

RENEGOTIATION AND COLLUSION IN ORGANIZATIONS

LEONARDO FELLI

*London School of Economics
Houghton Street, London WC2A 2AE, UK
lfelli@econ.lse.ac.uk*

J. MIGUEL VILLAS-BOAS

*University of California, Berkeley
Haas School of Business, Berkeley, CA 94720
villas@haas.berkeley.edu*

It has been argued that collusion among the members of an organization may lead to inefficiencies and hence should be prevented in equilibrium. This paper shows that whenever the parties to an organization can renegotiate their incentive scheme after collusion, these inefficiencies can be greatly reduced. Moreover, it might not be possible to prevent collusion and renegotiation in equilibrium. Indeed, if collusion is observable but not verifiable, then the organization's optimal incentive scheme will always be renegotiated. If, instead, collusion is not observable to the principal, both collusion and renegotiation will occur in equilibrium with positive probability. The occurrence of collusion and renegotiation should therefore not be taken as evidence of the inefficiency of an organization.

1. INTRODUCTION

1.1 OVERVIEW

The mechanism design literature has provided us with two basic principles that govern the specification of an optimal mechanism to regulate the structure of an organization or a firm: the *renegotiation-proof*

This is a revised version of a working paper entitled "Friendships in Vertical Relations" (Felli and Villas-Boas, 1996). We are grateful to two anonymous referees, a coeditor of the journal, Luca Anderlini, Francesca Cornelli, Drew Fudenberg, John Groenewegen, Oliver Hart, Thomas Mariotti, Steven Matthews, Nicola Persico, Sönje Reiche, Mark Shankerman, Lars Stole, Jean Tirole, Steve Umlauf, Birger Wernerfelt, Jeff Zwiebel, and seminar participants at the Universidade Católica Portuguesa for their useful discussions and comments. Errors remain our own responsibility. A revision of this paper was prepared while the first author was visiting the Department of Economics of the University of Pennsylvania. Their generous hospitality is gratefully acknowledged.

principle (Dewatripont, 1989; Maskin and Tirole, 1992; Dewatripont and Maskin, 1995) and the *collusion-proof principle* (Tirole, 1986, 1992; Laffont and Martimort, 2000). These two principles guarantee that whenever the members of an organization have the opportunity to collude against or renegotiate the optimal incentive scheme that governs the working of the organization, in equilibrium, both collusion and renegotiation can be prevented by the optimal incentive scheme. Indeed, this can be done by taking explicitly into account the parties' opportunities to get involved in collusion or renegotiation and offering the parties the net payoff they would derive from the renegotiation and collusion transactions.

In this paper, we show that in a dynamic setting, if the members of an organization have the opportunity to both collude against the optimal incentive scheme and then renegotiate it, it is not possible in equilibrium to prevent both collusion and renegotiation. This is true provided that the principal of this organization cannot offer optimal incentive schemes that are contingent on the employees' collusive agreements.

In particular we show that if collusion is observable, but not verifiable, to everybody in the organization, the optimal incentive scheme *might* be collusion-proof but *cannot* be renegotiation-proof. On the other hand, if collusion is not observable to any party but the colluding ones, the optimal incentive scheme cannot be either collusion-proof or renegotiation-proof.

Finally, we show that while the presence of collusion is harmful to the surplus of the organization, if following collusion the parties to the organization have the opportunity to renegotiate their incentive scheme, then these inefficiencies can be greatly reduced. In particular, if collusion is observable, this inefficiency can be completely eliminated, whereas some inefficiencies remain if collusion is only observable to the colluding parties.

Consider an organization in which at least two employees (an agent and a supervisor) operate under the same center (principal). If the principal wants to induce these employees to complete separate tasks using an incentive scheme, he will accomplish this by introducing some risk in their remuneration schedule. However, since the two employees do not necessarily have to perform perfectly correlated tasks, the risks introduced in the two remuneration schedules may differ. Hence, the employees' shares of surplus will, in general, differ in different states of nature. If both employees operate in the same working environment, and at least one of them is risk-averse, they will then have an incentive to get involved in a collusion contract, which takes the form of a risk-sharing agreement. In this way, in fact, they can share the risk the principal is imposing on them. This agreement

alters their incentives, introducing some potential efficiency losses in the organization.

If, in addition, the productive activity of the organization takes time and (say) the supervisor completes her task before the agent completes his own, it is now the principal's turn to have an incentive to renegotiate the wage schedule of the supervisor so as to share with her the risk that the original contract and the collusive agreement with the agent together impose on the supervisor's remuneration. Foreseeing this renegotiation, agent and supervisor might readjust their collusive risk-sharing agreement, introducing additional efficiency losses in the organization.

If, however, the principal is aware of both the collusion and the renegotiation opportunities of the agent and the supervisor, these efficiency losses may be reduced. Indeed, we show that if collusion is observable and renegotiation occurs after collusion, the principal can take into account the redistribution of surplus that these collusion and renegotiation agreements yield when designing the optimal incentive scheme. The result is that, when all transfers are realized, the share of surplus each individual is left with is exactly the one that induces optimal incentives. Hence, in this way, all efficiency losses are eliminated and the only role of collusion and renegotiation is to decentralize the allocation of the optimal shares of surplus to the members of the organization.

