inspection-corruption stage admits several interpretations: The supervisor and the agent can be acting simultaneously, or the supervisor may be inspecting the agent ex-post, through the accepted application files. If the supervisor's incentive compatibility constraint is satisfied, the agent knows that he is actually, or will be in the future, inspected with probability one. The agent will determine accordingly the optimal number of bribe solicitation and the timing of inspection is therefore not a crucial issue.

3 Collusion and its prevention

3.1 Honest supervisor

As a reference point, we consider first the case where the supervisor *does not collude* with the agent, i.e., the supervisor is honest or incorruptible just like the external source of detection.

because these penalties never accrue to the principal. Instead, the nonmonetary penalties appear only *indirectly* in the principal's overall cost-minimization objective through possible side-transfers to the supervisor (if collusion is allowed). That is, we assume the agent's wealth to be sufficiently large to (weakly) exceed the maximum permissible nonmonetary penalty.

The principal's objective is to minimize the overall costs of corruption consisting of the social costs and the enforcement costs:

$$TC = w_A + w_S + \sum_{j=1}^{n} n_{c_j} q^j (1-q)^{n-j} [r(b(j)) \cdot p_S + b(j)\delta], \tag{1}$$

where $b(j)$ is the number of bribes accepted by the agent from a total turnout of j low-quality applicants and $r(\cdot)$ is the supervisor's corresponding detection probability.

Consider the agent's bribe solicitation strategy. Suppose the agent has all the bargaining power in determining the size of the bribe for accepting a low-quality applicant. Then any low-quality applicant, when asked for a bribe, simply pays z, his entire monetary benefit from obtaining a license.[10] The agent's expected payoff from taking k bribes, $1 \leq k \leq j$, when the supervisor exerts effort and monitors, is

$$U_A(k) = w_A + kz - [\alpha + (1-\alpha)r(k)]F_A; \tag{2}$$

when no bribe is taken, the payoff is w_A.

Define $k^*(F_A)$ as the solution to the agent's first-order condition,[11]

$$z = (1-\alpha)r'(k)F_A. \tag{3}$$

Where clearly understood, we suppress the argument and denote $k^*(F_A)$ simply as k^*. Next, when the supervisor is employed and given incentives to monitor, define $x(F_A)$ to be the smallest positive real number such that

$$x(F_A) \cdot z \geq [\alpha + (1-\alpha)r(x(F_A))]F_A. \tag{4}$$

For the agent to solicit bribes, the turnout of low-quality applicants must (weakly) exceed $x(F_A)$. If no such $x(F_A)$ exists, the agent takes no bribes. We avoid this uninteresting case by assuming a finite $x(\overline{F}) < n$ exists. When the supervisor is not employed or does not monitor, define $\hat{x}(F_A)$ satisfying $\hat{x}(F_A) \cdot z = \alpha F_A$ and interpret it the same way as $x(F_A)$.

[10] Alternatively, z can be split half-half in the Nash bargaining fashion as in Basu et al. (1992) or Besley and McLaren (1993); then the bribe z can be redefined accordingly. Marjit and Shi (1998) have shown that the structure of bargaining (Nash or take-it-or-leave-it) often determines the *degree* of effectiveness of various corruption control measures. Our results, mostly qualitative in nature, are robust with respect to alternative bargaining schemes.

[11] The second-order condition is satisfied: $-(1-\alpha)r''(\cdot)F_A < 0$.

Lemma 1 **(i)** *Always* $k^*(F_A) \geq x(F_A)$*;*

(ii) $x(F_A)$ *satisfies (4) with equality;*

(iii) $k^*(F_A)$ *is decreasing and* $x(F_A)$ *is increasing in* F_A*, and both are continuous in* F_A*.*

Clearly, $x(F_A) > \hat{x}(F_A)$.

Below we present the optimal bribe solicitation strategy with k^* as a continuous variable (refer section 2), but the optimal strategy can also be stated after converting any non-integer $k^*(F_A)$ to its next higher or lower integer value, whichever yields the agent a higher expected payoff.

[Insert Figure 1 about here]

Bribe solicitation. *Suppose the supervisor is employed and monitors the agent. Assuming the critical* $x(F_A) > 0$ *exists, the agent's optimal bribe solicitation strategy is as follows:*

$$b^*(j, F_A) = \begin{cases} 0, & \text{if } j < x(F_A); \\ \min\{j, k^*(F_A)\}, & \text{if } j \geq x(F_A). \end{cases}$$

(See Figure 1). Also, the strategy, $b^*(j, F_A)$*, is (weakly) decreasing and continuous in* F_A*.*

When the supervisor is not employed or does not monitor, the agent would solicit the maximum number of bribes, if at all, and award a license to all low-quality applicants:

$$\hat{b}(j, F_A) = \begin{cases} 0, & \text{if } j < \hat{x}(F_A); \\ j, & \text{if } j \geq \hat{x}(F_A). \end{cases}$$

Optimality of $\hat{b}(j, F_A)$ is obvious. Optimality of $b^*(\cdot, \cdot)$ follows from strict concavity of the agent's expected payoff in k. That $b^*(\cdot, \cdot)$ is (weakly) decreasing in F_A is straightforward, given Lemma 1, and continuity of $b^*(\cdot, F_A)$ is implied by continuity of $x(F_A)$ and $k^*(F_A)$ and the $\min\{., .\}$ function. Henceforth we suppress the argument F_A and denote $b^*(j, F_A)$ simply as $b^*(j)$, where there is no confusion.

To avoid a different type of confusion, it is worth remarking that the agent's bribe solicitation strategy formulated above for the honest supervisor case will be no different when the supervisor is dishonest, that is, open to collusion with the agent. Later on we will denote the agent's strategy using alternative notations such as $b^{CP}(j)$ and $b^{BH}(j)$, depending on whether a collusion-proof or a bounty hunter arrangement is being analyzed.

Given $b^*(j)$, the ex-ante[12] expected payoff and participation constraint of the agent is

$$U_A = w_A + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} [b^*(j) \cdot z - (\alpha + (1-\alpha) r(b^*(j))) \cdot F_A] \geq 0. \tag{2'}$$

Since the agent can guarantee himself w_A by remaining honest and the agent's bribe solicitation does not depend on w_A, *the principal sets* $w_A^* = 0$.

Consider now the supervisor's problem. The supervisor will exert the monitoring effort and incur the cost η if and only if both the incentive compatibility constraint (in short, *ICC*) given in (5) and the participation constraint (PC_S) in (6) hold:

$$w_S + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} [r(b^*(j)) p_S - (1 - r(b^*(j))) \alpha F_S] - \eta$$
$$\geq w_S - \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} \alpha F_S$$

$$\text{i.e.,} \qquad \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} r(b^*(j)) (p_S + \alpha F_S) \geq \eta; \tag{5}$$

$$w_S + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} [r(b^*(j)) p_S - (1 - r(b^*(j))) \alpha F_S] - \eta \geq 0. \tag{6}$$

We start with the following observations regarding the supervisor's optimal incentive scheme:

Lemma 2 *To minimize the overall costs, the principal must choose p_S and F_S to bind the effort incentive constraint (5). Specifically, the penalty F_S should be set at zero so that the reward p_S satisfying (5) with equality is minimized, thereby minimizing (1).*

Also, the supervisor's optimal wage is $w_S^* = 0$.

