Collusion and the Organization of the Firm*

Alfredo Burlando† and Alberto Motta‡

May 11, 2014

Abstract

This paper shows that the threat of collusion between a productive agent and the auditor in charge of monitoring production can influence a number of organizational dimensions of the firm, including outsourcing decisions and the allocation of production costs. We find that the optimal organizational response to *internal collusion* lets the agent choose between working outside the firm with no monitoring, or working within the firm with monitoring. In equilibrium, there are no rents due to collusion and the efficient worker works outside the firm. The results are robust to a number of extensions.

JEL classification: D82; C72; D23
Keywords: Collusion; Supervision; Mechanism design; Theory of the Firm; Outsourcing; Insourcing

1 Introduction

If the market is an efficient method of resource allocation, why so many transactions take place within firms? This point lies at the heart of the economic literature that studies the

*We are especially grateful to Dilip Mookherjee and Antonio Nicolò for their help, suggestions and advice. We also wish to thank Richard Holden, Masatoshi Tsumagari, James Morley, Chris Ellis, Jaques Lawarrée, seminar participants at Boston University, University of Padua, University of Oregon, CEPET conference attendees, and two anonymous referees for their very helpful and insightful comments. Financial support from the Australian School of Business is gratefully acknowledged. Any remaining errors are ours.

†University of Oregon, e-mail: burlando@uoregon.edu.
‡University of New South Wales, e-mail: motta@unsw.edu.au

nature and boundaries of firms. In the attempt to rationalize the existence of firms, various theories have emerged.[1] Our paper departs from the rest of the literature by focusing on a particular form of moral hazard: internal collusion. We show that this alone can influence a number of organizational dimensions including the outsourcing decisions that a firm makes, the allocation of claims over production costs, and the ability of auditors to affect production. We study a version of the Laffont and Tirole (1991) asymmetric information model where a principal hires an agent to perform a task. The principal does not know the agent's productivity, but can hire a supervisor or auditor to gather this information. Collusion arises when the agent offers a bribe to the auditor to prevent her from reporting certain information. This type of collusion generates costly information rents.

Given this setup, we ask: "What is the optimal organizational response to collusion?" We find that the optimal mechanism consists of a menu of organizational structures. We present a simple example, where the agent can be either efficient or inefficient, and where he is offered the choice of two organizational structures. In the first, which we refer to as the "outsourcing contract," the agent is a residual claimant: he receives a fixed payment from the principal, takes responsibility over all production costs, and therefore need no monitoring by the principal. In the second, which we refer to as the "insourcing contract", the residual claimant is the principal. Here, monitoring is needed because the principal owns the production process, and so it must pay the realized costs of production. The principal offers both types of contract to the agent, and it is the agent that makes the organizational choice. In equilibrium, we find that high productivity agents select the outsourcing contract and work outside the firm, whereas low productivity agents choose to be in the firm and select the insourcing contract.[2] As a result, the threat of collusion is eliminated at zero cost, and the outcome achieved by the principal is "second best" —that is, it mimics one in which the principal has complete control over the auditor's actions. This is as a consequence of the equilibrium outcome, where auditors have no ex-post informational advantage over the principal and, consequently, are unable to retain any information rents.

The result of our paper—that collusion is solved without a penalty for the principal—is valid whenever the auditor and the agent cannot collude on their participation decisions,

[1] Gibbons (2005) provides an insightful comparison of a number of these theories, which are distilled from important contributions by Hart, Holmstrom, Klein, Williamson, and others. See also Aghion and Holden (2011) for an excellent literature review.

[2] Our optimal solution satisfies the "no distortion at the top" property. The most efficient type's effort is not subject to any distortion because no one has an incentive to mimic him. Hence, his optimal effort is the same in all states of the world, irrespective of the auditor's signal. This renders the auditor completely useless to the principal on the equilibrium path when the most efficient type is realized (in our case, the outsourcing contract). But it is still crucial to have monitoring off the equilibrium path in case the most efficient agent tries to misreport his type.

meaning that they cannot exchange information and coordinate their reports to the principal before agreeing to the contract. It is thus applicable in those contracting environments where both auditors and productive agents are uninformed about one another, and learn about their respective identity only after they have agreed to work for the principal. It can explain why principals might not want to disclose (or even select) the auditing firm during negotiations with agents. It can also explain the practice of forced job rotation, where auditors and agents are continually reassigned. Our mechanism is weakened if the auditor and the agent can devise a side agreement before the contracting stage with the principal, a situation we refer to as *ex-ante collusion*. In many instances, this situation is all but inevitable. For example, the principal might need an expert opinion from an auditor with significant connections to the contracting agent, perhaps because the task can be evaluated only by someone with "insider" knowledge. Another situation might arise in repeated games, in which the auditor and the agent interact over the course of several contract cycles. In those situations, the auditor and the agent may be able to create a scheme in which an efficient agent chooses to be monitored, and the auditor never reports to the principal that the agent is efficient (and, consequently, foregoes any incentive pay associated with that report). This side agreement, however, may be difficult to implement. As is discussed in section 4.2, there is a potential hold up problem in which the agent needs to deliver on the side contract early (i.e., at the signing of the contract), whereas the auditor delivers later (when it reports to the principal.) If this hold up problem exists, we find that our flexible organization can achieve the second best under certain conditions.

Finally, we show that alternative information structures can affect the performance of the contract. For instance, if the agent is fully informed at the participation stage (i.e., he knows not only his type but also the auditor's information), the flexible organization setup cannot achieve the second best. On the other hand, we find that it does not matter when the auditor receives her private information.

The results from the two types model generalize naturally to the multiple types case. In a more complex environment with $N > 2$ types, the optimal flexible organization offers N possible contracts to the agent. All agent types self-select into their type-specific contract, auditors earn no information rents, the mechanism is collusion-proof, and the principal achieves the second best outcome; other elements of the 2-type mechanism are partially retained. However, we do find that monitoring is generally required, and it is completely inconsequential only when the agent is the most efficient type.[3]

The rest of the paper is organized as follows. The next section explains the relation

[3]Our model is thus quite dissimilar from the one proposed by Baron and Besanko (1984), in which agents also selected among a menu of contracts but where the selection reduces the probability of an audit.

to the literature. Section 2 defines the model and the benchmarks of collusion-free and collusion-proof supervision when there are two types of agent. Section 3 introduces the flexible organization, and measures it against the benchmarks. Section 4 discusses the extensions: 4.1 relaxes assumptions on the participation constraints, 4.2 discusses the importance of ex-ante collusion, 4.3 highlights the importance of the timing of the auditor's signal and 4.4 discusses the model with multiple agent types. Section 5 concludes.

1.1 Relation to the Literature

This paper contributes to two different strands of literature: the theory of the firm, and, more specifically, the mechanism design literature on collusive supervision.

1.1.1 Mechanism Design Literature

Our contribution builds on the classic Laffont and Tirole (1991) model, hereafter LT, where in contrast we assume that the auditor's signal and the agent's type are not revealed simultaneously. In this context, we show that allowing the principal to offer a menu of contracts (or organizational structures) can improve the principal's profit.

Faure-Grimaud, Laffont and Martimort (2003)—hereafter FLM—and Celik (2009) consider a similar principal-auditor-agent framework where the agent and auditor cannot collude on their participation decisions. Nonetheless, there are a number of underlying differences in the timing and structure of information and in the bargaining protocol. They look at a framework where (a) monitoring is soft (the auditor does not possess hard information), (b) there is no shadow-cost of transferring payments from the agent to the auditor (i.e., no transaction costs within the coalition), (c) both the agent and the auditor know the auditor's signal before making their participation decisions, and (d) there is some residual asymmetric information within the coalition (after the auditor's learn her signal).

Owing to these differences, their results does not apply to the setup considered in this paper. For example, we don't require residual asymmetric information within the coalition for monitoring to be beneficial. Also, the distribution of bargaining power within the coalition is irrelevant in our framework. Finally, the principal's expected profits under the optimal mechanism in FLM (2009) and Celik (2009) fall in the presence of risk aversion, because the optimal mechanism leverages uncertainty—stemming from (d)—to deal with the collusion problem. On the contrary, the performance of our mechanism improves with risk aversion, because our mechanism provides full insurance to the agent and the auditor. Indeed, in the extreme case where the agent and the auditor are infinitely risk averse it achieves the first best. Motta (2012) considers the same framework as FLM (2003) and Celik (2009) and

shows that, even in those environments, allowing for menus of mechanisms can improve the principal's outcome. Kessler (2000) shows similar improvements when the auditor is tasked with monitoring the agent's *actions* as opposed to their costs.

1.1.2 The Theory of the Firm

Gibbons (2006) provides an excellent framework to contextualize our contribution. He defines and compares elemental versions of four theories of the firm: (1) rent seeking (starting with Williamson, 1971), (2) property rights (starting with Grossman and Hart, 1986), (3) incentive system (starting with Holmstrom and Milgrom, 1991, 1994), (4) adaptation (starting with Simon, 1951). Our paper speaks mostly to (3), but there are some analogies with the other theories that are worth mentioning.

As in the adaptation theory, our model considers two parties that face some uncertainty in the production process and must choose between (a) agreeing on precise allocations before uncertainty is resolved or (b) assigning authority to one party, who will then take allocation decisions after uncertainty is resolved (a situation referred to as employment contract in the adaptation literature). This framework implies a tradeoff between flexibility and exploitation: the subordinate can sacrifice flexibility by selecting (a), or she can allow the party in charge to decide later, and risk exploitation.[4] In our framework, uncertainty arises from the agent and the principal not knowing the auditor's information at the stage where the contracts are signed. In this situation, we show in section 3.3.1 that the tradeoff between flexibility and exploitation is an outcome of the optimal mechanism. Our solution then endogenizes the tradeoff at the core of the adaptation theory.[5] Our optimal mechanism offers a choice between (a) negotiating before uncertainty is resolved (i.e., the outsourcing contract where the auditor's information is not available) or (b) allocating a measure of authority to the agent after the uncertainty is resolved (i.e., the insourcing contract where the auditor's information is available). Our insourcing contract is peculiar in that the agent and the auditor shares some influence over production, which is further limited by the design of the principal's contract.

From this perspective, our model is closer to the "incentive system" theory. The latter focuses on the incentive problem between a principal and an agent. By analyzing internal incentives, this theory offers a potential explanation for the observation that incentives offered to employees in firms are low-powered relative to the high-powered incentives offered to independent contractors in markets. In the incentives theory, firms choose low-powered

[4] Note that the parties choose not only between (a) and (b), a theory of employment, but also the allocation of decision rights in (b), a theory of the firm.