Clearly, if one of the individuals operating in the organization is risk-neutral, the way to induce optimal incentives is simple. It will be enough for the principal to make this member of the organization the residual claimant of the other member, in effect selling the organization to him/her. Neither collusion nor renegotiation will be observed in equilibrium.

This paper generalizes the result to the case in which both members of the organization are strictly risk-averse, so that making one of them residual claimant may become very expensive for the principal. As mentioned above, we show in this case that if the principal cannot offer to the supervisor and the agent contracts that are contingent on the allocation induced by their collusive agreement, the optimal incentive scheme yields an allocation of surplus between the employees that might not be colluded upon but will be renegotiated in equilibrium. In other words, it is not possible to have an optimal incentive scheme that is both collusion- and renegotiation-proof.

The main intuition for this result can be described as follows.

Consider the following timing. The principal first offers an incentive scheme to both the agent and the supervisor. Then the supervisor and agent collude by signing a risk-sharing agreement. After this the supervisor completes her task. Then renegotiation takes place

between the principal and the supervisor before the agent completes his own task.

In such a situation, whatever the supervisor's remuneration, given that the supervisor is risk-averse and the principal is risk-neutral and at the renegotiation stage the supervisor has completed her task, the principal will provide the supervisor with full insurance. One might conclude from this observation that the supervisor, foreseeing the outcome of the renegotiation with the principal, will provide the agent with full insurance at the collusion stage, eliminating any incentive from the agent's remuneration and maximizing the efficiency losses to the organization. This intuition however is not correct in our setting. Indeed, since renegotiation follows collusion, by subgame perfection the supervisor and the agent will internalize the renegotiation agreement between the principal and the supervisor and its effect on the agent's effort choice when choosing their optimal, collusive agreement. In other words, by leaving some risk to the agent at the collusion stage, the supervisor and the agent will maximize the stakes of collusion and in this way internalize the efficiency losses that renegotiation and collusion might generate.

The principal, then, by fine-tuning the risk present in the joint remuneration of the supervisor and the agent, may actually restore second-best efficiency, reducing to zero any efficiency losses induced by collusion and renegotiation. Of course, since it is not possible to offer the supervisor and the agent a risky joint remuneration and at the same time prevent the supervisor from taking at least some risk at the collusion stage, any initial contract offered to the supervisor will necessarily be renegotiated. This is not true for the collusion agreement.

If collusion is observable to every party in the organization, in solving for the optimal incentive scheme, only the joint remuneration offered to the supervisor and the agent by the initial contract matters. In other words, there exists a degree of freedom in the way in which the initial contract allocates the joint remuneration between the agent and the supervisor. This implies that by offering, at the initial stage, both the agent and the supervisor the same net remuneration they will be left with after collusion, any collusion will be prevented in equilibrium, although the contract will still be renegotiated. This degree of freedom disappears if collusion is not observable to the principal.

It is important to notice that it is not possible to restore second-best efficiency if the outcome of the renegotiation between the principal and the supervisor is independent of the collusion agreement between the supervisor and the agent. This is true if, for example, the renegotiation contract is agreed upon before collusion takes place. In this case the supervisor will be left after the collusion stage with a

risky payoff which translates into a loss of efficiency, which means that the renegotiation contract is not *ex post* optimal.

This observation is key to the intuition of why, in our framework, when collusion is not observable to the principal, collusion occurs in equilibrium and the principal suffers from the presence of both collusion and renegotiation. Indeed, since in our framework the principal has all the bargaining power at the renegotiation stage and collusion is not observable, the renegotiation stage looks like an adverse-selection model in which the type of the supervisor is characterized by the collusive agreement she accepted from the agent. However, since the way in which the principal can separate different types of supervisor is through their willingness to accept the renegotiation offer, everything is as if the renegotiation contract and the collusion contract were chosen simultaneously. This implies that the renegotiation contract is given when collusion is chosen. The intuition discussed above then applies, and there does not exist a pure-strategy equilibrium of the collusion-and-renegotiation subgame in which the agent exerts a strictly positive effort. We further show that the only equilibrium compatible with the agent's incentives to exert a positive effort is a mixed-strategy equilibrium in which both collusion and renegotiation occur in equilibrium with positive probability.

For some parameter values this is the only equilibrium compatible with the optimal incentive scheme for the organization. It is worth noticing that in the case in which collusion is not observable, the fact that both collusion and renegotiation occur in equilibrium with strictly positive probability is independent of our assumption that the principal cannot offer contracts contingent on the outcome of the collusive agreement.

The rest of the paper is organized in the following way. We start by presenting the structure of the model (Sec. 2), the timing (Sec. 3), and the benchmark incentive scheme, that is, the optimal incentive scheme when collusion is not feasible for exogenous reasons (Sec. 4). Sections 5 and 6 describe, respectively, collusion and renegotiation in our framework. The optimal incentive scheme in the case in which collusion is observable to all the members of the organization is derived in Section 7. An example of this incentive scheme in the case in which contracts are restricted to be linear and the state of nature is normally distributed is presented in Section 8. In Section 9 we characterize the features of the optimal incentive scheme in the case in which collusion is not observable to the principal. Section 10 concludes the paper.