Let us explain the intuition for why the optimal F_S is zero. From the participation constraint (6) it follows that the enforcement costs (supervisor's wage plus expected rewards) in total expected costs must be at least η, strictly exceeding η

[12] That is, before the agent knows j, the actual number of low-quality applicants.

if $F_S > 0$. Note that setting $F_S = 0$ and choosing p_S to bind the incentive compatibility constraint also satisfies the participation constraint with equality if $w_S = 0$. This way, enforcement costs are minimized, which is optimal for the principal. In short, in choosing among the instruments to influence the supervisor's monitoring incentives, *the carrot (i.e., p_S) is better than the stick* because the deadweight loss of nonmonetary penalty can be justified only for its deterrence role; however, the supervisor gets penalized even when he exerts the effort but fails to uncover corruption.

The agent's penalty, F_A, despite its nonmonetary nature and the associated deadweight loss, must be set at a positive level unlike the penalty on the supervisor because it is the sole instrument available to control the agent. But the choice of F_A involves a potential tradeoff: Setting a large F_A decreases the number of bribes but requires a larger reward promise to the supervisor, suggesting an ambiguous impact on the principal's cost objective. However, because rewards can always be adjusted to keep the principal's expected (reward) cost of using the supervisor equal to η, the first effect always dominates, thus a larger F_A is always beneficial as it lowers corruption and the associated social costs:[13]

Lemma 3 *Given $F_S^* = w_S^* = w_A^* = 0$, decreasing the agent's penalty will increase the overall costs for the principal, hence $F_A^* = \overline{F}$.*

Finally, the agent's participation constraint ($2'$) need not bind: with the optimal incentives already determined there is no other instrument left for the principal to run down the agent's rent.

We summarize the results obtained so far as follows:

Proposition 1 *If the supervisor is employed, in the absence of collusion the principal minimizes overall expected costs of corruption by setting the maximal penalty $F_A^* = \overline{F}$ on the agent and inducing the most conservative bribe solicitation strategy $b^*(j, \overline{F})$.*

The supervisor's penalty is set at $F_S^* = 0$ and reward p_S^* is just large enough to induce effort, satisfying (5) with equality.

[13]Note that the conclusion that F_A must be set at the maximal level relies on the premise that increasing F_A would strictly lower $b^*(j, F_A)$ by continuously lowering k^*. However, due to the discrete nature of bribe solicitation possibilities it is possible that F_A is not set maximally; instead, F_A is increased only to the point beyond which any further impact on the agent's bribe solicitation strategy would vanish.

Both wages are set to bind the limited liability constraints for the supervisor and the agent: $w_A^* = 0$ *and* $w_S^* = 0$.

The alternative option for the principal is to dispense with the supervisor and rely exclusively on the external source of detection. As in the case where the supervisor is employed, the principal sets $\hat{w}_A = 0$ and F_A maximal at $\overline{F}$ inducing maximal $\hat{x}(F_A)$ and minimizing expected social costs of corruption:

$$TC_{\hat{b}} = \sum_{j \geq \hat{x}(\overline{F})}^{n} n_{c_j} q^j (1-q)^{n-j}\, \hat{b}(j, \overline{F})\delta.$$

Using the incentive scheme characterized in Proposition 1, the principal's total expected costs will be lower when the supervisor is employed if and only if

$$[\sum_{j \geq \hat{x}(\overline{F})} n_{c_j} q^j (1-q)^{n-j} \hat{b}(j, F_A) - \sum_{j \geq x(\overline{F})} n_{c_j} q^j (1-q)^{n-j} b^*(j, \overline{F})]\delta \geq \eta,$$

which will hold if the number of bribes induced by employing the supervisor is sufficiently low to compensate for the monitoring cost η incurred by the principal. We assume that this condition holds, so that the supervisor is hired, in the rest of the paper.

3.2 Dishonest supervisor

We now consider the supervisor to be dishonest, who will collude with the agent if there is a surplus from doing so. To dissuade the supervisor from accepting side-transfers from the agent, the supervisor's rewards should satisfy the *collusion-proofness constraint* (in short, CPC):

$$p_S \geq (1-\alpha)F_A - \alpha F_S. \tag{7}$$

The agent cannot guarantee escaping punishment by making a side-transfer to the supervisor because the media can uncover bribery. Also, the fact that the supervisor is penalized for not reporting bribery that is uncovered by the media puts a check on the supervisor in accepting bribes, lowering the required rewards for the supervisor.

The principal now minimizes (1) subject to (7), the effort incentive constraint (5) and the participation constraint (6), with the agent choosing his bribe solicitation strategy optimally. The following proposition describes certain features of the optimal collusion-proof incentive scheme for the principal.

Proposition 2 *Suppose that the (*)-mechanism described in Proposition 1 violates the constraint (7). Under the optimal collusion-proof (CP) mechanism, the principal's costs are larger. The optimal mechanism has the following features:*

[*i*] $w_A^{CP} = w_S^{CP} = 0.$

[*ii*] *The collusion-proofness constraint (7) is always binding.*

[*iii*] *The supervisor's participation constraint (6) must also bind unless $F_S^{CP} = \overline{F}$. To bind (6) the principal should substitute the rewards p_S with an increase in the penalties F_S, whenever possible (i.e., without violating (7)).*

[*iv*] *Either $F_A^{CP} < \overline{F}$ or $F_S^{CP} > 0$ must hold (or possibly both hold).*

[*v*] *Both the supervisor and the agent may earn some positive rent.*

Although not a complete characterization, Proposition 2 describes a number of important features that derive from potential collusion. First, the collusion-proofness constraint must always bind. This is intuitive, given that collusion is the main reason why the principal must depart from the (*)-mechanism derived under the honest supervisor assumption. Second, the agent's maximal penalty property of Proposition 1 may no longer hold; alternative to lowering F_A (which increases bribery, thus, social costs), the principal must either increase p_S or increase F_S, both of which are also costly directly or indirectly. Third, despite the associated deadweight loss, the principal may set the penalty F_S at a positive level in contrast to the honest supervisor model of Proposition 1, where he could rely more on direct rewards to the supervisor for provision of monitoring incentives. Fourth, the principal may have to leave a positive surplus to the supervisor in order to eliminate the possibility of collusion. These last three factors combine to make collusion prevention costly.

What is not possible, however, is to identify a clear pecking order in the choice between a positive F_S and a non-maximal F_A (refer part [*iv*] above). The two adjustments have different cost implications – positive F_S involves deadweight loss and non-maximal F_A involves both higher enforcement costs and higher social costs of corruption. Depending on the rate of increases in overall costs resulting from each type of adjustment, one or both instruments could be relied upon by the principal. Finally, the principal's preference for any adjustment by increasing F_S

and lowering of p_S, whenever possible (i.e., without violating any of the constraints), is understandable given that it economizes on direct reward payments and eases the collusion-proofness constraint.

4 The bounty hunter mechanism

In this section we evaluate the performance of the *bounty hunter* (in short, BH) mechanism, whereby the principal allows collusion and replaces direct rewards p_S by the potential side-transfer the supervisor may obtain from the agent upon detection of bribery.

Throughout the analysis we assume that the supervisor can destroy any credible evidence of bribery to reach a side-transfer agreement with the agent. Therefore, unless bribery is uncovered by the media, the colluding parties guarantee no punishment.[14] We also assume that the supervisor has all the bargaining power in determining the side-transfer, which is therefore given by $(1-\alpha)F_A$, the maximum that the agent is willing to pay. For a positive surplus from collusion, this amount should exceed $p_S + \alpha F_S$, the minimum the supervisor must be paid for destroying the evidence.[15]

The BH-mechanism economizes on the monitoring cost η. However, it may generate a potential cost due to the loss of the instrument, p_S. The supervisor's "reward", now determined by the right-hand side of (7) as $(1-\alpha)F_A - \alpha F_S$, is bounded above by maximal penalties, thus may not be large enough to satisfy the supervisor's incentive and participation constraints.