[5] Another common feature is that our model does not require the notion of specific investments. This is instead the focus of the rent-seeking and property-rights theories.

incentives because the agent responds to a given contract differently as an employee than she would as a contractor. We obtain the same result but the underlying mechanism is different. Instead of being due to asset ownership, our result is a byproduct of the adverse selection problem, where efficient types work outside the firm (and are offered high-powered incentives) and inefficient types work within the firm (and obtain low-powered incentives.)

Finally, let us consider the rent-seeking theory. The latter highlights the presence of quasi-rents that originate when different firms engaged in a relationship own assets that are more valuable within the relationship than outside of it (Holmström and Roberts, 1998). Since this induces firms to engage in costly haggling for appropriation of such quasi-rents, it is more efficient to have one player appropriating all assets and thus incorporating all quasi-rents without engaging in rent-seeking behavior. A key assumption is that rent-seeking is more pronounced across firms than within firms. Our model looks at this problem from a different perspective: while collusion does create information rents, these arise *within* the firm and not between firms. Perhaps unsurprisingly, then, our model reaches the opposite conclusion than the transaction cost literature: outsourcing is a way to *reduce rents within the firm*. While our setup does not consider transaction costs between firms, the model could be integrated in the future in standard transaction costs models and used to evaluate the relative importance of internal versus external rent-seeking behavior.

2 The Model

2.1 Setup

We begin with a discussion of an economic environment similar to the one presented in LT. The problem under consideration involves a firm (which we call the principal) that markets one unit of a good with value G, and hires an agent to produce the good. The agent's underlying efficiency is unknown to the principal, which is why it hires a manager (whom we call the auditor) to find out. All parties are risk neutral.

The agent: The agent produces the good at cost

$$C = \beta - e \tag{1}$$

where β represents the technology parameter, which can take one of the two values: "efficient" ($\underline{\beta}$) with probability v and "inefficient" ($\overline{\beta}$) with probability $(1-v)$. The agent knows the realization of β. By exerting effort $e \in \mathbb{R}_+$, the agent reduces cost of production but he incurs an increasing and convex disutility $\psi(e)$, where $\psi' > 0$ and $\psi'' > 0$. The principal pays the agent's costs and it collects the revenue. Let t denotes the transfer from the principal to

the agent. We assume that the agent is able to quit at any time with no penalty other than whatever costs he has already incurred. Due to the timing of the model (which is explained in detail below), this will give rise to the following ex-post participation constraint:

$$PC_a : U = t - \psi(e) \geq 0, \tag{2}$$

where we normalized the agent's reservation utility to 0.[6]

The auditor: The agent's manager (auditor) is hired to obtain information on the agent's cost structure, and pass that information along to the principal. The auditor receives a signal σ which provides verifiable information about the true efficiency level ($\sigma = \beta$) with probability ξ, and no information at all ($\sigma = \varnothing$) with probability $1-\xi$. The auditor receives a payout s from the firm. Like the agent, the auditor can quit at any time, which entails that the payoff to her must be sufficiently high to meet the reservation utility (which we normalize to zero) in every state of the world:

$$PC_s : V = s \geq 0 \tag{3}$$

The principal: The principal observes neither β nor σ, although it knows the distribution function of each and it observes the cost C. The principal's payoff for a given cost of production C, transfer level t and payout s is

$$W(C, t, s) = G - C - t - s.$$

We assume that G is large enough such that the principal always prefers to produce irrespective of the agent's type.

Timing: We now highlight the timing of the model, which differs from LT in that the auditor's signal σ and the agent's type β are not revealed simultaneously: the agent learns his type before participating, but σ is revealed only after the participation decisions have been made.[7] This is consistent with a situation where the auditor performs her monitoring activities after signing the principal's contract, and therefore covers many situations that were not included in the original model. With this crucial distinction, the timing of the model is as follows. At stage 0, the agent learns about his type, and every one learns the distribution of β and σ. At stage 1, a contract is offered to the agent and the auditor. In the same period, they independently and without being able to communicate send participation

[6]The same participation constraint would arise under the assumption that the agent is covered by limited liability such that he must be assured non-negative rents.

[7]Our results would not change if only the auditor learns her signal at the participation stage, provided that she cannot communicate this information to the agent at that stage–see section 4.3.

messages. We call this the *participation* or *interim stage*. At stage 1.5, the agent and the auditor learn the signal σ, and can collude—that is, they can decide on side contracts and coordinated messages. We call this the *collusion stage*. At stage 2, the agent and the auditor exchange messages with the principal. At stage 3, the agent chooses the effort level, output is produced, and the principal earns G and pays s, t, and C.[8]

The cost parameter β together with the signal σ received by the auditor defines four *ex-post* states of the world:

$$\begin{aligned} p_1 &= \Pr(\beta = \underline{\beta}, \sigma = \underline{\beta}) = \nu\xi \\ p_2 &= \Pr(\beta = \underline{\beta}, \sigma = \varnothing) = \nu(1-\xi) \\ p_3 &= \Pr(\beta = \overline{\beta}, \sigma = \varnothing) = (1-\nu)(1-\xi) \\ p_4 &= \Pr(\beta = \overline{\beta}, \sigma = \overline{\beta}) = (1-\nu)\xi, \end{aligned}$$

where p_i is the probability of each correspondent state. Thus, the agent is efficient in states 1 and 2 and inefficient in states 3 and 4; and the auditor has full information in states 1 and 4 and no information in states 2 and 3.

It is important to highlight that the problem we consider requires two communication phases: stage 1 messages (prior to the signal being revealed) and stage 2 messages (after the signal is revealed). A general mechanism that takes these various communication phases into account can therefore be written as follows:

$$\Gamma = \{C(m_{a1}, m_{a2}, m_{s1}, m_{s2}), t(m_{a1}, m_{a2}, m_{s1}, m_{s2}), s(m_{a1}, m_{a2}, m_{s1}, m_{s2})\},$$

where $m_{i1} \in M_{i1}$ are i's stage 1 messages and $m_{i2} \in M_{i2}$ are i's stage 2 messages observed after σ is revealed.

Before proceeding, let us highlight some notable implications of our assumptions:

(i) Since the agent and the auditor could decline participation in the contract at any stage, M_{aj} and M_{sj} must include the exit option or, more formally, $Exit \in M_{aj}$ and $Exit \in M_{sj}$, for $j = 1, 2$. For simplicity, we assume that production is shut down and all payoffs are 0 if either $m_{aj} = Exit$ or $m_{sj} = Exit$. That is, both the auditor and the agent are indispensable for production.[9]

(ii) The principal can ignore any "soft message" from the auditor. This point will be proven formally later in the paper. Here we offer an intuitive explanation. Since the auditor

[8] LT adopts the same timing except that both the agent and the auditor learn the signal σ at stage 0.

[9] Assuming that the auditor is not indispensable for production would not change our main results, but it would slightly complicate the exposition.

has no private information at the participation stage, without loss of generality $M_{s1} = \{Participate, Exit\}$. On the other hand, the stage 2 message space M_{s2} depends on the information σ obtained by the auditor and can be conditioned on the messages (m_{a1}, m_{s1}) exchanged at the participation stages, i.e., $M_{s2}(\sigma, m_{a1}, m_{s1})$. If $\sigma = \varnothing$ the auditor and the principal share the same information, i.e., they are equally uninformed. Since the principal cannot benefit from communicating with the auditor, it is without loss of generality that the auditor message space can be reduced to $M_{s2}(\sigma = \varnothing, m_{a1}, m_{s1}) = \{\varnothing, Exit\}$ for any pair (m_{a1}, m_{s1}). If $\sigma = \beta$, the auditor is perfectly informed and obtains hard information. In this case, the principal always prefers to receive hard information from the auditor, rather than non-verifiable soft information. Hence, the auditor's message space can be reduced to $M_{s2}(\sigma = \beta, m_{a1}, m_{s1}) \subseteq \{\sigma, \varnothing, Exit\}$ without loss of generality. Moreover, the fact that signals are revealed sequentially does not create any further complication in our framework.[10]

(iii) The agent observes whether or not the auditor obtains the hard information σ, but he does not possess it himself. As in LT, the information available to the agent is soft in its entirety. This entails that M_{a2} does not depend on σ. However, M_{a2} can be conditioned on the participation stage messages and so $M_{a2}(m_{a1}, m_{s1})$.

(iv) Since the agent and the auditors can quit after learning their types and having exchanged messages with the principal and before effort has been applied, the relevant participation constraints are ex-post.[11]

We can now present the formal definition of a *single contract* mechanism:

Definition 1. *A mechanism is single contract if* $M_{a1} = \{Participate, Exit\}$.

Our focus on single contract mechanisms as a benchmark is justified by two observations. First, in existing models where parties acquire their private information simultaneously, it is without loss of generality that one can focus on single contracts; indeed, LT and subsequent models focus on this class of mechanisms. It is interesting to benchmark our model against

[10] In a sequential game, players might obtain information by observing the behavior of other players in previous stages (e.g., Myerson [1986]). In our multi-stage framework this complication does not arise. The only state of the world where the auditor has a non-trivial message space M_{s2} is also the state where the auditor is perfectly informed about the agent's type. There is no asymmetric information left between the agent and the auditor, so the fact that the auditor observes the message sent by the agent at stage 1 does not affect her incentive to report in stage 2. When the auditor observes $\sigma = \varnothing$, her belief about the agent's type at stage 2 may be updated following the observation of the agent's message choice at stage 1, something that could potentially complicate the analysis. However, whenever $\sigma = \varnothing$, the principal does not need the auditor's report, since the principal and the auditor share the same information.

[11] We discuss how our results would change with ex-ante participation constraints in section 4.1.

these. Second, in the following section we show that single contract mechanisms continue to achieve the upper bound welfare even under our information structure—provided that there is no collusion between the parties.[12] However, this class of mechanisms does not achieve the upper bound welfare under collusion, while a more complex mechanism based on multiple contracts does. This leads to the main result of the paper: the threat of internal collusion generates alternative organizational structures.

2.2 Collusion-free Monitoring

We first consider the benchmark case, in which the auditor always reports truthfully and cannot misreport or hide the signal from the principal. We denote the optimal outcome implementable by the principal under these circumstances as *second best*, to distinguish it from the *first best* case where the principal knows the agent's type β. The second best scenario is akin to a situation where the principal pays s to receive—and directly control—a monitoring technology allowing it to learn the true efficiency of the agent with probability ξ. Since the auditor always reports what she knows, the efficient contract requires the auditor to receive her outside option in all circumstances.[13] Under the assumptions presented in the previous section:

Proposition 1. *In the absence of collusion, it is without loss of generality to focus on direct revelation mechanisms with single contract.*

Proof. See Appendix. □

This result has an important implication: in the absence of collusion, the principal does not benefit from eliciting information at the participation stage. Without loss of generality, all the communication between the principal and the agent can take place in stage 2. In the next section we show that this is no longer the case when collusion is allowed. Thus, more complex organizational structures—including the option between outsourcing and insourcing—emerge as a response to the threat of internal collusion.