1.2 RELATED LITERATURE

Two strands of literature are related to the analysis of this paper: the literature on collusion and the literature on renegotiation, in particular the papers that analyze renegotiation in agency contracts.

The literature on collusion is essentially divided in two groups of papers. The first group are papers—pioneered by Tirole (1986)—that analyze adverse-selection models. In these papers the stake of collusion consists of the surplus that employees can capture by not revealing to the principal the private information they have and are supposed to report. In this setting the *collusion-proof principle* holds. The optimal incentive scheme can be implemented by preventing any collusive agreement in equilibrium.¹ Our analysis differs from these papers in that, as discussed below, we focus on a different stake of collusion and renegotiation. Moreover, in our setting, when collusion is not observable to the members of the organization, it is not possible to implement the optimal incentive scheme in a collusion-proof manner.

The second group are papers that analyze collusion in agency models (Holmström and Milgrom, 1989; Varian, 1989; Itoh, 1993). In these papers the stake of collusion comes from the risk that the incentive scheme imposes on individual employees' remuneration and that can be profitably shared among risk-averse employees. This is the type of collusion we analyze in this paper. In this literature the collusion-proof principle holds as well.² We differ from these papers in introducing the possibility of renegotiation between the principal and one of the employees (the supervisor). This renegotiation could be interpreted as an additional collusion (risk-sharing) opportunity, this time between the principal and the supervisor.

The other literature of relevance for our analysis is the one on renegotiation. Renegotiation was first identified as a constraint for an optimal incentive scheme by Dewatripont (1989) in an adverse-selection setting. In this setting renegotiation opportunities arise because of dynamic changes in the parties' information structure. This is not the type of renegotiation we focus on.

1. A notable exception to the collusion-proof principle is Kofman and Lawarrée (1996). In their setting preventing collusion is too costly with respect to the efficiency losses that collusion introduces in the organization. As a result, for certain parameter values it is optimal for the designer of the incentive scheme to let the parties collude.

2. In this literature a failure of the collusion-proof principle is presented in Itoh (1993). In that paper it is optimal for the principal to allow the parties to collude. The reason is that the parties have superior information to the principal and by colluding they use this superior information efficiently. In other words, collusion is beneficial to the organization, since it induces the parties to better exploit their private information. In our setting this is not the case. The colluding parties have the same information structure as the principal, and hence collusion is potentially harmful to the organization.

We model renegotiation as a risk-sharing agreement between the principal and the supervisor. The stake of renegotiation is created by the risk that the possibility of collusion introduces in the supervisor's remuneration that the principal has an *ex post* interest to insure. Subgame perfection however implies that by insuring the supervisor the principal indirectly provides the agent with (partial) insurance as well. Indeed, at the collusion stage both the agent and the supervisor can foresee the outcome of the future renegotiation and share their risk accordingly. In this respect the renegotiation we model is closer in nature to the one analyzed in the literature on renegotiation in agency contracts (Fudenberg and Tirole, 1990; Hermalin and Katz, 1990; Ma, 1994; Matthews, 1995). The difference with the latter group of papers is that these papers allow the principal to renegotiate directly with the agent, while this is only indirectly possible in our setting, where a renegotiation opportunity is introduced by the possibility of collusion between the employees (agent and supervisor) of the organization.

A recent paper on renegotiation that is related to our analysis is Reiche (1999). That paper analyzes an optimal contract between two asymmetrically informed parties in the presence of renegotiation. One of its main results is that the renegotiation-proof principle does not hold. In other words, for certain parameter values the unique continuation equilibrium of the renegotiation subgame is such that renegotiation occurs in equilibrium with strictly positive probability. Although it focuses on a different model, the logic behind the fact that in Reiche (1999) parties renegotiate the optimal contract with strictly positive probability is similar to the logic behind the mixed-strategy equilibrium of the collusion-and-renegotiation subgame that we analyze in Section 9 below when collusion is not observable.

The last aspect of the literature on collusion that needs to be mentioned here is the enforcement mechanism of collusive contracts. At first the literature on collusion has simply assumed that side contracts are regular contracts and can be enforced using a court, not being modeled, in the background (Tirole, 1986). A small literature has recently developed that models explicitly the enforcement mechanism of side contracts (Felli, 1996; Acemoglu, 1996; Martimort, 1997). Given the convert nature of a collusive agreement, a side deal needs to be self-enforcing. This enforcement mechanism can be, for example, an exogenous penalty that each party can impose on his/her counterpart if he/she does not perform according to the side deal (Felli, 1996) or a punishment strategy that each party can use to enforce a given equilibrium behavior of the other colluding party (Acemoglu, 1996; Martimort, 1997). Being explicit on this mechanism allows the principal to use more effective and possibly cheaper ways to prevent collusion.