We begin the analysis of the BH-mechanism with some basic observations about incentives in the hierarchy. Note that the agent's expected payoff and participation constraint under collusion is exactly as given in ($2'$): If the agent is bribed and

[14]Collusive agreements can be enforced through internal mechanisms involving credible threats of retaliations; it can be endogenized in a multi-period model, but this is beyond the scope of this paper.

[15]The assumption that the supervisor has all the bargaining power can be replaced by a general bargaining process where the agent's transfer is an increasing function of the supervisor's disagreement utility, as in the Nash bargaining solution. Under the BH-mechanism the principal can influence the supervisor's disagreement utility through the unpaid, official, reward p_S^{BH}. If p_S^{BH} is chosen arbitrarily close to but less than $(1-\alpha)F_A^{BH}$ (and choosing optimally $F_S^{BH} = 0$; see Lemma 4), the supervisor will have almost all the bargaining power and the transfer he receives from the agent will be $(1-\alpha)F_A^{BH}$.

detected by the supervisor, under collusion he pays $(1-\alpha)F_A$ to the supervisor for not reporting, but risks paying an additional F_A to the principal with probability α if detected by the media; without collusion he pays the entire penalty F_A to the principal, which yields the same expected payoff expression. Then, given F_A, the agent's bribe solicitation strategy is also unchanged, as stated in the honest supervisor case following Lemma 1: $b^{BH}(j, F_A) = b^*(j, F_A)$. Under the BH-mechanism the supervisor's incentive and participation constraints are respectively (8) and (9), the analogues of (5) and (6) where p_S is replaced by $(1-\alpha)F_A - \alpha F_S$:

$$\sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} \cdot r(b^*(j, F_A))(1-\alpha)F_A \geq \eta; \tag{8}$$

$$w_S + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} \cdot [r(b^*(j, F_A))(1-\alpha)F_A - \alpha F_S] - \eta \geq 0. \tag{9}$$

Our analysis of the BH-mechanism to follow will be for two different settings: in the first, the (*)-mechanism is not collusion-proof so that the alternative to the BH-mechanism is the CP-mechanism; in the second, the (*)-mechanism *is* collusion-proof, thus, is also the alternative to the BH-mechanism. The following result characterizes optimal BH-mechanism common to both settings.

Lemma 4 *Under any optimal bounty hunter mechanism, the principal always sets* $F_S^{BH} = 0$ *along with* $w_S^{BH} = w_A^{BH} = 0$.

Intuitively, why penalize the supervisor for not reporting bribery uncovered by the media if, after all, the incentives are especially designed to induce the supervisor to collude with the corrupt agent? In more detail, the penalty F_S has no effect on the supervisor's incentive compatibility constraint: A positive penalty decreases the supervisor's expected payoffs from not monitoring and monitoring the agent by the same amount, because he does not report bribery in either case. But a positive value for F_S^{BH} has a negative impact on the supervisor's overall expected payoff which, to keep the participation constraint satisfied, must be compensated for by an increase in the base wage w_S and so brings in an additional cost for the principal. Hence, $F_S^{BH} = 0$.

4.1 When the (*)-mechanism is *not* collusion-proof

When the (*)-mechanism fails to be collusion-proof, collusion prevention is costly as shown in Proposition 2. Then the principal's options are either to *prevent* collusion through the CP-mechanism, or to *allow* collusion through the BH-mechanism. The following proposition compares these two options.

Proposition 3 *Suppose the (*)-mechanism is not collusion-proof, that is,* $p_S^* < (1-\alpha)\overline{F}$. *Then, in the optimal bounty hunter mechanism the principal sets the agent's penalty maximal at* $F_A^{BH} = \overline{F}$ *and induces the bribe strategy* $b^{BH}(j, \overline{F}) \leq b^{CP}(j, F_A^{CP})$. *As a result, total expected costs of corruption under the optimal bounty hunter mechanism will be strictly less than under the optimal CP-mechanism.*

We emphasize the following implication of Proposition 3:

When penalties are mainly nonmonetary in nature and collusion prevention is costly, the principal would strictly benefit to control corruption by allowing collusion.

There is a simple intuition to why the BH-mechanism should be preferred. In the collusion-proof scheme the (nonmonetary) penalties that are enforced would have been lost from the system, whereas under the bounty hunter arrangement those penalties effectively finance the monitoring costs. Thus, overall the society is better off by at least the monitoring cost η. At least, because, in addition the BH-mechanism can induce a smaller expected number of bribes with $b^{BH}(j, \overline{F}) < b^{CP}(j, F_A^{CP})$, if $F_A^{CP} < \overline{F}$.

Perhaps more striking is that, when the (*)-mechanism fails to be collusion-proof, the principal would not only allow collusion, he should even actively *encourage* collusion! That is, the principal should hire a dishonest supervisor rather than an honest supervisor. Hiring an honest supervisor implies that the total cost will be according to the (*)-mechanism. Hiring a dishonest supervisor and implementing the BH-mechanism actually reduces the total cost because the principal induces the same bribe solicitation strategy, $b^{BH}(j, \overline{F}) = b^*(j, \overline{F})$, without having to pay any reward to the supervisor.[16]

[16]This logic can be extended to the case where the supervisor's type (honest or dishonest) is private knowledge. Though we do not solve for the optimal mechanism under uncertainty about the supervisor's honesty, based on the results above we can conclude that a dishonest supervisor would be preferred to a supervisor who is honest with positive (but less than one) probability. In the latter case, to induce an equilibrium where the agent is monitored with probability one, the principal has to rely on CP-mechanism because the alternative of BH-mechanism may not work if the supervisor turns out to be honest, while the (*)-mechanism will not work if the supervisor is dishonest. Total costs will then be smaller under a dishonest supervisor operating under the BH-mechanism. This is in line with a result in Kofman and Lawarrée (1996) that we discuss below.

We are not aware of any obvious theoretical justification for why the principal may want to actively encourage 'cover-up' of corruption (i.e., collusion). The general emphasis has been (Tirole, 1986, 1992; Laffont and Tirole, 1991) that collusion harms the principal. In a different context Olsen and Torsvik (1998) show that collusion between the supervisor and the agent may benefit the principal by alleviating an intertemporal contractual commitment problem. The benefits of the BH-mechanism are reminiscent of a finding by Itoh (1993) that the principal may prefer side-contracting amongst agents. Kessler (2000) provides a setup in which collusion imposes no cost on the principal but her model differs substantially from the present one. Moreover in her model inducing collusion never dominates the collusion-free mechanism. Kofman and Lawarrée (1996) show that collusion may be allowed under incomplete information about the supervisor's type, if ex-ante the supervisor is likely to be honest, to economize on the cost of deterring collusion. In their setup collusion is always harmful and the principal prefers the honest, non-collusive supervisor type.

In our context, the principal's preference for a dishonest supervisor partly relies on the assumption that the (*)-mechanism is not collusion-proof. When the (*)-mechanism *is* collusion-proof, the case we analyze next, the BH-mechanism would induce an increase in bribery above $b^*(j, \overline{F})$. However the savings on enforcement costs, η, could still be large relative to the increased social costs of bribery to justify the use of a dishonest supervisor. Thus, the main intuition in favor of collusion (or encouragement of collusion) remains valid.