In the proof of Proposition 1 we show that it is without loss of generality to restrict attention to a direct revelation mechanism where the auditor reports $m_{s2} = \sigma$ and the agent reports $m_{a2} = \beta$ (or equivalently the agent could be rewarded on the basis of m_{s2} and the observable C). Hence, we can focus on incentive schemes of the form $t(C(m_{a2}), m_{s2})$, $s(C(m_{a2}), m_{s2})$. Here we offer a simple description of the optimal mechanism. For simplicity,

[12] In fact, in the absence of collusion, the upper bound under our information structure is the same as in LT.

[13] Of course, the assumption that the outside option of the auditor is 0 can be relaxed; in that case, a centralized organization could be expensive but necessary.

we denote by t_i, s_i, and C_i respectively the transfers and the cost when the agent and auditor report messages consistent with state i. Moreover, we denote by e_i the effort level that is implied by C_i when the agent's type is consistent with state i. Note that in states 2 and 3, monitoring does not provide any information on the cost of the agent. If the agent is efficient, then, he can report $\overline{\beta}$ and choose an effort level e that would mimic the production cost of the inefficient agent: $\underline{\beta} - e = C_3 = \overline{\beta} - e_3$. In that case, the effort that the efficient agent would provide is $\psi(e_3 - \triangle\beta)$, where $\triangle\beta = \overline{\beta} - \underline{\beta}$. To discourage this behavior, the principal needs to offer an incentive compatible contract to the efficient agent:[14,15]

$$IC_a: \quad t_2 - \psi(e_2) \geq t_3 - \psi(e_3 - \triangle\beta). \tag{4}$$

Finally, the ex-post participation constraints (2) and (3) must be met. The optimal contract with the agent solves the following Collusion-Free programme (CF):

$$\max_{\{e_i,t_i\}_{i=1,\ldots,4}} W = G - \sum_{i=1}^{4} p_i(s_i + t_i + \beta_i - e_i) \tag{5}$$

$$\begin{aligned}
s.t. & \\
PC_a: & \quad t_i - \psi(e_i) \geq 0, \ for\ all\ i \\
IC_a: & \quad t_2 - \psi(e_2) \geq t_3 - \psi(e_3 - \triangle\beta) \\
PC_s: & \quad s_i \geq 0, \ for\ all\ i
\end{aligned}$$

The solution to the maximization program (5), involves the following:

(a) agents earn zero rents when they are inefficient (state 3 and 4) and when they are efficient but the signal is informative (state 1): $t_i = \psi(e_i)$ for $i = 1, 3, 4$;

(b) the auditor never earns any rents: $s_i = 0$ for $i = 1, 2, 3, 4$;

(c) when the signal is not informative and the agent is efficient (state 2), the agent is paid $t_2 = \psi(e_2) + \Phi(e_3)$, meaning that he gains positive information rents, where

$$\Phi(e_3) = \psi(e_3) - \psi(e_3 - \triangle\beta) > 0.$$

[14] If e_3 is sufficiently small the efficient type might be unable to mimic the inefficient one. For the sake of exposition, we rule out this possibility by focusing on parameters values such that the optimal e_3 is larger or equal than $\triangle\beta$, that is, $\psi'(\triangle\beta) \leq (1-\nu) - \nu\psi'(0)$.

[15] The mechanism has a second incentive compatibility constraint that ensures that the inefficient agent never chooses the efficient agent's cost. As usual in these cases, this constraint is never binding.

(d) effort levels of the agent solve the following first order conditions:

$$\begin{aligned} \psi'(e_i^*) &= 1 \qquad for\ i = 1,2,4 \\ \nu\Phi'(e_3^{CF}) + (1-\nu)\psi'(e_3^{CF}) &= (1-\nu), \end{aligned} \tag{6}$$

where the superscript CF denotes the collusion-free outcome. The equilibrium outcome is typical for this type of problems: the efficient agent's effort is always at the optimal level, whereas the inefficient agent receives low-powered incentive in one state of the world, since $e_3^{CF} < e_4^*$. The optimized profit function under a single contract when auditors cannot be bribed is

$$W_{CF}^* = G - \sum_{i=1,2,4} p_i \left\{\psi(e_i^*) + \beta_i - e_i^*\right\} - p_3 \left\{\psi(e_3^{CF}) + \overline{\beta} - e_3^{CF}\right\} - \nu(1-\xi)\Phi(e_3^{CF}) \tag{7}$$

where $\beta_i = \underline{\beta}$ for $i = \{1,2\}$ and $\beta_i = \overline{\beta}$ for $i = \{3,4\}$.

2.3 Collusion-proof Monitoring: Single Contract Mechanism

In this section we allow for collusion. Let us denote $K \equiv \{(\beta,\sigma) \mid \beta \in \{\underline{\beta},\overline{\beta}\}, \sigma \in \{\underline{\beta},\overline{\beta},\varnothing\}\}$. The side agreement is defined as follows:

Definition 2. *A side agreement is a set of functions $\{m(\beta,\sigma), b(\beta,\sigma)\}$ where $m : K \to M_{a2} \times M_{s2}$ denotes a coordinated message (m_{a2}, m_{s2}) sent to the principal, and $b : K \to \Re$ denotes the associated side payment from the agent to the auditor.*

As standard in the literature on collusion, the side agreement is assumed to be enforceable. If the agent or the auditor refuses the side agreement, the principal's mechanism is played non-cooperatively. In addition, we follow standard practice by assuming a shadow cost $\lambda > 0$ of transferring payments to the auditor, arising for instance from the fact that effort must be exerted to avoid detection by the principal.[16]

The mechanism presented in Section 2.2 is vulnerable to this kind of side agreements. The problem arises when the agent is efficient: he gains $\Phi(e_3)$ if the auditor reports $m_{s2} = \varnothing$, and gains nothing if she reports $m_{s2} = \underline{\beta}$. In that case, collusion involves a side payment from the efficient agent to the auditor and the coordinated message $m(\underline{\beta},\underline{\beta}) = (\underline{\beta},\varnothing)$.

With this problem, the features of the optimal single contract mechanism under collusion (hereafter, CP) are described in the following proposition.

[16] In many instances, detection can be more easily avoided when the bribe is in the form of a non-monetary gift. In those cases, λ can be thought of as the difference between the monetary cost and the utility value of the gift.

Proposition 2. *In the presence of collusion and within the class of single contract mechanisms, it is without loss of generality to focus on collusion-proof mechanisms. The threat of collusion is costly to the principal.*

Proof. See the appendix.

The intuition for the result is as follows. Denote by U_i and V_i the ex-post utility of the agent and the auditor in state i. Note that state 1, $(\beta = \underline{\beta}, \sigma = \underline{\beta})$, is the only state where there is an incentive to collude. If the (informed) auditor hides her information and reports $m_{s2} = \varnothing$, she foregoes payment s_1 and receives instead payment s_2. She is willing to do this only if $s_2 + b(\underline{\beta}, \underline{\beta}) \geq s_1$. On the other hand, the agent is willing to offer the bribe only if $U_2 - U_1 \geq (1+\lambda)b(\underline{\beta}, \underline{\beta})$. Since $U_2 - U_1 = \Phi(e_3)$, there is no side agreement in state 1 if:

$$IC_s : s_1 \geq s_2 + \frac{1}{1+\lambda}\Phi(e_3). \tag{8}$$

Note that in state 1 there is no asymmetric information within the coalition agent-auditor. Hence, the notion of collusion proofness embedded in this incentive constraint must be a strong one: it does not require any restriction on the allocation of bargaining power inside the coalition, and it does not hinge on the identity of the coalition member who offers and initiates the collusive agreement.

We now describe the outcomes under the optimal contract. First, the CP contract is characterized by points (a) and (c) from the CF solution. Second, auditor's payoffs are as follows:

(b') $s_i = 0$ for $i = 2, 3, 4$; $s_1 = \frac{1}{1+\lambda}\Phi(e_3)$.

That is, under CP, constraint (8) is binding, the auditor earns rents in state 1, and the profits to the principal are reduced. Note that, if (8) is violated, a collusive agreement exists such that the agent and the auditor misreport state 1 as state 2. But the principal does not benefit from the auditor's signal in the remaining states 2, 3 and 4. Hence auditing is not beneficial to him and the problem reduces to a standard principal-agent contract. The fact that the optimal solution requires the principal to hire the auditor and to ensure that (8) is satisfied owes to the the presence of the transaction cost $\lambda > 0$. With this cost, the information rent required to prevent the coalition from mispreporting state 1 is $\frac{1}{1+\lambda}\Phi(e_3)$, whereas the information rent in a standard principal-agent model in the same state is $\Phi(e_3)$. Hence, the principal always prefer to hire the auditor and prevent collusion.

Finally, effort levels under CP are as follows:

(d')

$$\begin{aligned} \psi'(e_i^*) &= 1, \quad i = 1,2,4 \\ \Phi'(e_3^{CP})\left[\frac{\xi\nu}{1+\lambda} + \nu(1-\xi)\right] + (1-\nu)(1-\xi)\psi'(e_3^{CP}) &= (1-\nu)(1-\xi), \end{aligned} \tag{9}$$

that is, there is an additional distortion away from efficiency that is due to the possibility of collusion between the auditor and the agent: note that $e_3^{CP} < e_3^{CF}$. This distortion is due to the trade off between allocation efficiency and the agent's information rents in state 2, $\Phi(e_3)$, as well as the auditor's transfer in state 1, $s_1 = \frac{1}{1+\lambda}\Phi(e_3)$. A comparison of CP and CF reveals that $W_{CF}^* \geq W_{CP}^*$: indeed, for a given e_3 we have,

$$W_{CF}(e_3) = W_{CP}(e_3) + \frac{\xi\nu}{1+\lambda}\Phi(e_3), \tag{10}$$

and from (10) we get that the cost of collusive monitoring is $\frac{\xi\nu}{1+\lambda}\Phi(e_3)$.

3 Flexible organization

Under the threat of collusion, we now show that the principal can implement a better outcome by not restricting attention to single contract mechanisms. By offering a menu of contracts the principal can implement the second best outcome (CF), making collusion costless to eliminate. We propose an intuitive explanation of the mechanism in section 3.1.