In our analysis we follow the original approach of the literature: we assume that side contracts are fully enforceable like regular contracts, and we do not model the enforcement mechanism. However, to respect at least the covert nature of collusive agreement, we assume that no other contract can be made contingent on the allocation induced by a collusive agreement. In other words, no contract might require the verifiability in court of a collusive agreement.

2. THE PARTIES

The framework of our analysis is a very simple three-level hierarchy. The top of the hierarchy is the residual claimant of profits generated by the whole structure: the principal (P). The bottom layer is the agent (A), the only level that actually produces any output. The intermediate layer is a supervisor (S), who is capable of collecting information on the agent's unobservable characteristics.

The *agent* is the productive unit of the structure; he controls a random technology that can generate two possible outcomes, which we normalize to zero for the low outcome, and one for the high outcome. When born, the agent is endowed with a productivity parameter θ , $\theta \in \{\bar{\theta}, \underline{\theta}\}$, $0 < \underline{\theta} < \bar{\theta} < 1$, which is his private information. He decides how much productive effort to exert: $e \in [0, 1 - \bar{\theta}]$. This effort is unobservable to third parties. For a given productivity level of the agent, the probability that the technology will generate a high outcome increased in the effort level exerted by the agent. In particular, we will assume

$$\bar{\theta} + e = \Pr\{x = 1|e, \bar{\theta}\}, \quad (1)$$

$$\underline{\theta} + e = \Pr\{x = 1|e, \underline{\theta}\}. \quad (2)$$

These equations translate the intuitive idea that the marginal productivity of effort increases in the productivity θ .

Preferences of the risk-averse agent are described by the following Von Neumann–Morgenstern utility function separable in income and effort: $U(w) - G(e)$. We assume that the utility of income, $U(\cdot)$, is bounded from above and satisfies $U'(\cdot) > 0$, $U''(\cdot) < 0$. The disutility of effort is assumed to have the following properties: $G(0) = 0$, $G'(0) = 0$, $G'(\cdot) \geq 0$, $G''(\cdot) > 0$, and $\lim_{e \rightarrow 1 - \bar{\theta}} G(e) = +\infty$. These properties also guarantee that the optimal effort level is always positive. The agent's reservation utility is $U^* = U(w^*)$.

The *supervisor* has a monitoring role in the structure. She does not contribute to the productive process, but just provides information. She has the time and the willingness to collect information on

the agent's productivity and, if requested, can supply such information to the principal. We model this by assuming that the supervisor observes costlessly and perfectly the agent's productivity and that this is *hard* information—in the way Tirole (1986) defines this term. In other words, we assume that the supervisor has to document every report she makes to the principal on the agent's productivity, and she has no way to produce enough supporting documentation for a false report.

Therefore any outside party—the principal in particular—can verify the truth of the supervisor's report.³

Preferences of the risk-averse supervisor are described by the following Von Neumann–Morgenstern utility function $V(s)$, strictly concave in income: $V'(\cdot) > 0$, $V''(\cdot) < 0$. The supervisor has an outside option with a reservation salary s^* .

The *principal* is a risk-neutral individual: he observes both the outcome of the productive process and the report of the supervisor which are both *verifiable* to third parties.

3. THE TIMING AND SOLUTION CONCEPT

In this section we describe the information structure and the extensive form of our model.

The information structure is such that before contracting the agent knows his unobservable productivity while the other parties share a common prior $q \equiv \Pr\{\theta = \bar{\theta}\}$. Negotiation takes place among the principal, the supervisor, and the agent. The principal is assumed to have all the bargaining power: he proposes a take-it-or-leave-it offer C (contract) to both the agent and the supervisor, which specifies a schedule of compensations for both employees as a function of the outcome and the supervisor's report. The agent and the supervisor observe each other's contracts and take the decision to accept or reject C , simultaneously and independently.

If the contract is accepted, then the supervisor learns the productivity of the agent, and collusion between the agent and the supervisor may take place. We assume, for simplicity, that in the collusion game the agent has all the bargaining power and makes a take-it-or-leave-it offer to the supervisor. The supervisor can only accept or reject the offer.

The supervisor then produces a report for the principal. This report is public information. Renegotiation between the principal and

3. This is just one example of a supervisor's task that is imperfectly correlated with the agent's task. Any other task with the same imperfect correlation will be compatible with our analysis. Indeed, in Section 8 below we leave the supervisor's task unspecified; we just require the principal to hire a supervisor and specify her reservation salary.

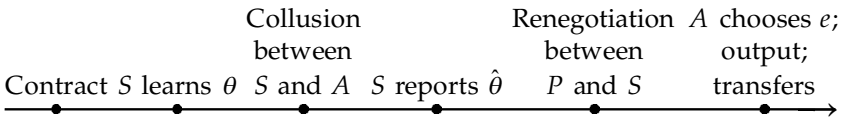


FIGURE 1. TIMING OF THE MODEL

the supervisor follows. Once again, for simplicity, we assume that the principal has all the bargaining power and makes a take-it-or-leave-it offer to the supervisor.⁴

Finally, the agent chooses effort, uncertainty is resolved, and the three parties exchange transfers according to the latest contractual agreements. The timing is summarized in Figure 1.

We look for a *perfect Bayesian equilibrium* of this game.