4.2 When the (*)-mechanism *is* collusion-proof

We now focus on the case $p_S^* > (1-\alpha)\overline{F}$ so that *collusion prevention is no longer costly.* In such situations the BH-mechanism potentially runs into a difficulty because the agent's penalty F_A performs two functions at the same time. It determines the agent's bribe solicitation strategy *and* the supervisor's reward to motivate monitoring. The principal may not be able to hit both targets with the same one instrument, that is, BH-mechanism may no longer be feasible.

BH-mechanism is feasible if there exists a penalty $F_A \in [0, \overline{F})$ such that, given the agent's bribe solicitation strategy $b^*(j, F_A)$,

$$\sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} r(b^*(j, F_A))(1-\alpha) F_A \geq \eta. \tag{10}$$

Condition (10) is same as (8), the supervisor's incentive compatibility constraint[17] which, recall, is satisfied with equality under the (*)-mechanism where $F_A = \overline{F}$ and p_S^* replaces $(1-\alpha)\overline{F}$.

Clearly, in the case $p_S^* > (1-\alpha)\overline{F}$ the principal cannot induce the same outcome through the BH-mechanism by setting the agent's penalty maximal at $F_A = \overline{F}$ for this will violate (10): Given the agent's bribe solicitation strategy $b^*(j, \overline{F})$, the maximum "reward" $(1-\alpha)\overline{F}$ the supervisor can get from the agent by monitoring and detecting bribery is not large enough to compensate the supervisor for the monitoring cost η. Then the principal has to modify the agent's penalty within the feasible range $[0, \overline{F})$ and see whether (10) can be satisfied to induce monitoring effort given the agent's optimal bribe solicitation strategy $b^*(j, F_A)$.

Lowering the penalty F_A below $\overline{F}$ will have opposing effects on the supervisor's monitoring incentives: the L.H.S. of (10) tends to increase due to the increase in $b^*(j, F_A)$, whereas the opposite happens due to the decrease in indirect rewards component $(1-\alpha)F_A$. Thus, even when the agent's penalty is lowered, the BH-mechanism may not be feasible. The next question is whether, when feasible, the optimal BH-mechanism would generate lower costs than the (*)-mechanism. If BH-mechanism were feasible for penalties *close enough to* $\overline{F}$, the principal would prefer using it to economize on expected reward payments (which amount to η) and accept a small increase in the incidence of bribery.[18] But as F_A is lowered sufficiently, the agent will switch to soliciting bribes from all low-quality applicants, setting $b(j, F_A) = j$. The principal's total costs must then exceed the costs under the (*)-mechanism.[19]

When BH-mechanism is feasible, let F^+ denote the maximal penalty $F_A < \overline{F}$ satisfying (10) given $b^*(j, F_A)$. We have the following proposition.

Proposition 4 *Suppose* $p_S^* > (1-\alpha)\overline{F}$ *so that the principal can implement the bribe solicitation strategy* $b^*(j)$ *collusion-proof by imposing the maximal penalty on the agent through the (*)-mechanism. If the bounty hunter mechanism is feasible and used, then* $w_A^{BH} = w_S^{BH} = F_S^{BH} = 0$ *and the optimal penalty for the agent is*

[17] It also represents the supervisor's participation constraint because, by Lemma 4, $w_S^{BH} = 0$.

[18] Lowering F_A slightly below $\overline{F}$ causes a discrete reduction in enforcement costs by η, whereas bribery may slightly increase, or even remain the same due to the discrete nature of bribe solicitation possibilities.

[19] Recall, we assumed that the costs under the (*)-mechanism are lower than the case in which the supervisor is not used (where the external source, the media, is the sole source of detection).

set at F^+, inducing the bribe solicitation strategy $b^{BH}(j, F^+) = b^*(j, F^+) > b^*(j)$. Total costs under the BH-mechanism are smaller than under the (*)-mechanism if

$$\left[\sum_{j \geq x(F^+)} n_{c_j} q^j (1-q)^{n-j} b^{BH}(j, F^+) - \sum_{j \geq x(\overline{F})} n_{c_j} q^j (1-q)^{n-j} b^*(j, \overline{F})\right] \delta < \eta.$$

The optimal penalty under the BH-mechanism is set as large as possible, thus equals F^+, because among all F_A that satisfy (10) and so induce the supervisor to monitor, the penalty F^+ minimizes the agent's bribe solicitation $b(j, F_A)$ and thereby minimizes also the principal's cost objective. An interesting feature of the BH-mechanism is that the agent's penalty is less than $\overline{F}$, the "official" upper bound on F_A.[20]

To clarify the role of nonmonetary penalties for the dominance of the bounty hunter arrangement, let us alternatively consider monetary penalties. A short argument suffices for our purpose. Fix any configuration of bounty hunter incentives. In any state where the supervisor detects and hides bribery, he receives a transfer of $(1-\alpha)F_A^{BH}$ from the agent and consequently the principal would lose penalties of expected value $(1-\alpha)F_A^{BH}$. However, if the principal chooses the reward $p_S = (1-\alpha)F_A^{BH}$, in any state of detection of the bribery the supervisor would no longer hide it and as a result the principal's (expected) penalty collection from the agent increases by $(1-\alpha)F_A^{BH}$ with which the reward p_S is financed. Thus the principal's gains and losses balance out, while no other constraints are affected. This shows that under monetary penalties there is an alternative, collusion-free mechanism that does at least as well as the BH-mechanism, hence the bounty hunter arrangement cannot dominate collusion-proof incentives. Furthermore, because the principal is free to adjust p_S, the feasible set of incentives under alternative collusion-free mechanisms is strictly larger than under the BH-mechanism, which may bring down the overall costs for the principal. Thus, the principal would never gain by choosing the BH-mechanism under monetary penalties. This reasoning continues to hold for the issue of hiring a dishonest supervisor or an honest supervisor: Under monetary penalties the principal will opt for an honest supervisor.

[20]While the literature on crime deterrence provides several reasons for why penalties may not be set maximal, none of these coincide with the explanation we provide in this paper, which, as mentioned, stems from the principal's motive to generate collusion between the supervisor and the agent at minimum cost.

Thus, changing the nature of penalties from monetary to nonmonetary can drastically modify the principal's choice of control mechanism. Reality is somewhere between the two extremes. Favoring monetary penalties will no doubt strengthen the case against the use of collusion in controlling corruption. If corrupt agents can conceal or find a way of protecting large fractions of their wealth from public authorities, penalties are de facto mostly nonmonetary. In such environments our results are in favor of the bounty hunter mechanism.

5 Results under alternative assumptions

Most of our assumptions are mainly to keep the analysis simple and focused with the basic intuitions robust to plausible modifications. Below we briefly discuss the importance of various assumptions and consider their modifications to check whether and how our results may change.