3.1 Intuition

We now show that our Flexible Organization (FO) can achieve the second-best outcome even in the presence of collusion. Before getting to the model, it is useful to revisit the payoffs to the auditor and the agent in the second best (CF) and third best (CP) outcome. When the agent is efficient (i.e., $\beta = \underline{\beta}$), payoffs to agents ($U$) and auditor ($V$), contingent on the signal σ, are represented by the following table:

Collusion free (CF)

signal σ	*state*	*prob.*	U	V
informative	1	ξ	0	0
uninformative	2	$1-\xi$	$\Phi(e_3)$	0

Collusion proof (CP)

signal σ	*state*	*prob.*	U	V
informative	1	ξ	0	$\frac{1}{1+\lambda}\Phi(e_3)$
uninformative	2	$1-\xi$	$\Phi(e_3)$	0

A couple of aspects are worth noticing. First, in expectation the efficient agent earns a rent of $(1-\xi)\,\Phi(e_3)$ in both cases, whereas the auditor earns a rent of $\frac{\xi}{1+\lambda}\Phi(e_3)$ in CP but of 0 in CF. Second, under CP, both the agent and the principal have reasons to dislike state 1. The agent dislikes it because it earns zero rents; the principal dislikes it because it must pay out the auditor.

The principal can use this mutual dislike of state 1 by offering the agent the opportunity to avoid monitoring altogether. The principal can offer him the expected information rent $(1-\xi)\,\Phi(e_3)$ in *both* states 1 and 2, provided that he agrees to an alternative contract without monitoring. We call this the outsourcing contract. We will show that the optimal outsourcing contract has several useful characteristics. First, the efficient agent avoids the monitoring lottery and earns his expected information rent with certainty. Second, it is unappealing to the inefficient agent. Intuitively, this must be the case because, if the inefficient agent chooses to avoid the efficient agent contract under monitoring, he should likewise avoid a contract that requires the same amount of high effort but a lower payoff. Third, it requires the efficient agent to reveal his type to the principal directly, in the absence of monitoring.

These intuitions translate to a more general case where there are many different possible types of agents (see section 4.4). In particular, what is relevant here is that the flexible organization mechanism elicits the private information of the agent at the participation stage, in exchange for a type-specific payoff that (in equilibrium) does not depend on the signal of the auditor.

3.2 The Contract

In this section we introduce the FO structure by removing the restriction that the principal needs to offer a single contract mechanism. We first establish an upper bound on expected profit by deriving few necessary conditions that must be satisfied in any equilibrium. We find that allowing for menu of contracts changes the upper bound in the following fashion:

Proposition 3. *The upper bound on expected profit is equal to the one the principal can achieve in the collusion-free benchmark.*

Proof. See the appendix. □

In a sense, the proposition is the first main result of the paper: it demonstrates that the presence of internal collusion need not reduce the upper bound on profits.

We next show the second important result: the principal can indeed attain this upper bound on expected profit by offering a simple FO mechanism that allows for multiple

organizational structures. The mechanism is constructed in the following way. At the participation stage, the agent is offered three options: an outsourcing contract with no monitoring, an insourcing contract with monitoring, or rejection of any contract (no participation): $M_{a1} = \{Insourcing, Outsourcing, Exit\}$. The auditor, on the other hand, simply accepts or refuses to participate.

The insourcing contract The insourcing contract is similar to the CP contract presented in Section 2.3. It features the same message spaces $M_{a2}(m_{a1} = Insourcing, m_{s1} = Participate) = \{\overline{\beta}, \underline{\beta}, Exit\}$, $M_{s2}(m_{a1} = Insourcing, m_{s1} = Participate) = \{\overline{\beta}, \underline{\beta}, \varnothing, Exit\}$ and it specifies the same transfer functions and effort levels, where the only difference is that the effort level in state 3 is now e_3^{CF} instead of e_3^{CP}. Thus, transfers are structured as follows: **(a)** $t_i = \psi(e_i^*)$ for $i = 1, 3, 4$; **(b)** $s_i = 0$ for $i = 2, 3, 4$; **(c)** $t_2 = \psi(e_2^*) + \Phi(e_3^{CF})$; **(c)** $s_1 = \frac{1}{1+\lambda}\Phi(e_3^{CF})$.
It is readily observed that the insourcing contract has similar features to the CP contract: in particular, the contract satisfies participation constraints (2) and (3), the incentive compatibility constraint (4), and the no-collusion constraint (8). Thus, if the agent selects the insourcing contract, both the agent and the auditor report truthfully in equilibrium; the agent earns positive rents only when he is efficient and the signal is uninformative; and the auditor earns rents only when she has hard information that the agent is efficient.

The outsourcing contract The outsourcing contract does not depend on any input of the auditor and the agent: $M_{a2}(m_{a1} = Outsourcing, m_{s1} = Participate) = \{\varnothing, Exit\}$ and $M_{s2}(m_{a1} = Outsourcing, m_{s1} = Participate) = \{\varnothing, Exit\}$. The principal makes an offer to the agent $(C_0, t_0) \in \Re_+^2$, where (i) $C_0 = \underline{\beta} - e_0$, (ii) $t_0 = \psi(e_0) + (1 - \xi)\Phi(e_3^{CF})$ and (iii) the effort level e_0 is first-best, i.e., it solves $\psi'(e_0) = 1$. The auditor is offered a payment $s_0 = 0$.

We are now ready to present our next proposition:

Proposition 4. *The principal can use the Flexible Organization to achieve the upper bound on expected profit derived in Proposition 3. Thus, the threat of collusion is eliminated at no cost and* $W_{FO}^* = W_{CF}^*$.

Proof. At the participation stage the agent decides which contract to select. This decision is taken non-cooperatively because collusion is not allowed at the participation stage. Given the structure of the FO, the efficient agent anticipates that by selecting the insourcing contract he won't be able to earn any information rent when the auditor's signal is informative: there is no side agreement that would convince the auditor to hide her information and would also leave the agent better off. Thus, the agent gains $t_1 - \psi(e_1^*)$ in case the signal is informative.

On the other hand, when the signal is uninformative, no information is revealed and the agent's payoff is $t_2 - \psi(e_2^*)$. If the agent instead were to choose the outsourcing contract, he would gain payoff $t_0 - \psi(e_0)$. In equilibrium, the efficient agent selects the outsourcing contract if condition

$$\underline{IC}_0 : t_0 - \psi(e_0) \geq \xi\,(t_1 - \psi(e_1^*)) + (1-\xi)(t_2 - \psi(e_2^*)) \tag{11}$$

is satisfied, where the right-hand side represents the agent's expected payoff in the insourcing contract and the left-hand side is the payoff in the outsourcing one. The outsourcing transfer t_0 and effort e_0 described above ensure that this condition is met with equality. The inefficient agent, on the other hand, selects the insourcing contract because the outsourcing one yields a negative payoff $t_0 - \psi(e_0 + \triangle\beta) < 0$.[17] Note also that, since there is no auditor in the outsourcing contract, the agent cannot hope to collude and avoid this negative payoff. Thus, the expected profit function under FO is

$$\begin{aligned} W_{FO}^* &= G - \nu(t_0 + s_0 + \underline{\beta} - e_0) - \\ &\quad -(1-\nu)[\xi(t_4 + s_4 + \bar{\beta} - e_4^*) - (1-\xi)(t_3 + s_3 + \bar{\beta} - e_3^{CF})]. \end{aligned} \tag{12}$$

Finally, plugging the FO transfers and effort levels in the expected profit function reveals that $W_{FO}^* = W_{CF}^*$. □

The auditor is useful in two ways: (i) it helps adjust the inefficient agent's effort level according to the realized state of the world; (ii) it discourages the efficient agent from selecting the insourcing contract. In addition, the FO mechanism is collusion-proof. The agent decides whether to select the outsourcing or the insourcing contract in a non-cooperative fashion because collusion is not allowed at the participation stage. In addition, the insourcing contract is by construction collusion-proof while under outsourcing there are no message exchanges (i.e., there is no reason for the auditor and the agent to establish a side agreement). Hence, a corollary of Proposition 4 is that there is no loss in generality in focusing on collusion-proof mechanisms.[18] Also, there is no loss in generality in assuming that the menu consists of only two contracts, i.e., $M_{a1} = \{Insourcing, Outsourcing, Exit\}$; more complex

[17] To see this, substitute $t_0 = \psi(e_0) + (1-\xi)\Phi(e_3^{CF})$ and rearrange to obtain $\psi(e_0 + \triangle\beta) - \psi(e_0) > (1-\xi)[\psi(e_3^{CF}) - \psi(e_3^{CF} - \triangle\beta)]$ which is true since, by convexity of $\psi(.)$ and the fact that $e_3^{CF} < e_0$, $\psi(e_0 + \triangle\beta) - \psi(e_0) > \psi(e_3^{CF}) - \psi(e_3^{CF} - \triangle\beta)$.

[18] The collusion-proof insourcing contract presented above has the additional benefit of maximizing the principal's profit in the (off the equilibrium) case where the efficient agent selects the insourcing contract. This could happen if the agent underestimates the probability of a successful audit ξ with a small probability. Given that the agent is indifferent between accepting the insourcing or the outsourcing contract, the slightest underestimation of ξ suffices. The same is not true for the inefficient type, whose choice is robust to small perturbations in the estimation of ξ.

mechanisms are not profit enhancing.

3.3 Remarks

3.3.1 Tradeoff Between Monitoring and Residual Claims

The optimal FO studied here introduces a tradeoff between monitoring and claims over production costs. In the outsourcing contract, the agent is not subject to monitoring but he is the residual claimant of the realized costs: the contract simply specifies a fixed payment, t_0, which depends neither on the agent's nor on the auditor's reports. Under the insourcing contract, on the other hand, the agent is no longer the residual claimant. For example, when $m_{s2} = \varnothing$, the agent can report that the realized cost is C_2 or C_3, and the principal must take note of that and react according to its contractual obligations.

3.3.2 Insurance

As we have already mentioned, our FO has some built-in elements of insurance in that the outsourcing contract ensures the efficient agent against the monitoring lottery. Naturally, allowing the agent to be risk averse would only improve FO's performance. Consider the extreme case where the agent is infinitely risk averse, so that he cares only about the worse payoff. The incentive compatibility constraint in (11) becomes

$$\underline{IC}_0 : U(t_0 - \psi(e_0)) \geq U(t_1 - \psi(e_1^*)), \tag{13}$$

which entails that the efficient agent is left with no rents in all states of the world and the principal implements the first-best outcome without asymmetric information.