4. THE COLLUSION-FREE CONTRACT

We start by analyzing the benchmark case in which collusion is not a feasible option (for some exogenous reason) for the parties involved in the main contract *C*. This benchmark is of particular interest in our analysis, since in Section 7 below we prove that even in the presence of collusion and renegotiation the parties will be able to achieve the same allocation of resources we identify in this collusion-free environment.

Since collusion is not an issue, the risk-neutral principal pays a constant salary s^* to the risk-averse supervisor, who accurately reports the agent’s productivity. The principal receives perfect information on the productivity of the agent and faces only the moral-hazard problem of inducing the agent to exert some level of unobservable effort.

Let us consider the problem in the case where the productivity of the agent is θ :⁵

$$\max_{\{w_1\}, e} (\theta + e)(1 - w_1) - (1 - \theta - e)w_0 \tag{3}$$

$$\text{s.t. } (\theta + e)U(w_1) + (1 - \theta - e)U(w_0) \geq U^* + G(e), \tag{4}$$

$$U(w_1) - U(w_0) = G'(e). \tag{5}$$

Problem (3) is quite standard. Equation (4) is the agent’s *individual rationality* constraint. It states that the agent must obtain at least his

4. It should be said that by assumption in our setting no other renegotiation between the members of the organization may occur. Indeed, in our model, the agent and the supervisor may want to renegotiate their collusive agreement after the supervisor-principal renegotiation. Moreover, once the collusive agreement is renegotiated, also the principal and the supervisor may want to renegotiate their contract, and so on for a possibly infinite sequence of renegotiations. For the sake of simplicity and tractability we choose to truncate their sequence at its first step.

5. We use w_x to indicate wages offered to the agent of a general type θ , where the subscript refers to the final outcome x .

reservation utility. Equation (5) is the agent's *incentive compatibility* constraint, making him prefer to exert effort e in equilibrium. The solutions to problem (3) are such that, for a given e , condition (4) is satisfied with equality, and condition (5) is satisfied. Condition (4), namely the agent's binding individual rationality constraint, is the consequence of our assumption that the whole bargaining power in the negotiation lies with the principal. These two conditions determine w_1 and w_0 as a function of e such that

$$U(w_0) = U^* + G(e) - (\theta + e)G'(e) \quad (6)$$

and

$$U(w_1) = U^* + G(e) + (1 - \theta - e)G'(e). \quad (7)$$

The optimal e can then be computed from the condition

$$1 - w_1 + w_0 - (\theta + e)(1 - \theta - e)G''(e) \left(\frac{1}{U'(w_1)} - \frac{1}{U'(w_0)} \right) = 0. \quad (8)$$

The three conditions (6), (7), and (8) show that a higher θ corresponds to a higher e and a lower w_0 . The fact that the optimal w_i changes with θ makes the information regarding θ valuable to the principal. Hence for a low enough s^* the principal wants to hire the supervisor in the organization.

5. COLLUSION

Collusion in our model takes the form of an agreement between two parties who exchange bribes with the sole purpose of redistributing the risk between themselves.⁶ We assume for the sake of simplicity that this collusion takes a monetary form, and that bribes transfer wealth between individuals.

We start by assuming that the collusion between the agent and the supervisor is observable by the principal but not verifiable. This means that the renegotiation transfers between the supervisor and the principal cannot be made contingent on the collusion transfers between the supervisor and the agent.⁷ The case where collusion is not observable to the principal is presented in Section 9 below.

6. This is the type of collusion considered in Holmström and Milgrom (1989), Varian (1989), and Itoh (1993) and discussed in Section 1.2 above.

7. For this purpose we also need to rule out, by assumption, the possibility for the renegotiating parties to write a contract contingent not on the collusion transfers but rather on the agent's and the supervisor's reports of the size of such transfers. In other words we rule out by assumption message-contingent contracts à la Maskin and Tirole (1999).

The type of collusion we consider requires the parties to exchange bribes contingent on the final outcome of the random production technology. Furthermore, we assume that the agent has all the bargaining power in the collusion game with the supervisor,⁸ and that the contract signed by the agent and the supervisor with the principal specifies the wage and salary schedules $w_i, s_i, i \in \{0, 1\}$, contingent on the report of the supervisor.⁹ Then at the collusion stage a type- θ agent computes the maximal expected utility he could reach when he exchanges bribes with the supervisor, adjusts his effort level, and leaves the supervisor with at least the same expected utility she would enjoy in the absence of collusion.