We assumed that the number of applicants, say m, is equal to n, the maximum number of licenses to be awarded. If $m > n$, a door opens for the agent to extract bribes from even the high-quality applicants. However, under our assumption that high-quality applicants' surplus from the specific license is equal to their competitive surplus elsewhere, the agent cannot hope to extract any bribe from these applicants. Suppose, then, that high-quality applicants derive a positive surplus specific to the license, say π_H, strictly less than π_L, a corresponding surplus for the low-quality applicants.[21] Now the government can impose a license fee $\kappa = \pi_H$, so that the agent is no longer able to extract bribes from the high-quality applicants. Then the net surplus to any low-quality applicant would be $z = \pi_L - \kappa > 0$, which is available for bribery. With this modification, the principal would minimize the following modified cost objective:

$$\begin{aligned}\widetilde{TC} = w_A + w_S \quad &+ \quad \sum_{j=1}^{m} m_{c_j} q^j (1-q)^{m-j} [r(b(j)) \cdot p_S + b(j)\delta \\ &+ \quad \min\{b(j), m - j + b(j) - n\} \cdot 1_{\{m-j+b(j)>n\}} \cdot \beta_H \\ &- \quad \min\{n, m - j + b(j)\} \cdot \kappa], \end{aligned} \tag{11}$$

[21]While a low-quality applicant derives a greater personal benefit from the license, the overall social surplus is larger if the license is awarded to a high-quality applicant. This is in contrast with Banerjee (1997), for example, where both personal and social benefits are higher for the high-quality applicants.

where β_H is the additional *social benefit* (over and above the applicant's personal benefit) from awarding a license to a high-quality applicant, and $1_{\{.\}}$ is the indicator function. In (11), the expression in the second line measures the social benefits that are lost (thus, stacking up extra social costs over and above the δ-costs) as some high-quality applicants miss out obtaining licenses when the agent awards licenses to some low-quality applicants instead. The expression in the third line of (11) represents expected license fees. These two expressions were absent in our original formulation of the principal's cost-minimization problem in (1) because there a high-quality applicant always received a license, for free.

The solutions to the principal's problem will be qualitatively affected only if the above modification substantively alters the principal's approach to containment of bribery through the penalties (both on the agent and the supervisor) and/or the supervisor rewards; the wages perform no essential role. Note that the sole impact of the modification in the cost objective is to increase the social cost of bribery.[22] The decisions of bribe solicitation, monitoring and collusion are not affected in any way, hence the incentive compatibility, collusion-proofness and participation constraints are unchanged. In the *honest supervisor model of section 3.1*, the principal sets F_A maximal and binds the supervisor's incentive and participation constraints; modifying these incentives would clearly increase both social costs (due to δ) and lost benefits (due to β_H) that are likely to exceed any increase in revenues from the granting of licenses to low-quality types. In the *dishonest supervisor case of section 3.2*, our findings that the agent's penalty may be less than the maximal and the supervisor's penalty possibly positive are derived based on the fact that always the CPC and often the PC_S are binding. These last two results involving the constraints (see parts [i] and [ii] of Proposition 2) are obtained by the method of contradictions without altering F_A and the induced bribery $b(j)$.[23] But because $b(j)$ is the only variable (under principal's indirect control) appearing in the two additional terms of the modified objective function (11), our method of proof by contradictions in Proposition 2 remains valid. Finally, the argument in our main result in Proposition 3 is very general and works equally well with the modification

[22] While the license fees do bring down the costs, granting licenses to low-quality applicants should never be the principal's objective.

[23] The only instance where F_A is altered is to prove that it must be maximal to bind the CPC. When the CPC is not binding and the penalty is non-maximal, the (logic of) improvement in the cost objective by increasing the penalty, F_A, remains valid for the modified cost objective.

in the objective function. The main qualitative result in Proposition 4 is that, even when collusion prevention is not costly the BH-mechanism may be of value to the principal, especially if the supervision cost η is large. The intuition behind this result should therefore continue to be valid under the modified objective function.

We assumed no confiscation of bribes. Suppose now that all bribes can be costlessly confiscated by the state (principal) – a polar opposite assumption. The monitored agent will reduce bribe solicitation because the potential loss of corrupt proceeds affects his incentives in the same qualitative way as an increase in the penalty F_A. Under the (*)-mechanism the principal modifies p_S to keep the supervisor's expected reward payments equal to the monitoring cost η. All other components of the mechanism are unchanged. Thus, if bribes can all be confiscated upon detection, under (*)-mechanism total costs will fall for two reasons, first because corruption is lower, and second, because the principal can finance part of reward payments to the supervisor through confiscated bribes. The possibility of collusion now depends on the number of bribes and on the supervisor's information about this number. Let us assume that the supervisor learns the number of bribes or their size if monitoring is successful, and that if the parties collude and corruption is detected by the media, the bribes can still be confiscated, this time from the supervisor. To achieve collusion-proofness, the principal has to consider the case where the agent pockets the sum $k^*(F_A)z$ and is not constrained by the turnout of low-quality applications. Then the collusion-proofness constraint will change, the R.H.S. increasing by $(1-\alpha)k^*(F_A)z$. Modifying the incentive scheme to satisfy collusion-proofness will increase costs under the CP-mechanism in accordance with our claim in Proposition 2. Consider now the BH-mechanism, where the bribed and detected agent is willing to transfer the sum $(1-\alpha)(F_A + min\{j, k^*(F_A)\}z)$ to the supervisor to avoid being reported. Clearly the agent's payoff and bribe solicitation strategy are exactly the same as under the new (*)-mechanism (i.e., the mechanism relevant for the confiscated-bribe version) because the agent gets the same payoff in every outcome of the monitoring game. So if the new (*)-mechanism is not collusion-proof and the BH-mechanism is used instead of the CP-mechanism, the principal economizes *at least* the monitoring cost η but will loose the expected bribe proceeds to the supervisor.[24] We conclude that the ranking of mechanisms

[24]But if in addition even the bribe-transfer to the supervisor can be confiscated (recall, bribe-transfer is only a part of the side-transfer), then the noted loss for the principal will not occur.

stated in Proposition 3 and 4 continues to hold unless expected bribe proceeds are very large.

We assumed for simplicity that licenses cannot be revoked once awarded. The possibility of revoking low-quality applicants' licenses reduces the bribe that these applicants are willing to offer, thus decreases the agent's private benefit from corruption and reduces his bribe solicitation given the penalty. If these licenses can be revoked before the social harm is realized, the social cost of corruption will be eliminated whenever corruption is detected. The advantage of BH-mechanism relative to the CP-mechanism is the same: to economize on monitoring costs η and incentive costs as mentioned in Proposition 2. Its disadvantage is that expected corruption and the associated costs are likely to be larger. The disadvantage will be small the more promptly the licenses are revoked, that is, *before* the social harm is inflicted.

Our assumption regarding the timing of the agent's bribe solicitation strategy and the internal supervisor's monitoring decision can be questioned. As is typically the case in inspection games, we assumed that the two parties act under incomplete information about each other's actions. They could thus be acting simultaneously, or sequentially, the agent first, followed by the supervisor under incomplete information about the agent's bribe proceeds. One could also introduce an additional ex-post stage of inspection, to be activated if the internal supervisor observes/reports some bribery. Such an extension would be worthwhile if the penalty is made contingent on the proportion of inspected applications awarded to undeserving applicants, in the spirit of "penalty fitting the crime." We do not pursue this line of inquiry for two reasons: *practicality* and *simplicity.* Rarely is it the case that the media or an internal audit inspects each and every application (or even a large proportion of applications) to determine the fraction of inappropriately awarded licenses. Such large scale inspection could be costly to administer. Also, contingent penalty schemes will complicate the analysis and we do not believe our main results will be affected qualitatively.