4 Extensions

4.1 Ex-ante Participation Constraints

So far we have assumed that it is not possible for the principal to provide the agent and the auditor with any payoff that is *ex-post* negative. Consider now the case where the participation constraints are *ex-ante*. This would arise, for instance, if the agent agrees to the contract before learning about the state of the world, and cannot exit the contract before having exerted effort. Thus, the agent's and the auditor's participation constraints

are respectively,

$$PC_a^{EA} \quad : \quad EU = \sum_{i=1}^{4} p_i U_i \geq 0, \tag{14}$$

$$PC_s^{EA} \quad : \quad EV = \sum_{i=1}^{4} p_i V_i \geq 0, \tag{15}$$

As usual in this type of models, the principal can achieve the first best outcome. This outcome can be implemented as a FO with an important caveat: the agent gets punished harshly whenever he selects the insourcing contract and the auditor subsequently reports that he is efficient. However, the punishment occurs out of equilibrium, so the agent is never actually taxed.[19]

To see why this matters, consider a more traditional approach where the principal offers a single contract mechanism. For example, take a contract fashioned after the one proposed by Faure-Grimaud et al. (2003). The auditor obtains a positive payoff in state 1 (i.e., the signal is informative and the type is efficient), which is taxed away when her signal is uninformative (state 2 or 3). Additionally, the efficient agent obtains an information rent when the auditor's signal is uninformative (state 2), but the rent is taxed away when the signal is informative (state 1). Lacking knowledge about the signal type, the agent and the auditor accept the contract provided that in expectation it leaves them with their outside option. Using this single contract mechanism, the principal can eliminate rents completely and implement the first-best.

There is an important difference between this single contract mechanism and the FO described above: in the former, the agent and auditor are actually taxed on the equilibrium path and therefore receive a negative payoff in some states of the world, whereas in the latter these negative payoffs never occur. This difference has an important implication: expected profits under a single contract mechanism decline in the presence of risk aversion, because the contract leverage uncertainty to deal with the collusion problem, while under FO they do not. To clarify this point, suppose the agent and the auditor are infinitely risk averse. Given that they care only about the worse payoff, any mechanism where a negative payoff is realized in equilibrium would violate their participation constraints. This would ultimately prevent a single contract mechanism from costlessly eliminating the threat of collusion (a point already made by Tirole [1986] and, in a soft information context, by Faure-Grimaud et al. [2003].) On the contrary, under FO the incentive compatibility constraint is given

[19] To see this point more formally, it is sufficient to revisit the proof of proposition 3 and note that, in the presence of ex-ante constraints, the principal can set $\breve{U}_1 = -\infty$, reduce the agent's information rent to zero in all states of the world ($\hat{U}_i = 0$ for all i), and implement first-best efforts ($\psi'(e_i^*) = 1$ for all i.)

by (13), payoffs are always non-negative and, as discussed in section 3.3.2, the contract continues to achieve the first best outcome.

4.2 Ex-ante Collusion and the Role of Commitment

An important assumption made is that the auditor can in no way influence the participation decision of the agent before the contracting stage. In other words, the auditor and the agent cannot collude before the contract is signed. What happens if they could collude ex-ante? It turns out that the answer depends on the commitment abilities of the two colluding parties. If the auditor is able to commit ex ante to a side deal that is ex-post suboptimal, our proposed contract breaks down.[20] If, however, the two actors lack this commitment ability, then the FO is (sometimes) robust to ex-ante collusion.

Before proceeding with the explanation, let us clarify the timing of the game under ex-ante collusion. At stage 0, the agent learns about his type, and every one learns the distribution of β and σ. At stage 1, the principal offers the mechanism. At that point, the agent and the auditor can collude over the selection of the participation stage messages (m_{a1}, m_{s1}). At stage 1.5, the agent and auditor learn the signal σ, and can again collude over the remaining messages (m_{a2}, m_{s2}). At stage 2, the agents reports m_{a2} and the auditor reports m_{s2}. At stage 3, output is produced and payoffs realized.

Full Commitment To begin, consider the case where the agent and the auditor can commit to any ex-ante agreement, even if such deal is ex-post inferior. In this case our FO unravels. The intuition is straightforward and is only sketched here. It is sufficient to note that the auditor could commit ex-ante to always report $m_{s2} = \varnothing$, in exchange for a payment $\frac{1}{1+\lambda}b$ from the agent in state 1. Any bribe $b < \xi\Phi(e_3)$ will make the agent earn higher rents under the (now collusive) insourcing contract than under outsourcing. The payoff to the agent in state 1 is thus $\Phi(e_3) - b > t_0 - \psi(e_0)$. The payoff to the auditor is $\frac{1}{1+\lambda}b > s_0 = 0$. The side contract is ex-ante beneficial to both the agent and the auditor. In order to make the contract ex-ante collusion proof, the principal would have to raise the payoff to the auditor, s_0. It is readily observed that the second best, as we described for the case of interim collusion, cannot be obtained. We leave the analysis of the optimal mechanism in the case of full commitment and ex-ante collusion to future research.

Partial Commitment Now suppose that the auditor and the agent cannot commit to an ex-ante side agreement, i.e., what the auditor and the agent agree in stage 1 can be reneged

[20] See Mookherjee and Tsumagari (2004) for a model with soft information and ex-ante collusion.

during stage 2. Once the agent has decided to accept the insourcing contract, and state 1 has been realized, any bribe that the agent would have agreed in stage 1 to pay is lower than s_1. Thus, the auditor will choose to break the side contract and report $m_{s2} = \underline{\beta}$. Knowing this, the agent will never agree to the side contract.[21] However, it is important to recognize that this result hinges both on the timing of information and on the assumption that the signal σ is public—that is, it is observed by both agent and auditor. For instance, if the signal is private information to the auditor, and if the auditor receives the information at the participation stage, then ex-ante collusion may be possible even in the absence of commitment. In that case our FO can still implement the second best under certain conditions.[22]

Other Considerations on Ex-ante Collusion We have seen that the mechanism presented is robust to ex-ante collusion if commitment is not possible. Commitment power might be difficult to establish in markets with many possible principals, auditors and agents, where the latter often negotiate the contract with the principal, and only later they learn the identity of the auditor. Our theoretical model suggests that this prevents collusion, and there is some empirical evidence to suggest that indeed the random assignment of auditors reduces corruption in regulated markets (Duflo et al, 2012). Forced job rotation, where auditors and workers are continually reassigned, achieves the same result (Felli and Hortala-Vallve, 2011).

4.3 Timing of Information: Signal in Stage 0

So far, we have assumed that the information structure is such that the signal is generated in the interim stage. In this section, we show what happens if the auditor's signal is generated in stage 0 (before participation).

Privately Informed Auditor Suppose the auditor privately learns σ during stage 0 but cannot use this information to collude. (That is, we exclude ex-ante collusion.) With this setup, the relevant binding constraints are unchanged. Thus, the second best and collusion-proof solution matches the discussion in section 2.2 and 2.3. It is straightforward to see that the FO outcome is unchanged, since the outcome hinges on the continuing participation of the auditor (guaranteed by (3)) and on constraint (11) for the agent, which remains unchanged.

[21] In order to avoid the commitment problem, the auditor could post a bond b^{EA} at time 1. Then, in period 2, both agent and auditor would report non cooperatively to the principal; if the signal σ is informative, the bond would be paid to the agent, whereas it would be returned to the auditor if $\sigma = \varnothing$. This workaround relies on the auditor being able to post the bond, i.e., without limited liability, and on the presence of a fully credible third party to handle the bond payment.

[22] Notes available from the authors upon request.

Public Signal Next consider the case where the signal σ is generated in stage 0, and the information is public to both agent and auditor. It is clear that our mechanism now cannot implement the second best outcome. The agent knows whether the signal is informative before choosing between the insourcing and the outsourcing contract. He will thus always choose the insourcing contract when $\sigma = \varnothing$, and the outsourcing contract when $\sigma = \underline{\beta}$. Thus, eliciting information at the participation stage is no longer beneficial to the principal.

4.4 Multiple Types of Agent

In this section we discuss the extension to multiple types of agents. A web appendix offers a rigorous analysis, including all the proofs and a detailed description of the optimal mechanism. Here we just present an informal overview of the main results. Qualitatively, the results from the 2-type generalize to the N-type case: (i) the optimal mechanism is collusion-proof, (ii) it achieves the same profit as the collusion-free benchmark (second-best), (ii) it includes as many contracts as different types of agents, such that each type i chooses a different contract in equilibrium; more complex mechanisms are not profit enhancing, (iv) the agent is fully insured along the equilibrium path.

In what follows, we present our N-types framework and we discuss how the results from the 2-type generalize. The agent's type is now $\beta \in \{\beta_1, \beta_2, ..., \beta_N\}$, where $\beta_i > \beta_j$ for $i > j$, and ν_i denotes the probability that the agent has type β_i, with $\sum_i \nu_i = 1$. The auditor receives one signal which may be informative of the agent's type. Denote by the superscript $w \in \{I, U\}$ the states of the world where the signal is informative (I) or uninformative (U). The signal is informative with probability ξ, and uninformative with probability $1 - \xi$. The auditor's signal belongs to the set $\{\sigma_1, \sigma_2, ..., \sigma_N, \varnothing\}$, where $\sigma_i^I = \beta_i$ for all i, and $\sigma^U = \varnothing$ irrespective of the agent's realized type.

An optimal mechanism specifies N type-specific contracts, such that in equilibrium, each type of agent selects a different contract. Thus, as in the 2-type model, the principal prefers to elicit the agent's type at the participation stage, before the auditor's signal is realized. Perhaps more surprising is that, to optimally elicit that information, the utility of the agent is *always* fully insured: $\hat{U}_i^U = \hat{U}_i^I$ for all i, where $\hat{U}_i^w$ represents the agent's utility when the type is β_i and the state is w.[23] Note that the threat of collusion arises when the agent prefers the auditor to misreport state I as U. Since the agent obtains the same payoff irrespective of the auditor's report, collusion is eliminated at no cost and the auditor's payoff can be set to zero in all states of the world. However, the auditor's signal does play a role: insurance applies to the agent's utility, but not to his effort level since the equilibrium efforts are

[23] Interestingly, this is optimal not only under the threat of collusion, but also in the collusion-free benchmark.

$e_i^I \neq e_i^U$. As in the two-type case, supervision is not necessary for type 1 agent, because $e_1^I = e_1^U$. The auditor's job is also important off the equilibrium path. In our construction, she must be induced to report her signal truthfully whenever the agent selects the contract that was not designed for him. For further details on how the N-type mechanism compares with the 2-type one, see the web appendix.