The collusion contract that the agent offers the supervisor, therefore, solves the following program:

$$\max_{\{b_i\}, e} (\theta + e)U(w_1 - b_1) + (1 - \theta - e)U(w_0 - b_0) - G(e) \tag{9}$$

$$\text{s.t. } (\theta + e)V(\tilde{s}_1(b_1, b_0) + b_1) + (1 - \theta - e)V(\tilde{s}_0(b_1, b_0) + b_0) \geq (\theta + \hat{e})V(\tilde{s}_1(0, 0)) + (1 - \theta - \hat{e})V(\tilde{s}_0(0, 0)), \tag{10}$$

$$U(w_1 - b_1) - U(w_0 - b_0) = G'(e), \tag{11}$$

$$U(w_1) - U(w_0) = G'(\hat{e}), \tag{12}$$

where $b_i, i \in \{0, 1\}$, denote the bribes that the supervisor and agent exchange, and $\tilde{s}_i(b_1, b_0)$ the equilibrium salary offers received and accepted by the supervisor at the renegotiation stage.¹⁰ These offers are contingent on b_i because collusion is observable and renegotiation occurs after collusion.¹¹ Finally, \hat{e} denotes the effort level of the agent if the collusion agreement is rejected by the supervisor.

8. Our results should generalize to different distributions of the bargaining power in the negotiation of the collusion contract. The difference is that in the more general case the original contract (which remains a take-it-or-leave-it offer) has to allow for the facts that the supervisor will gain some surplus at the collusion stage and that the collusion bribes being exchanged are now different.

9. The definition of the symbols used to indicate salaries for the supervisor is symmetric to the one of the symbols that indicates wages for the agent (see footnote 5).

10. Notice that there are w_i and s_i such that the first-order conditions of problem (9) determine the optimum. In order to see this, note that the constraint (10) is binding at the optimum, and that the constraints (10) and (11) determine b_1 and b_0 uniquely as smooth functions of e , for all $e \in [0, 1 - \bar{\theta}]$. Then problem (9) can be set up as a problem of maximizing a smooth function of e for $e \in [0, 1 - \bar{\theta}]$. The case of $e = 1 - \bar{\theta}$ cannot be the optimum, because $G(1 - \bar{\theta}) = \infty$ and $U(\cdot)$ is bounded from above. The objective function is increasing at $e = 0$ if $w_1 - w_0$ and $s_1 - s_0$ are sufficiently large. Then, there must be an $e \in (0, 1 - \bar{\theta})$ that satisfies the first-order conditions and is an optimum.

11. Given our assumption that the principal has all the bargaining power at the renegotiation stage, we will be able to replace $\tilde{s}_i(b_1, b_0)$ in (10) by s_i , the payments to the supervisor in the original contract. Both sets of salaries yield the level of utility $V(s^*)$ to the supervisor.

The solution to problem (9) determines whether the agent makes a collusive offer to the supervisor, as well as how he adjusts his effort level. Indeed, it is possible that collusion between the supervisor and the agent induces the agent to exert a different effort level from the one (naively) considered by the principal in the initial contractual offer.

6. RENEGOTIATION

Renegotiation of the initial contract between the principal and the supervisor follows the collusion game between the agent and the supervisor. We assume, once again for simplicity, that the principal has all the bargaining power in this renegotiation.¹² Therefore the principal's offer to the supervisor solves the following minimization problem:

$$\min_{\{\tilde{s}_i\}} (\theta + e)\tilde{s}_1 + (1 - \theta - e)\tilde{s}_0 \tag{13}$$

$$\begin{aligned} \text{s.t. } & (\theta + e)V(\tilde{s}_1 + b_1) + (1 - \theta - e)V(\tilde{s}_0 + b_0) \\ & \geq (\theta + e)V(s_1 + b_1) + (1 - \theta - e)V(s_0 + b_0), \end{aligned} \tag{14}$$

where $s_i, i \in \{0, 1\}$, is the initial salary schedule offered to the supervisor (contingent on her own report), while the side transfers to which principal and supervisor agree at the renegotiation stage are given by $\tilde{s}_i - s_i, i \in \{0, 1\}$. On the other hand, $b_i, i \in \{0, 1\}$, are the bribes the supervisor and the agent agreed upon at the collusion stage. The solution to this problem determines the $\tilde{s}_i(b_1, b_0)$ that we use in the statement of problem (9). Notice that the first-order conditions of problem (13) imply that

$$\tilde{s}_1(b_1, b_0) + b_1 = \tilde{s}_0(b_1, b_0) + b_0 \quad \forall b_1, b_0. \tag{15}$$

Condition (15) allows us to prove the following result.

LEMMA 1: *Every renegotiation-proof optimal contract fully insures the supervisor against the uncertainty on the final outcome of the production process.*

Proof. Given (15), every renegotiation-proof contract needs to specify $s_1 + b_1 = s_0 + b_0$. □

12. Our results should generalize to situations with different distributions of the bargaining power at the renegotiation stage. The difference is that in the more general case the original contract offered by the principal takes account of the gain in surplus by the supervisor at the renegotiation stage. The renegotiation contract will still fully insure the supervisor at the reservation salary s^* .

The message of Lemma 1 is very intuitive. A risk-neutral principal will always fully insure a risk-averse supervisor when the principal does not need to provide the supervisor with any incentive (the role of the supervisor in the organization is completed: she does not exert any effort or report any new information). Obviously, if the original contract is collusion-proof (that is, $b_1 = b_0 = 0$), Lemma 1 implies that the only renegotiation-proof contract between the principal and the supervisor requires a constant remuneration for the supervisor $s_1 = s_0$.