We assumed the probability of detection α by the external source, media, to be constant. Allowing α to be (weakly) increasing in the number of bribe solicitations, k, is perhaps more realistic but the analysis becomes much more involved. Instead of making conjectures about the detailed implications for our analysis, we would indicate how the strategies of the two main players – the agent and the supervisor

– are affected and what it might mean for the principal. If one assumes $\alpha(k)$ to be concave in k (similar to the strict concavity of $r(k)$) and $\alpha'(n)$ sufficiently small to satisfy the second-order condition for the agent's problem, then the agent's optimal $k^*(F_A)$ will be unique and decreasing in F_A. Also it is easy to check (as in Lemma 1 proof) that the bribe trigger $x(F_A)$, solving $x \cdot z = [\alpha(x) + (1 - \alpha(x))r(x))]F_A$, will be increasing in F_A. These two facts together imply that the agent's optimal bribe solicitation strategy, $b^*(j, F_A)$, will be qualitatively the same as before. As for the supervisor's incentives, the reward p_S and the penalty F_S basically work in the same way as before. The principal may still not want to impose a positive penalty on the supervisor when collusion is not an issue, because the penalty is still a deadweight loss that can be saved. However, how the principal should design either the collusion-proof or bounty hunter incentives are relatively difficult issues. In both these cases the principal's approach to the supervisor's incentives requires more detailed considerations, not knowing the exact number of low-quality applicants, j, and hence the extent of actual bribery $b^*(j, F_A)$. Does our main intuition favoring the principal's use of a collusion-inducing program change? We like to think not, though only further research can satisfactorily resolve this issue.

6 Conclusion

A well-known feature of cost-effective control of corruption since Becker (1968) and Becker and Stigler (1974) is optimal management of the trade-off between the level of corruption and the resources spent on enforcement of anti-corruption legislation. While keeping this feature present in our model, we focused on the choice between preventing and allowing, or even encouraging, collusion within public organization hierarchies. This choice depends on the nature of penalties, that is, the extent to which penalties on detected corrupt officials take the form of nonmonetary sanctions. Though in some public organizations collusion can generate the benefit of avoiding the problem of double-marginalization as noted by Shleifer and Vishny (1994), its benefits in our setup stem from economizing on rewarding of enforcement efforts as well as avoiding the costs of preventing collusion under nonmonetary sanctions. Just how significant are nonmonetary sanctions as a fraction of the overall penalty on corrupt officials is an empirical question and the answer would no doubt vary across jurisdictions.

Appendix

Proof of Lemma 1. Part (i) follows from the definitions of $k^*(F_A)$ and $x(F_A)$. Part (ii) follows from continuity of the expressions on both sides of (4) with respect to $x(F_A)$.

That $k^*(F_A)$ is decreasing and continuous in F_A are straightforward. To show that $x(F_A)$ is increasing in F_A, suppose not. Let F_A be increased from f to f'. Clearly, $x(f') = x(f)$ is impossible because the R.H.S. of (4) will be increased while the L.H.S. will remain unchanged, contradicting part (ii). Suppose, then, $x(f') < x(f)$. By definition of $x(f')$ we have

$$x(f') \cdot z = [\alpha + (1-\alpha)r(x(f'))]f',$$

which implies

$$x(f') \cdot z > [\alpha + (1-\alpha)r(x(f'))]f.$$

Then we can decrease $x(f')$ slightly to some $\tilde{x}$ so that

$$\tilde{x} \cdot z > [\alpha + (1-\alpha)r(\tilde{x})]f,$$

which contradicts the definition of $x(f)$. Hence, $x(f') \geq x(f)$. Combining with the fact that $x(f') = x(f)$ is impossible establishes our claim, $x(f') > x(f)$.

Continuity of $x(F_A)$ is straightforward. **Q.E.D.**

Proof of Lemma 2. Suppose the incentive compatibility constraint (5) is not binding for the cost-minimizing incentives chosen by the principal; for this lemma we leave F_A unspecified as the argument holds for any F_A chosen by the principal. If both $p_S > 0$ and $F_S > 0$, the principal is able to lower p_S and F_S slightly without violating (5) or (6), which will lower costs[25] – a contradiction. If $p_S = 0$ and $F_S > 0$ so that $w_S > 0$ to satisfy the participation constraint, the principal can lower F_S and w_S slightly to lower the overall costs, again a contradiction. If $p_S > 0$ and $F_S = 0$, the principal can lower p_S slightly and satisfy both the incentive compatibility constraint and the participation constraint; this will lower the overall costs, again a contradiction.

[25]It is sufficient to choose $\Delta p_S < 0$ and $\Delta F_S < 0$, both very small in magnitudes, such that $r(b^*(j))\Delta p_S - (1 - r(b^*(j)))\alpha\Delta F_S \geq 0 \ \forall j \geq x(F_A)$. This will be achieved if $\frac{\Delta p_S}{\Delta F_S} \geq \frac{(1-r(b^*(n)))}{r(b^*(n))}$.

Suppose now $F_S > 0$. Given that the ICC is binding, rewrite the participation constraint:

$$w_S + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} [r(b^*(j))(p_S + \alpha F_S) - \alpha F_S] - \eta \geq 0,$$

$$\text{i.e.,} \qquad w_S - \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} \alpha F_S \geq 0.$$

To minimize total costs the principal must set $w_S = \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} \alpha F_S$, and total costs are:

$$\begin{aligned} TC_0 &= w_S + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} [r(b^*(j)) p_S + b^*(j)\delta] \\ &= \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} [(1 - r(b^*(j)))\alpha F_S + b^*(j)\delta] + \eta. \end{aligned}$$

Now consider total costs if $F_S = 0$ and p_S is chosen to bind (5), that is, $\sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} r(b^*(j)) p_S = \eta$. Observe that the participation constraint is now automatically satisfied (with equality) by choosing $w_S = 0$. This mechanism yields the total cost

$$TC_1 = \eta + \sum_{j \geq x(F_A)} n_{c_j} q^j (1-q)^{n-j} b^*(j)\delta,$$

which is lower than TC_0. Therefore F_S cannot be positive.

Since w_S does not affect supervision effort and appears as a cost item in the principal's objective (1), it is optimal to set $w_S^* = 0$. **Q.E.D.**

Proof of Lemma 3. We provide only an heuristic argument. Because in the optimal mechanism the supervisor's effort incentive constraint (5) must be binding and $F_S^* = 0$ (Lemma 2), the supervisor's expected rewards is always set equal to η by appropriate choice of p_S. Thus, F_A affects the principal's cost objective only through the expected social costs associated with agent's optimal bribe solicitation strategy $b^*(j)$. Since $b^*(j)$ is (weakly) decreasing in F_A, overall costs will be minimized by setting $F_A^* = \overline{F}$.[26] **Q.E.D.**

[26] More precisely, an increase in F_A will lower the expected social costs if it induces a sufficiently large increase in the bribe trigger $x(F_A)$ and a fall in the interior optimal number of bribes $k^*(F_A)$ so that for at least some j values (the number of low-quality applicants) the agent switches from soliciting j bribes to no bribe at all, while for large j turnouts the agent finds it optimal to accept a smaller number of bribes. See Figure 1.

Proof of Proposition 2. The set of feasible incentives under collusion-proofness is a proper subset of the feasible incentives in the absence of collusion. Therefore the principal's expected overall costs under collusion-proof incentives cannot decrease. To show that the costs would strictly increase, consider the following two mutually exclusive cases under the optimal CP-mechanism: (A) $F_A^{CP} < \overline{F}$; (B) $F_A^{CP} = \overline{F}$. In case (A), $b^{CP}(j) > b^*(j)$ (assuming the agent's bribe solicitation is strictly decreasing in the penalty F_A) would increase social costs of corruption with no reduction in expected enforcement cost for the principal (because (6) will have to be satisfied), thus pushing up the overall costs. In case (B), $p_S^{CP} + \alpha F_S^{CP} \geq (1-\alpha)\overline{F} > p_S^* + \alpha F_S^*$, which implies either $p_S^{CP} > p_S^*$ or $F_S^{CP} > F_S^*$ (or both): since $b^{CP}(j) = b^*(j)$, the first clearly increases the expected enforcement cost for the principal and thus the overall costs; the latter implies $F_S^{CP} > 0$, so to satisfy the participation constraint in (6) the expected enforcement cost must again strictly exceed η, pushing up the overall costs. We verify below the remainder of the proposition.