A final point is worth noting. The 2-type FO had a feature that added to its robustness: In the (out-of-equilibrium) event that the efficient agent selected the insourcing contract and the auditor's signal was uninformative, the agent was still incentivized to truthfully reveal his type and to exert optimal effort. This could be an important feature if one is concerned with the possibility of random realization of out of equilibrium events, which could occur if the agent makes a mistake in evaluating the probability of a successful audit.[24] Being a problem inherent to our FO, it is potentially important for our mechanism to address this aspect. Implementing this feature in the N-types framework is no longer costless. In the web appendix we show a particular version of the optimal contract with out of equilibrium payoffs that alleviates this problem. As an example, take contract Γ_i, which is selected in equilibrium by type β_i. After the contract has been selected, the principal can allow all types $j < i$ to report that the realized cost is indeed C_j. However, in order to make this costless to the principal, reporting a cost C_j is allowed with a certain probability τ. With probability $1 - \tau$, the principal pays the agent a fixed fee $C_i = \beta_i - e_i^U$. The web appendix also shows that allowing all types $j > i$ to report their costs is always costly to the principal, no matter how small τ is. While not central to the main argument of this paper, the refinement retains an interesting feature of the FO outcome: namely, the selection of a contract is akin choosing the amount of residual claims that are retained by the agent. With a certain probability, the principal is the residual claimant for costs, up to the maximum cost indicated by the agent.

5 Conclusion

This paper shows that the threat of collusion alone can influence a number of organizational dimensions. To show this point we considered a simple modification of Laffont's and Tirole's (1991) model and mechanism in hierarchical structures where an agent and his auditor can collude at the expense of the principal. By letting the agent choose over a menu of contracts or organizational structures, our model yields second-best results. In particular, we let the agent choose between an *outsourcing contract* that requires no monitoring, and an *insourcing contract* subject to auditing. The optimal contract has well-defined characteristics. First,

[24] This possibility does not affect the LT mechanism. In LT the agent makes his report at the stage where the auditor's signal is already realized and common knowledge.

it involves highly productive workers choosing to produce outside of the firm, with no monitoring, while inefficient workers remain inside the firm. Second, when the principal/firm outsources production, the agent is the residual claimant of production costs, while when production remains within the firm, the principal is. Third, it provides insurance both to the agent and the auditor.

Our mechanism provides a novel explanation for the practice of outsourcing of production that is based on the presence of collusion *within* the firm, and delivers several insights on the importance of commonly-used assumptions in principal-agent models of corruption. In particular, the model makes clear that the timing of collusion and information matters in determining the optimal contract. Our model delivers second-best outcome when the agent and the auditor do not collude in the participation decision; however, when collusion takes place before they agree to the contract, our model is sensitive to the ability to commit of the auditor. Finally, we find that the model is robust to changes in the timing of information available to the auditor, and can be extended to multiple types of agent.

6 Proofs

6.1 Proof of Proposition 1

6.1.1 Single Contract

Let us first focus on the class of single contract mechanisms when there is no collusion (the auditor cannot misreport or lie about the signal to the principal). For any arbitrarily chosen single contract Γ, the mechanism offered by the principal determines some *equilibrium allocation*, one for each of the four possible ex-post states of the world (β, σ) described in Section 2.1: $(\underline{\beta}, \underline{\beta})$, $(\underline{\beta}, \varnothing)$, $(\overline{\beta}, \varnothing)$ and $(\overline{\beta}, \overline{\beta})$. Let us index this equilibrium allocation by a hat: $\{\hat{t}_i, \hat{s}_i, \hat{U}_i, \hat{V}_i\}_{i=1}^4$, where we follow LT and focus on deterministic final allocations. The same reasoning can be extended to random final allocations. The actual transfers from the principal to the auditor and the agent are denoted respectively by s_i and t_i in state i. Hence,

$$\hat{s}_i = s_i \tag{16}$$

$$\hat{t}_i = t_i \tag{17}$$

$$\hat{U}_i = \hat{t}_i - \psi(e_i) \tag{18}$$

$$\hat{V}_i = \hat{s}_i. \tag{19}$$

Having this schedule in place, we first derive a couple of necessary conditions that must be satisfied by the final allocation in any equilibrium. Then, we derive an upper bound on expected profit by (i) writing the expected profit as a function of the final allocation and (ii) maximizing the expected profit subject to the limited set of necessary conditions we derived. Finally, we show that the upper bound can be reached by a direct revelation mechanism. Recall that our definition of single contract implies $M_{a1} = \{Participate, Exit\}$. Participation of both the auditor and the agent is required for production by assumption. Thus, the set of constraints that need to be satisfied for all i is:

$$\hat{V}_i \geq 0 \tag{20}$$

$$\hat{U}_i \geq 0 \tag{21}$$

$$E(\hat{U}|\underline{\beta}) = \frac{p_1}{p_1 + p_2}\hat{U}_1 + \frac{p_2}{p_1 + p_2}\hat{U}_2 \geq 0 \tag{22}$$

$$E(\hat{U}|\overline{\beta}) = \frac{p_3}{p_3 + p_4}\hat{U}_3 + \frac{p_4}{p_3 + p_4}\hat{U}_4 \geq 0 \tag{23}$$

$$E(\hat{V}) = \sum_i p_i \hat{V}_i \geq 0. \tag{24}$$

The first two inequalities ensure that the final allocations satisfy the ex-post participation constraints. The latter three inequalities are ex ex-anteante participation constraints that need to be satisfied before either the auditor or the agent learn about each other's information. It is clear that

constraints (22)-(24) are always satisfied when (20) and (21) are satisfied—so they can be ignored.

In States 1 and 4 the principal learns the agent's type directly (and costlessly) from the auditor, so misreporting is not a concern. However, since in state 2 only the agent knows that $\beta = \underline{\beta}$, he can mimic the behavior of type $\overline{\beta}$ and get utility $\hat{U}_3 + \Phi(e_3)$, where $\Phi(e_3) = \psi(e_3) - \psi(e_3 - \triangle\beta) > 0$. Thus, the equilibrium outcome must satisfy the following condition:

$$\hat{U}_2 \geq \hat{U}_3 + \Phi(e_3) \tag{25}$$

Note that we are not restricting the message space M_{a2} available to the agent. For example, the underlying mechanism that determines the equilibrium allocations $\{\hat{t}_i, \hat{s}_i, \hat{U}_i, \hat{V}_i\}_{i=1}^4$ could instruct the agent to report not only β but also σ. Here we remain agnostic.
The expected profit function is

$$W = G - \sum_{i=1}^{4} p_i(t_i + s_i + \beta_i - e_i). \tag{26}$$

Use (16), (17), (18), and rearrange the profit function to obtain,

$$W = G - \sum_{i=1}^{4} p_i(\hat{s}_i + \beta_i - e_i + \hat{U}_i + \psi(e_i)). \tag{27}$$

Next, we find the upper bound $W^{\max}$ for W when the constraints (20), (21), and (25) are imposed on the decision variables $\{e_i, \hat{s}_i, \hat{U}_i\}_{i=1}^4$. That is, we ignore other potential necessary conditions for the moment. Owing to the fact that rents are costly, optimality requires

$$\hat{s}_i = 0 \qquad \text{for } i = 1, 2, 3, 4, \tag{28}$$

$$\hat{U}_1 = \hat{U}_3 = \hat{U}_4 = 0. \tag{29}$$

Moreover, (25) is satisfied with equality for the same reason, thus $\hat{U}_2 = \Phi(e_3)$. The maximization with respect to e is as announced in the main text. Effort is socially optimal, i.e., $\psi'(e_i^*) = 1$, except in state 3, when, using (25), the optimal effort solves

$$\nu\Phi'(e_3) + (1-\nu)\psi'(e_3) = (1-\nu)$$

To complete the proof we have left to show that the upper bound can be reached using a direct revelation mechanism where the principal costlessly learns σ from the auditor and the agent reports m_{a2}, where $M_{a2} = \{\underline{\beta}, \overline{\beta}\}$. Let $i = 1$ denote the state in which $\sigma = \underline{\beta}$, and $m_{a2} = \underline{\beta}$ etc. Now

consider the following transfers and cost targets

$$s_i = \hat{s}_i \tag{30}$$

$$t_i = \hat{U}_i + \psi(e_i) \tag{31}$$

$$C_i = \beta_i - e_i \tag{32}$$

where $\beta_i = \underline{\beta}$ for $i = \{1, 2\}$ and $\beta_i = \overline{\beta}$ for $i = \{3, 4\}$ and $\{e_i, \hat{s}_i, \hat{U}_i\}_{i=1}^4$ are the solutions to the maximization of (26) derived above. Clearly, in no state of nature do the agent have an incentive to individually misreport or lie.[25] Thus, the upper bound can be reached by a direct revelation mechanism.

6.1.2 General Contract

In the general mechanism, the agent can use the message space M_{a1} to communicate more than just his participation decision. At the participation stage, the agent knows his type (i.e., he can distinguish between states 1-2 and 3-4). Hence, the agent can use the information about β in choosing his message, i.e., $m_{a1}(\beta)$. Let us denote by $m^*_{a1}(\underline{\beta})$ and $m^*_{a1}(\overline{\beta})$ the equilibrium messages induced by the mechanism in states 1-2 and 3-4 respectively. Unlike in the single contract case, the principal can design a mechanism such that the agent reveals his information at the interim stage, when the auditor's signal is still unrealized. One necessary restriction on the equilibrium is generated by the fact that the efficient agent $\underline{\beta}$ can misreport his type by selecting $m^*_{a1}(\overline{\beta})$. The $\underline{\beta}$-agent expects to obtain $\frac{p_1}{p_1+p_2}\hat{U}_1 + \frac{p_2}{p_1+p_2}\hat{U}_2$ by reporting his type. On the other hand, he expects to obtain $\frac{p_1}{p_1+p_2}\breve{U}_1 + \frac{p_2}{p_1+p_2}\left[\hat{U}_3 + \Phi(e_3)\right]$ by misreporting, where $\breve{U}_1$ denotes the (out-of-equilibrium) payoff in the case where the efficient agent reports $m^*_{a1}(\overline{\beta})$ and the principal learns that the true type is $\underline{\beta}$. This entails that the interim incentive compatibility constraint for the efficient agent is

$$\frac{p_1}{p_1+p_2}\hat{U}_1 + \frac{p_2}{p_1+p_2}\hat{U}_2 \geq \frac{p_1}{p_1+p_2}\breve{U}_1 + \frac{p_2}{p_1+p_2}\left[\hat{U}_3 + \Phi(e_3)\right]. \tag{33}$$

Having this new necessary condition in place, we now establish an upper bound on expected profit—we ignore other potential necessary conditions for the moment. The set of binding constraints is then (20), (21), and (33). Given that rents are costly, the principal wants to set $\breve{U}_1$ as low as possible to discourage the agent from misreporting. Because the agent has the option to quit before applying effort in the state of the world where his misreport is uncovered by the principal, the lowest payoff is $\breve{U}_1 = 0$. Does the principal prefer to design a mechanism where screening occurs at the interim stage (i.e., [33] needs to be satisfied) or a mechanism where screening occurs ex-post (i.e., [25] is required)? It is easy to the see that the principal is indifferent. Because rents are costly, (33) must be binding. After substituting $\breve{U}_1 = 0$ and rearranging, the binding (33) reduces

[25] Note that the second incentive compatibility constraint that ensures that the inefficient agent never mimics the efficient one (state 3) is never binding in equilibrium.

to $\hat{U}_2 = \left[\hat{U}_3 + \Phi(e_3)\right] - \frac{p_1}{p_2}\hat{U}_1$. Substitute into (27) and note that $\hat{U}_1$ cancels out. The upper bound is the same as the one for the single contract case. It is then without loss of generality to focus on direct revelation mechanism with single contract.