7. EQUILIBRIUM CONTRACTS

In this section we characterize the optimal incentive scheme for our model. In particular, we identify the optimal renegotiation, collusion, and initial contracts. This will allow us to prove our main results.

First, we show that collusion-proof and renegotiation-proof contracts for the agent and the supervisor are incompatible with any positive effort exerted by the agent (Proposition 1). We then show that, in spite of the fact that at the renegotiation stage the principal provides the supervisor with full insurance, at the collusion stage the agent and the supervisor optimally agree on leaving some risk in the agent's remuneration so as to provide the agent with the incentives to exert a strictly positive effort (Proposition 2). We then proceed to show that by adjusting the risk that the initial contract leaves in the joint remuneration of both the agent and the supervisor, the principal can align the agent's incentives so as to restore second-best efficiency (Proposition 3).

We proceed by solving our model backwards. In order to do this we compute the optimal collusion contract given the continuation equilibrium in the renegotiation stage, which we solved for in Section 6 above. Given that the principal has all the bargaining power at the renegotiation stage, we have that

$$\begin{aligned} & (\theta + e)V(\tilde{s}_1(b_1, b_0) + b_1) + (1 - \theta - e)V(\tilde{s}_0(b_1, b_0) + b_0) \\ & = (\theta + e)V(s_1 + b_1) + (1 - \theta - e)V(s_0 + b_0) \quad \forall b_1, b_0. \end{aligned} \quad (16)$$

Equation (16) implies that we can rewrite problem (9), which characterizes the optimal collusion contract between the agent and the supervisor, by replacing the individual rationality constraint (10) with the one that follows.¹³

13. Alternatively—and leading obviously to the same results—one can solve problem (9) with the original constraint (10), keeping in mind the expression for $\partial \tilde{s}_i / \partial b_j$ for $i, j = 0, 1$, as obtained from the binding constraint of problem (13).

$$\begin{aligned} & (\theta + e)V(s_1 + b_1) + (1 - \theta - e)V(s_0 + b_0) \\ & \geq (\theta + \hat{e})V(s_1) + (1 - \theta - \hat{e})V(s_0). \end{aligned}$$

Notice that this new condition is exactly the one we would have in the absence of any renegotiation.

We now proceed to compute the equilibrium bribes b_1 and b_0 for given w_1 , w_0 , s_1 , and s_0 . For this purpose we first obtain from (11)

$$\frac{\partial e}{\partial b_0} = \frac{U'(w_0 - b_0)}{G''(e)}, \quad \frac{\partial e}{\partial b_1} = -\frac{U'(w_1 - b_1)}{G''(e)}.$$

Then from the first-order conditions of problem (9) we get

$$\frac{V'(s_1 + b_1)}{U'(w_1 - b_1)} = \frac{V'(s_0 + b_0)}{U'(w_0 - b_0)} + \frac{V(s_1 + b_1) - V(s_0 + b_0)}{(1 - \theta - e)(\theta + e)G''(e)}. \quad (17)$$

Condition (17) represents an efficient way for two risk-averse individuals to share risk. The modification to the classic coinsurance rule—the second term on the right-hand side—is due to the moral-hazard constraint.

The following lemma can be proved directly from condition (17).

LEMMA 2: *Every collusion-proof optimal contract needs to satisfy the following modified version of an optimal coinsurance rule:*

$$\frac{V'(s_1)}{U'(w_1)} = \frac{V'(s_0)}{U'(w_0)} + \frac{V(s_1) - V(s_0)}{(1 - \theta - e)(\theta + e)G''(e)}. \quad (18)$$

Using Lemma 1 and Lemma 2, we can now prove our first result.

PROPOSITION 1: *No equilibrium contract between the principal and the supervisor can be at the same time collusion-proof and renegotiation-proof and provide the agent with enough incentives to exert positive effort.*

Proof. From Lemma 1, whenever the original contract is both collusion-proof and renegotiation-proof, $s_1 = s_0$. Substituting $s_1 = s_0$ in (18), one obtains $w_1 = w_0$, which is a wage schedule that yields zero effort level. \square

Proposition 1 identifies the basic trade-off of our model. On the one hand, the need of the principal to induce the agent to exert a productive effort requires a risky wage schedule for the agent. On the other hand, since the principal cannot commit not to renegotiate with the supervisor, he will provide the supervisor with full insurance at the renegotiation stage. The result is that the only way in which the

principal can offer the agent a wage schedule that is collusion- and renegotiation-proof is to offer a constant wage schedule that is clearly incompatible with any positive effort choice.

The result is that it is not possible to achieve all three objectives: a positive effort by the agent, a collusion-proof contract, and a renegotiation-proof contract. Given that $G'(0) = 0$, the principal will always choose to offer an incentive scheme where the agent exerts a positive effort. In the next result of this section we establish which of the two features of an optimal contract, renegotiation-proofness or collusion-proofness, the principal has to give up.

Notice first that from (17), whenever the agent exerts effort in equilibrium, the contract between the principal and the supervisor will be renegotiated on the equilibrium path. Indeed, from Lemma 1 no renegotiation implies $s_1 + b_1 = s_0 + b_0$, which in (17) implies $w_1 - b_1 = w_0 - b_0$, that is, no positive effort.