[*i*] $w_A^{CP} = 0$ follows by the same reasoning as in the (*)-mechanism. That $w_S^{CP} = 0$ follows from the observation that the principal will do no worse by adjusting p_S rather than setting w_S positive to satisfy (6); adjusting p_S rather than w_S has the additional benefit of facilitating ICC and CPC.

[*ii*] First we claim that *in the optimal CP-mechanism at least one of the two constraints – CPC and ICC – must bind.* Suppose not so that (7) and (5) both hold with strict inequality; the participation constraint (6) may hold with or without equality. Let us maintain F_A at its optimal CP-level so that $b^{CP}(j)$ is unaffected. Consider now two cases: (1) both $p_S^{CP} > 0$ and $F_S^{CP} > 0$; (2) $p_S^{CP} > 0$ and $F_S^{CP} = 0$.[27] In case (1), lower both p_S and F_S slightly in the same way as in Lemma 2 (see footnote 25) so that the constraints (5) and (6) are satisfied. The collusion-proofness constraint (7) is satisfied because the changes in p_S and F_S are small. Overall, the principal's costs fall because p_S is smaller, contradicting the optimality of the proposed solution. In case (2), because ICC is non-binding and $F_S^{CP} = 0$, the participation constraint (6) must be holding with strict inequality. So if p_S is lowered slightly, the principal's expected costs will fall while all three constraints (5), (6) and (7) continue to be satisfied, a contradiction.

[27]The third case, $p_S^{CP} = 0,\ F_S^{CP} > 0$, does not arise as it violates (6).

Now we show that, in fact, the CPC would always bind. Suppose not, so that $p_S^{CP} + \alpha F_S^{CP} > (1-\alpha) F_A^{CP}$. This implies, given our argument above, that the ICC must be binding:

$$\sum_{j \geq x(F_A^{CP})} n_{c_j} q^j (1-q)^{n-j} r(b^{CP}(j))(p_S^{CP} + \alpha F_S^{CP}) = \eta.$$

Also, it must be that $F_A^{CP} = \overline{F}$. Suppose not. Then F_A can be increased to induce a fall in $b^{CP}(j, F_A)$, and an appropriate increase in p_S will keep ICC binding. As a result, the principal's overall costs will fall (on balance enforcement costs remain unchanged while social costs are lowered), contradicting the optimality of the CP-mechanism. Therefore, $F_A^{CP} = \overline{F}$, $b^{CP}(j, F_A^{CP}) = b^*(j)$ and

$$p_S^{CP} + \alpha F_S^{CP} = \frac{\eta}{\sum_{j \geq x(\overline{F})} n_{c_j} q^j (1-q)^{n-j} \cdot r(b^*(j))} = p_S^* < (1-\alpha)\overline{F},$$

which contradicts collusion-proofness (the first equality follows from the ICC being binding under the optimal CP-mechanism). Hence, the CPC must be binding.

[iii] Suppose $0 \leq F_S^{CP} < \overline{F}$ and the participation constraint does not bind. Then maintaining F_A^{CP} unchanged (so that bribery equals $b^{CP}(j)$), lower p_S (which must be positive because $w_S^{CP} = 0$) by a small $\epsilon > 0$ and correspondingly increase F_S by ϵ/α. These adjustments leave $p_S + \alpha F_S$ unchanged and thus would satisfy both (7) and (5). Also, in the supervisor's participation constraint (6) the change in the bracketed term under summation can be expressed as $r(\cdot)\cdot(\Delta p_S) - (1-r(\cdot))\alpha(\Delta F_S)$, which is $-\epsilon$. Since ϵ is small, (6) will not be violated. Given $b^{CP}(j)$ unchanged, the modification above reduces the principal's overall costs in (1), contradicting that the original CP-configurations were optimal. Hence (6) must be binding.

Now it is easy to verify that the principal should substitute away from the direct rewards p_S and rely on the penalty F_S to the extent possible (while satisfying (6)) if the supervisor is going to earn a positive rent; this adjustment lowers enforcement costs and total costs.

[iv] We claim that, if $F_S^{CP} = 0$ then $F_A^{CP} < \overline{F}$. To show this, suppose $F_S^{CP} = 0$ but $F_A^{CP} = \overline{F}$. By part [iii], the participation constraint (6) must bind, implying, given $w_S^{CP} = 0$, that the ICC will be binding and hence

$$p_S^{CP} = \frac{\eta}{\sum_{j \geq x(\overline{F})} n_{c_j} q^j (1-q)^{n-j} \cdot r(b^{CP}(j))}.$$

But with $F_A^{CP} = \overline{F}$ and therefore $b^{CP}(j) = b^*(j)$, it follows that $p_S^{CP} = p_S^*$. Also, by part $[ii]$

$$p_S^{CP} = (1-\alpha)\overline{F},$$

contradicting $p_S^* < (1-\alpha)\overline{F}$, thus establishing our claim.

From the above claim it also follows that, if $F_A^{CP} = \overline{F}$ then $F_S^{CP} > 0$.

Therefore, it must be that either $F_A^{CP} < \overline{F}$ or $F_S^{CP} > 0$.[28]

$[v]$ Finally, it is easy to see, given the result in part $[iii]$, that when $F_S^{CP} = \overline{F}$ the supervisor may have to be given a rent which the principal cannot capture under the limited liability (i.e, non-negative wages) assumption. That the agent may have to be given a rent follows from our analysis of the honest supervisor case in section 3.1. **Q.E.D.**

Proof of Lemma 4. Since the agent gets a nonnegative expected payoff in any equilibrium where he is induced to take a positive number of bribes, the principal sets $w_A^{BH} = 0$.

Next we show that $F_S^{BH} = 0$. Under the BH-scheme, the supervisor's incentive compatibility constraint is given by (8), which is independent of F_S. However, the penalty F_S appears in the supervisor's participation constraint (9). If (9) is not binding and $F_S > 0$, the principal can set $F_S = 0$ at no additional cost, because F_S appears nowhere else in the principal's problem, including the objective cost function. If (9) is binding and $F_S > 0$, again, setting $F_S = 0$ brings in no additional cost and does not violate any constraint; it will, however, strictly decrease costs if, in addition, $w_S > 0$: then F_S and w_S can both be decreased suitably without violating the supervisor's participation constraint, which reduces the wage bill, hence, costs.

Given the result $F_S^{BH} = 0$, observe that if the supervisor's incentive compatibility constraint (8) holds, the participation constraint stated above will also hold for any $w_S \geq 0$. Therefore $w_S^{BH} = 0$. **Q.E.D.**

Proof of Proposition 3. First note that while $p_S^* < (1-\alpha)\overline{F}$ so that the (*)-mechanism is not collusion-proof, by definition the configurations $(p_S^*, w_S^* = 0, F_S^* =$

[28] That both possibly may hold can be understood from the different cost implications of the two alternatives: positive F_S involves deadweight loss for which the supervisor must be adequately compensated (his participation constraint must be satisfied), whereas lowering of F_A would increase both enforcement costs as well as social costs of corruption; these two types of cost increases may be such that to balance the principal may compromise little bit in both directions.

$0, F_A^* = \overline{F}$) satisfy the supervisor's incentive compatibility and participation constraints, respectively (5) and (6).