6.2 Proof of proposition 2

Here we show that the mechanism presented in section 2.3 results in the third best outcome. This proof is similar to the one provided in the original LT paper, under our new timing. Strictly speaking, proving our claim would just require showing that the binding constraints match the ones found in the LT paper. For the sake of completeness, we will also closely follow their proof to derive final outcomes and demonstrate that (a) the auditor reports σ truthfully; (b) transfers are based on $(C(m_{a2}), \sigma)$ where $M_{a2} = \{\overline{\beta}, \underline{\beta}\}$ and more complex mechanisms are not profit enhancing; (c) the auditor's salary depends only on her report; (d) there is no bribe-exchange in equilibrium. The proof strategy in LT consists of the following steps: (1) establish an upper bound on expected profit by applying a few necessary conditions, and then maximizing the constrained profit. As in the LT case, we will find that the optimal bribes are zero. (2) We show that this upper bound can be reached by an incentive scheme that is collusion proof and that satisfies properties (a)—(d).

Let us start with some notation. For any arbitrarily chosen single contract Γ, the mechanism offered by the principal determines a side contract between the agent and the auditor, and some *equilibrium allocation*. Let us index this equilibrium allocation by a hat: $\{\hat{t}_i, \hat{s}_i, \hat{U}_i, \hat{V}_i\}_{i=1}^4$. This equilibrium allocation might include the equilibrium bribes, if any. The actual transfers from the principal to the auditor and the agent are denoted respectively by s_i and t_i in state i. Let us denote by $\tilde{s}_i$ the agent's bribe to the auditor, which is assumed to be positive without loss of generality.[26] Hence,

$$\hat{s}_i = s_i + \tilde{s}_i \tag{A.1}$$

$$\hat{t}_i = t_i - (1+\lambda)\tilde{s}_i \tag{A.2}$$

$$\tilde{s}_i \geq 0 \tag{A.3}$$

$$\hat{U}_i = \hat{t}_i - \psi(e_i) \tag{A.4}$$

$$\hat{V}_i = \hat{s}_i. \tag{A.5}$$

Here, we follow LT and focus on deterministic final allocations. The same reasoning can be extended to random final allocations. The restrictions (i)-(iv) fleshed out in Section 2.1 imply that the set

[26] It is possible to show that negative bribes are always suboptimal.

of constraints that need to be satisfied for all i is:

$$\hat{V}_i \geq 0 \tag{A.6}$$

$$\hat{U}_i \geq 0. \tag{A.7}$$

$$E(\hat{U}|\underline{\beta}) = \frac{p_1}{p_1 + p_2}\hat{U}_1 + \frac{p_2}{p_1 + p_2}\hat{U}_2 \geq 0 \tag{A.8}$$

$$E(\hat{U}|\overline{\beta}) = \frac{p_3}{p_3 + p_4}\hat{U}_3 + \frac{p_4}{p_3 + p_4}\hat{U}_4 \geq 0 \tag{A.9}$$

$$E(\hat{V}) = \sum_i p_i \hat{V}_i \geq 0. \tag{A.10}$$

The first two inequalities ensure that the final allocations satisfy the ex-post participation constraints. The latter three inequalities are ex-ante participation constraints that need to be satisfied before either the auditor or the agent learn about each other's information. It is clear that constraints (A.8)-(A.10) are always satisfied when (A.6) and (A.7) are satisfied—so they can be ignored.

Since in state 2 only the agent knows that $\beta = \underline{\beta}$, he can mimic the behavior of type $\overline{\beta}$ and get utility $\hat{U}_3 + \Phi(e_3)$. Thus, the final allocation must meet the following condition:

$$\hat{U}_2 \geq \hat{U}_3 + \Phi(e_3) \tag{A.11}$$

Finally, the auditor and the agent can always agree on a side-contract in state of nature 1 that would suggest that the true state of nature is 2. This imposes the following condition on their equilibrium allocations:

$$(1 + \lambda)(\hat{s}_1 - \hat{s}_2) \geq \hat{U}_2 - \hat{U}_1. \tag{A.12}$$

Note that we are not restricting the message space available to the agent and the auditor at stages 1 and 2. The agent could be asked to report any kind of message, including β and σ. And the auditor could be asked to report not only her hard information but also any kind of unverifiable message. Irrespective of the message space, (A.6)-(A.7) and (A.11)-(A12) must be satisfied.

Thus, the set of binding constraints is (A.6)-(A.7) and (A.11)-(A12), which is the same set identified by LT. The LT allocation is thus obtained. This proves our statement. However, for completeness, we provide a very close replication of the LT proof.
The expected profit function is

$$W = G - \sum_{i=1}^{4} p_i(t_i + s_i + \beta_i - e_i). \tag{A.13}$$

Use (A.1), (A.2), (A.4), and rearrange the profit function to obtain,

$$W = G - \sum_{i=1}^{4} p_i(\lambda \tilde{s}_i + \hat{s}_i + \beta_i - e_i + \hat{U}_i + \psi(e_i)). \tag{A.14}$$

Next, we find the upper bound $W^{\max}$ for W when the constraints (A.3), (A.6), (A.7), (A.11), and (A.12) are imposed on the decision variables $\{e_i, \tilde{s}_i, \hat{s}_i, \hat{U}_i\}_{i=1}^4$. That is, we ignore other potential constraints for the moment. Owing to the fact that rents are costly, optimality requires,

$$\hat{s}_i = 0 \qquad \text{for } i = 2, 3, 4, \tag{A.15}$$

$$\hat{U}_3 = \hat{U}_4 = 0. \tag{A.16}$$

Moreover, (A.11) and (A.12) are satisfied with equality for the same reason. An inspection of (A.14) also reveals that

$$\tilde{s}_i = 0 \qquad \text{for all } i, \tag{A.18}$$

meaning that the final allocation does not include a side payment. To show that $\hat{U}_1 = 0$, it suffices to note that from (A.11) and (A.12) follows

$$(1 + \lambda)\hat{s}_1 = \Phi(e_3) - \hat{U}_1.$$

Thus, maximizing W with respect to $\hat{U}_1$ is equivalent to maximizing $-(\frac{\Phi(e_3)-\hat{U}_1}{(1+\lambda)} + \hat{U}_1)$ subject to $\hat{U}_1 \geq 0$. Thus, $\hat{U}_1 = 0$. The maximization with respect to e is as announced in the main text. Effort is socially optimal, i.e., $\psi'(e_i^*) = 1$, except in state 3, when, using (A.11) and (A.12), the optimal effort solves

$$\Phi'(e_3)\left[\frac{\xi\nu}{1+\lambda} + \nu(1-\xi)\right] + (1-\nu)(1-\xi)\psi'(e_3) = (1-\nu)(1-\xi)$$

To complete the proof we have left to show that the upper bound can be reached. To this purpose, take a direct revelation mechanism where the auditor reports his signal, $m_{s2} = \sigma$, and the agent reports his type, $m_{a2} = \beta$. Let $i = 1$ denote the state in which $m_{s2} = \underline{\beta}$, and $m_{a2} = \underline{\beta}$ etc. Now consider the following transfers and cost targets

$$s_i = \hat{s}_i \tag{A.19}$$

$$t_i = \hat{U}_i + \psi(e_i) \tag{A.20}$$

$$C_i = \beta_i - e_i \tag{A.21}$$

where $\beta_i = \underline{\beta}$ for $i = \{1, 2\}$ and $\beta_i = \overline{\beta}$ for $i = \{3, 4\}$ and $\{e_i, \hat{s}_i, \hat{U}_i\}_{i=1}^4$ are the solutions to the maximization of (A.13) derived above. Clearly, in no state of nature do the auditor and the agent have an incentive to collude against this scheme, or to individually misreport or lie. Thus, the

upper bound can be reached by a pair of contracts that satisfy (a)—(d), as claimed above.

6.3 Proof of proposition 3

We use the same proof strategy used in the proof of proposition 2: (1) we establish an upper bound on expected profit by deriving a couple of necessary conditions that must be satisfied by the final allocation in any equilibrium, and (2) we maximize the constrained profit. We find in particular that optimal bribes are equal to zero, as in the LT single contract case.

Under the general contract, the participation decision is no longer a binary decision, i.e., *Participate* or *Exit*. The principal can offer a menu of contracts / mechanisms. Participation decisions become richer and can now be used as a screening device by the principal. Thus, the *equilibrium allocation* should also specify the equilibrium contract selected by the agent at the participation stage, i.e., $\{\hat{t}_i, \hat{s}_i, \hat{U}_i, \hat{V}_i, \hat{\Gamma}_i\}_{i=1}^4$, where $\hat{\Gamma}_i$ represents the equilibrium message $m^*_{a1}(\beta_i)$ selected by the agent in state i with $\beta_i = \underline{\beta}$ for $i = \{1, 2\}$ and $\beta_i = \overline{\beta}$ for $i = \{3, 4\}$. Clearly, $\hat{\Gamma}_i$ need not be the same in all states of the world. However, since the agent cannot distinguish between states 1 and 2 at the participation stage, the equilibrium contract selected by the agent must be the same in these two states of the world. The same applies to states 3 and 4. Let us denote by $\hat{\Gamma}_{\underline{\beta}}$, and $\hat{\Gamma}_{\overline{\beta}}$, the equilibrium contracts in states 1-2 and 3-4 respectively.

It is easy to see that the necessary conditions (A.6) and (A.7) must still be satisfied by the final allocation in any equilibrium. (A.12) also remains a necessary condition because, after the participation decisions are made, the agent and the auditor know the state i. If (A.12) were violated, in state 1 the agent would select $\hat{\Gamma}_{\underline{\beta}}$ and the coalition auditor-agent would be better off colluding and reporting that the state is 2.