Under the BH-mechanism, the "reward" the supervisor expects from detecting bribery is $(1-\alpha)F_A$ (by Lemma 4, $F_S^{BH} = 0$). Setting $F_A = \overline{F}$ maximizes the (indirect) reward and at the same time minimizes bribery through (3). Then the agent's strategy is $b^{BH}(j, \overline{F}) = b^*(j, \overline{F}) \leq b^*(j, F_A^{CP}) = b^{CP}(j, F_A^{CP})$. Given $(1-\alpha)\overline{F} > p_S^*$, the supervisor's incentive compatibility constraint (8) (the modified version of (5)) is satisfied for $b^{BH}(j, \overline{F}) = b^*(j, \overline{F})$. Also, given $w_S^{BH} = 0$ by Lemma 4, the participation constraint (9) (the modified version of (6)) is satisfied.

Since the principal incurs no direct reward costs under the BH-mechanism (which are incurred under both the (*)-mechanism and the CP-mechanism) and $w_A^{BH} = 0$ (by Lemma 4), compared to the CP-mechanism total expected costs will fall by at least η, and may even fall further if $b^{BH}(j, \overline{F}) < b^{CP}(j, F_A^{CP})$. **Q.E.D.**

References

[1] Banerjee, A.V., 1997. A theory of misgovernance. Quarterly Journal of Economics 112, 1289-1332.

[2] Bardhan, P., 1997. Corruption and development: A review of issues. Journal of Economic Literature 35, 1320-1346.

[3] Basu, K., Bhattacharya, S., and Mishra, A., 1992. Notes on bribery and the control of corruption. Journal of Public Economics 48, 349-359.

[4] Becker, G., 1968. Crime and punishment: An economic approach. Journal of Political Economy 76, 169-217.

[5] Becker, G. and Stigler, G.J., 1974. Law enforcement, malfeasance and the compensation of enforcers. Journal of Legal Studies 3, 1-19.

[6] Besley, T. and McLaren, J., 1993. Taxes and bribery: The role of wage incentives. Economic Journal 102, 119-141.

[7] Hindriks, J., Keen, M. and Muthoo, A., 1999. Corruption, extortion and evasion. Journal of Public Economics 74, 395-430.

[8] Itoh, H., 1993. Coalitions, incentives, and risk sharing. Journal of Economic Theory 60, 410-427.

[9] Kessler, A.S., 2000. On monitoring and collusion in hierarchies. Journal of Economic Theory 91, 280-291.

[10] Klitgaard, R., 1988. Controlling corruption. University of California Press, Berkeley.

[11] Kofman, F. and Lawarrée, J., 1993. Collusion in hierarchical agency. Econometrica 61, 629-656.

[12] Kofman, F. and Lawarrée, J., 1996. On the optimality of allowing collusion. Journal of Public Economics 61, 383-407.

[13] Laffont, J.J. and Tirole, J., 1991. The politics of government decision making. Quarterly Journal of Economics 106, 1089-1127.

[14] Marjit, S. and Shi, S., 1998. On controlling crime with corrupt officials. Journal of Economic Behavior and Organization 34, 163-172.

[15] Mookherjee, D. and Png, I., 1995. Corruptible law enforcers: How should they be compensated? Economic Journal 105, 145-159.

[16] Olsen, T.E. and Torsvik, G., 1998. Collusion and renegotiation in hierarchies: A case of beneficial corruption. International Economic Review 39, 413-438.

[17] Shleifer, R. and Vishny, A., 1994. Corruption. Quarterly Journal of Economics

[18] Tirole, J., 1986. Hierarchies and bureaucracies: On the role of collusion in organizations. Journal of Law, Economics and Organization 2, 181-214.

[19] Tirole, J., 1992. Collusion and the theory of organizations. In: J.J. Laffont, ed., Advances in Economic Theory, Sixth World Congress 2, 151-206, Cambridge University Press, Cambridge.

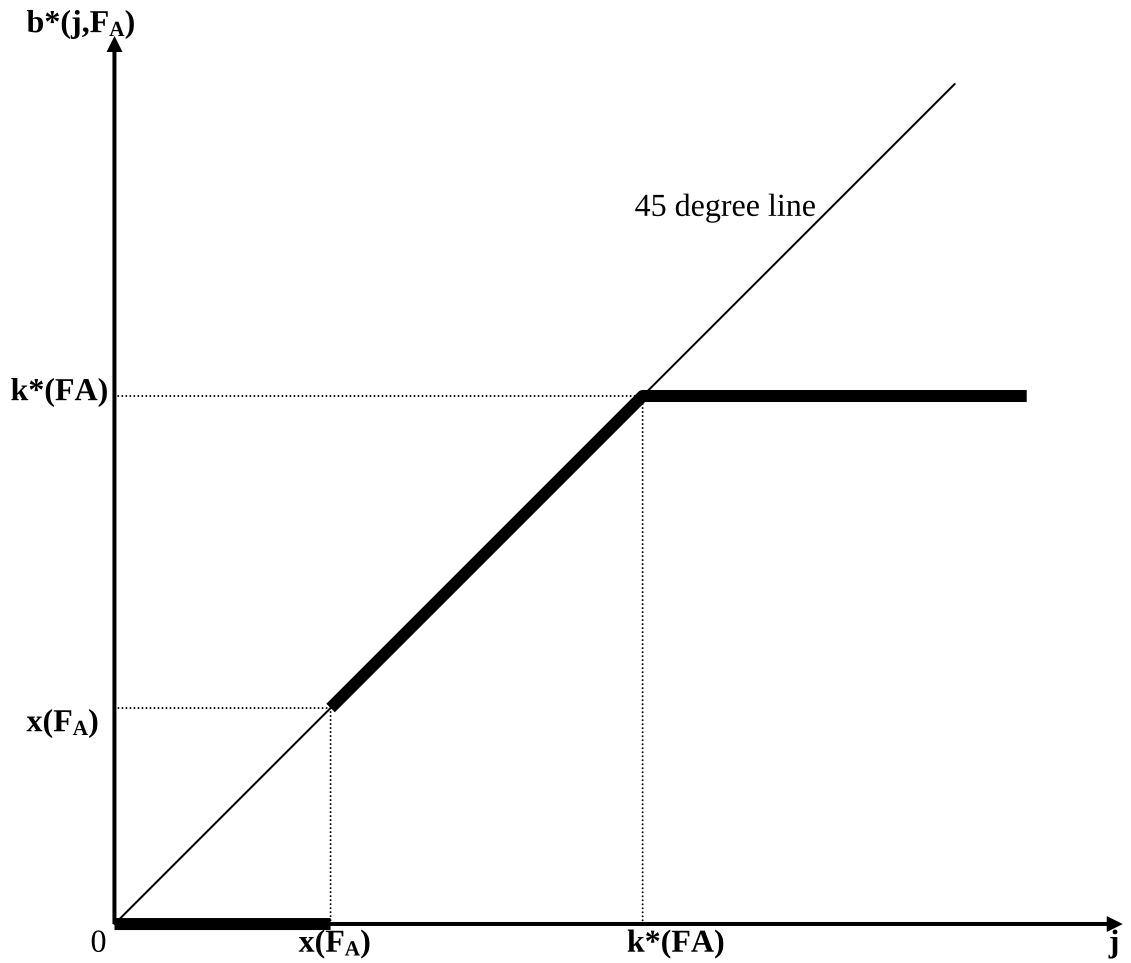

Figure 1: The agent's bribe solicitation strategy as a function of the number of low-quality applicants given the penalty F_A.