But is (A.11) also necessary? At the participation stage, the agent knows his type, so he can distinguish between states 1-2 and 3-4. Hence, the equilibrium contract need not be the same in states 1-2 and 3-4. Suppose it were instead the same. Then (A.11) would still be a necessary condition because, in state 2, the agent is the only one to know that $\beta = \underline{\beta}$, and so he can mimic the behavior of type $\overline{\beta}$ and get utility $\hat{U}_3 + \Phi(e_3)$. However, the principal can design a mechanism such that, in equilibrium, the agent selects a different contract in states 1-2 and 3-4. In that case, condition (A.11) is no longer required, and a new necessary condition emerges, which is effectively an interim incentive compatibility constraint.

To see the point formally, note that the agent has to make his participation decision without knowing the auditor's signal. One necessary condition is then the following: the principal has to ensure that the agent selects the contract $\hat{\Gamma}_{\underline{\beta}}$ when the type is $\underline{\beta}$. The $\underline{\beta}$-agent expects to obtain $\frac{p_1}{p_1+p_2}\hat{U}_1 + \frac{p_2}{p_1+p_2}\hat{U}_2$ by selecting $\hat{\Gamma}_{\underline{\beta}}$. On the other hand, he expects to obtain $\frac{p_1}{p_1+p_2}\breve{U}_1 + \frac{p_2}{p_1+p_2}\left[\hat{U}_3 + \Phi(e_3)\right]$ by selecting $\hat{\Gamma}_{\overline{\beta}}$, where $\breve{U}_1$ denotes the (out-of-equilibrium) payoff in the case where the agent selects $\hat{\Gamma}_{\overline{\beta}}$ and the auditor learns that the type is $\underline{\beta}$. This entails that the new

necessary condition is

$$\frac{p_1}{p_1+p_2}\hat{U}_1 + \frac{p_2}{p_1+p_2}\hat{U}_2 \geq \frac{p_1}{p_1+p_2}\breve{U}_1 + \frac{p_2}{p_1+p_2}\left[\hat{U}_3 + \Phi(e_3)\right]. \tag{A.11b}$$

We will show later in the proof that the relevant necessary condition is (A.11b) and not (A.11). For the time being, we remain agnostic.[27]

In the next part of the proof we establish an upper bound on expected profit by applying the necessary conditions, and then maximizing the constrained profit. The expected profit function is

$$W = G - \sum_{i=1}^{4} p_i(t_i + s_i + \beta_i - e_i). \tag{A.22}$$

Use (A.1), (A.2), (A.4), and rearrange the profit function to obtain,

$$W = G - \sum_{i=1}^{4} p_i(\lambda\tilde{s}_i + \hat{s}_i + \beta_i - e_i + \hat{U}_i + \psi(e_i)). \tag{A.23}$$

Does the principal prefer to design a mechanism where screening occurs at the participation stage (i.e., [A.11b] is the necessary condition) or a mechanism where all types select the same contract (i.e., [A.11] is the necessary condition)? Let us denote the two kind of mechanisms by (a) and (b) respectively. It is easy to the see that the principal prefers the former. To see this point note that—owing to the agent's ability to exit before applying effort—the principal wants to set $\breve{U}_1$ as low as possible compatibly with the ex-post participation constraints, thus $\breve{U}_1 = 0$.[28] Now we prove our point by contradiction. Suppose instead that (A.11) is the necessary condition, namely, the preferred mechanism is (b).The proof of proposition 2 then implies that the optimal allocation must specify $\hat{U}_1 = 0$. Note that when $\hat{U}_1 = 0$ equation (A.11b) reduces to (A.11). Thus, the allocation that the principal implements in (b) can also be implemented in (a). However, the opposite is not true, and (as we show next) optimality requires $\hat{U}_1 \neq 0$.

Next, we find the upper bound $W^{\max}$ for W when the constraints (A.3) and (A.6), (A.7), (A.11b), and (A.12) are imposed on the decision variables $\{e_i, \tilde{s}_i, \hat{s}_i, \hat{U}_i\}_{i=1}^4$. That is, we ignore other potential constraints for the moment. Owing to the fact that rents are costly, optimality

[27] We are also not imposing any restriction on the message space available to the agent and the auditor at stages 1 and 2. The agent and the auditor could be asked to report any kind of message (including soft and hard information). Irrespective of the message space, the conditions presented above are necessary. Finally, note that the auditor could use the message m_{a1} sent by the agent at stage 1 to update her beliefs about the agent's type. This might affect condition (A.12). However, in state 1 the auditor knows the agent's type. Thus, the auditor does not use the message m_{a1} to update her beliefs and (A.12) is unaffected by this possibility.

[28] Clearly, in the (out-of-equilibrium) case where the agent selects $\hat{\Gamma}_{\overline{\beta}}$ and the auditor reports that the type is $\underline{\beta}$, the auditor should also be suitably incentivized to report her signal. Because this occurs out-of-equilibrium, incentivizing the auditor is costless for the principal.

requires,

$$\hat{s}_i = 0 \qquad \text{for } i = 2, 3, 4, \tag{A.25}$$

$$\hat{U}_3 = \hat{U}_4 = 0. \tag{A.26}$$

Maximization of (A.23) also reveals that

$$\tilde{s}_i = 0 \qquad \text{for all } i, \tag{A.27}$$

meaning that there are no side payments. Moreover, (A.11b) and (A.12) are satisfied with equality for the same reason (i.e., rents are costly.) Note that from (A.12) follows

$$\hat{s}_1 = \frac{\hat{U}_2 - \hat{U}_1}{(1+\lambda)}. \tag{A.28}$$

Hence, $\hat{U}_2 = \hat{U}_1$ minimizes $\hat{s}_1$. Moreover, for any pair $(\tilde{U}_2, \tilde{U}_1)$ such that $\tilde{U}_1 \neq \tilde{U}_2$ and (A.11b) is satisfied with equality, there is another pair such that (A.11b) is still satisfied with equality but $\hat{U}_2 = \hat{U}_1$. Thus, optimality requires that $\hat{U}_2 = \hat{U}_1$. From (A.11b)

$$(\hat{U}_2 =)\hat{U}_1 = \frac{p_2}{p_1 + p_2}\Phi(e_3), \tag{A.29}$$

and,

$$\hat{s}_1 = 0. \tag{A.30}$$

Plugging (A.25), (A.26), (A.27), (A.29) and (A.30) into the function (A.23), we get an unconstrained profit function that depends only on effort levels $e_1, ..., e_4$:

$$\max_{\{e_i\}_{i=1}^4} W^{\max} = G - \sum_{i=1}^{4} p_i(\beta_i - e_i + \psi(e_i)) + p_2\Phi(e_3).$$

The first order conditions of this function are such that effort is socially optimal in all states, i.e., $\psi'(e_i^*) = 1$, except 3, where the optimal effort solves

$$\upsilon\Phi'(e_3) + (1-\nu)\psi'(e_3) = (1-\nu).$$

Note that this first order condition is the same as (6), so $e_3 = e_3^{CF}$. The maximized profit function takes the form:

$$W^{\max} = G - \sum_{i=1,2,4} p_i \left\{\psi(e_i^*) + \beta - e_i^*\right\} - p_3 \left\{\psi(e_3^{CF}) + \overline{\beta} - e_3\right\} - \nu(1-\xi)\Phi(e_3) = W_{CF}^*.$$

It is thus proved that the upper bound on expected profits under the threat of collusion is equal to

the optimal profit in the collusion-free problem, equation (7).

References

[1] Aghion, Philippe, and Richard Holden (2011): "Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?," Journal of Economic Perspectives, 25(2): 181 97.

[2] Baron, David and David Besanko (1984): "Regulation, Asymmetric Information, and Auditing." The RAND Journal of Economics 15(4), 447-470.

[3] Celik, G. (2009). "Mechanism Design with Collusive Supervision," Journal of Economic Theory, 144 (1), 69-95.

[4] Faure-Grimaud, Antoine, Jean-Jacques Laffont and David Martimort (2003): "Collusion, Delegation and Supervision with Soft Information," Review of Economic Studies 70, 253-275

[5] Felli, Leonardo and Rafael Hortala-Vallve (2011): "Preventing Collusion through Discretion," CEPR Discussion Paper No. DP8302.

[6] Gibbons, Robert (2005): "Four Formal(izable) Theories of the Firm?" Journal of Economic Behavior and Organization, 58, 200-245.

[7] Grossman, Sanford J., and Oliver D. Hart. (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," Journal of Political Economy, 94(4): 691-719.

[8] Holmstrom, Bengt and John Roberts (1998): "The Boundaries of the Firm Revisited," The Journal of Economic Perspectives, 12(4), 73-94.

[9] Holmstrom, Bengt and Paul Milgrom (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," Journal of Law, Economics, and Organization 7:24-52.

[10] Holmstrom, Bengt and Paul Milgrom (1994): "The Firm as an Incentive System," American Economic Review 84: 972-91.

[11] Jensen, Michael C., and William H. Meckling (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," Journal of Financial Economics, 3(4): 305-360.

[12] Khalil, Fahad and Jacques Lawarrée (2006): "Incentives for Corruptible Auditors in the Absence of Commitment," The Journal of Industrial Economics, 54(2), 269-291.

[13] Khalil, Fahad, Jacques Lawarrée and Troy J. Scott (2011): "Private Monitoring, Collusion and the Timing of Information," University of Washington Working Paper.

[14] Laffont, Jean-Jacques and David Martimort (2002): The Theory of Incentives: The Principal-Agent Model, Princeton, NJ: Princeton University Press.

[15] Laffont, Jean-Jacques and Jean Tirole (1991): "The Politics of Government Decision-Making: A Theory of Regulatory Capture", The Quarterly Journal of Economics, Vol. 106, No. 4 (Nov., 1991), pp. 1089-1127

[16] Laffont, Jean-Jacques and Jean Tirole (1993): A Theory of Incentives in Procurement and Delegation (Cambridge, MA: MIT Press).

[17] Lambert-Mogilianksy, Ariane (1998): "On Optimality of Illegal Collusion in Contracts," Review of Economic Design, 3, 303-348

[18] McAfee, P and McMillan, J (1995): "Organizational Diseconomies of Scope, "Journal of Economics and Management Strategy, 4, 399-426

[19] Mookherjee, Dilip and Masatoshi Tsumagari (2004): "The Organization of Supplier Networks: Effects of Delegation and Intermediation," Econometrica, 72(4), 1179-1219

[20] Motta, Alberto (2012): "Collusion and Selective Supervision," SSRN Working Paper.

[21] Myerson, R. (1986): "Multistage Games with Communication," Econometrica, 54(2), 323-358

[22] Simon, H. (1951):"A formal theory of the employment relationship," Econometrica, 19, pp. 293-305

[23] Tirole, Jean (1986): "Hierarchies and Bureaucracies," Journal of Law, Economics, and Organization, 2, 181-214

[24] Williamson, Oliver (1971): "The Vertical Integration of Production: Market Failure Considerations," American Economic Review, 61(2), 112-123.