What about collusion? It turns out that in this case the set of optimal contracts includes a collusion-proof contract. To see this, assume that an optimal contract, $(w_1^a, w_0^a, s_1^a, s_0^a)$, involves some level of collusion (b_1^a, b_0^a) satisfying (17). The level of effort exerted in this contract is determined by $U(w_1^a - b_1^a) - U(w_0^a - b_0^a) = G'(e)$. The payoff for the agent is $w_1^a - b_1^a$ if production occurs, and $w_0^a - b_0^a$ if production does not occur. The payoff for the supervisor, on the other hand, is $s^* = \tilde{s}_1^a + b_1^a = \tilde{s}_0^a + b_0^a$, which yields her the same expected utility as if she got $s_1^a + b_1^a$ if production occurred and $s_0^a + b_0^a$ if production did not occur. Notice that this would be the supervisor's payoffs in the absence of any renegotiation stage in the game. Finally, the payoff for the principal is $1 - w_1^a - (s^* - b_1^a)$ if production occurs and $-w_0^a - (s^* - b_0^a)$ if production does not occur. Notice also that if renegotiation is not a feasible opportunity, then the payoff to the principal is, respectively, $1 - w_1^a - s_1^a$ and $-w_0^a - s_0^a$.

Can we replicate this contract with a collusion-proof contract that yields the same payoffs to all the participants, and the same effort level in equilibrium? Consider the contract $(w_1^b, w_0^b, s_1^b, s_0^b)$ such that $w_1^b = w_1^a - b_1^a$, $w_0^b = w_0^a - b_0^a$, $s_1^b = s_1^a + b_1^a$, and $s_0^b = s_0^a + b_0^a$. Then, the agent and the supervisor do not collude on the equilibrium path, since condition (18) is satisfied. Furthermore, the agent ends up exerting the same effort level as in the contract that involves collusion, and the payoffs to both the agent and supervisor remain exactly the same in each event. Finally, the principal gets now $1 - w_1^b - s^* = 1 - w_1^a + b_1^a - s^*$ if production occurs, and $-w_0^b - s^* = -w_0^a + b_0^a - s^*$ if production does not occur, which are exactly the principal's payoffs when the optimal contract involves collusion. We summarize these results in the following proposition.

PROPOSITION 2: *If positive effort is exerted in equilibrium, every optimal contract between the principal and the supervisor is renegotiated. Furthermore, the set of optimal contracts between the principal and the agent includes an optimal collusion-proof contract.*

This proposition allows us to restrict attention to the collusion-proof contract between the principal and the agent (whether or not renegotiation is feasible).

We can now proceed to show that when the principal and the supervisor have the opportunity to renegotiate, the principal is able to achieve the same payoff as in the collusion-free environment (Sec. 4 above).

Consider a triple (w_1, w_0, e) that solves problem (3). When collusion is not feasible, then the payoff to the principal is $1 - w_1 - s^*$ with probability $\theta + e$ and $-w_0 - s^*$ with probability $1 - \theta - e$.

Assume now that collusion and renegotiation are both feasible. Using the triple (w_1, w_0, e) , construct a pair (s_1, s_0) satisfying (18) and

$$(\theta + e)V(s_1) + (1 - \theta - e)V(s_0) = V(s^*).$$

Then, by Lemma 2, no collusion occurs in equilibrium. Notice that this implies that $s_1 \neq s_0$. Therefore, at the renegotiation stage the principal renegotiates the payments to the supervisor from (s_1, s_0) to (s^*, s^*) , so as to provide her with full insurance. Then the payoff to the principal is $1 - w_1 - s^*$ with probability $\theta + e$ and $-w_0 - s^*$ with probability $1 - \theta - e$. This is exactly the same payoff the principal obtains in the case in which collusion is not feasible.

Notice also, that renegotiation is essential for this result. Indeed, without renegotiation the principal gets a payoff of $1 - w_1 - s_1$ with probability $\theta + e$ and of $-(w_0 + s_0)$ with probability $1 - \theta - e$. This is worse than the payoff the principal gets when renegotiation is feasible:

$$(\theta + e)V(s_1) + (1 - \theta - e)V(s_0) = V(s^*) < V((\theta + e)s_1 + (1 - \theta - e)s_0).$$

In other words, given the concavity of $V(\cdot)$,

$$s^* < (\theta + e)s_1 + (1 - \theta - e)s_0.$$

In this way we have proved that when collusion is observable to the principal, the feasibility of renegotiation makes the efficiency losses of collusion disappear. We summarize this result in the following proposition.

PROPOSITION 3: *Collusion, when observable, is harmful to the principal. However, when renegotiation follows collusion, the principal can obtain the second-best payoff, as in the collusion-free environment.*

Notice that the situation in which collusion is observable to the principal is similar to a simple moral-hazard problem between a principal and an agent in which the agent's effort is observable and renegotiation is feasible after effort is exerted and before the state of nature is revealed (see Hermalin and Katz, 1990). Indeed, in both cases the presence of renegotiation allows the principal to improve his payoff. For example, in Hermalin and Katz (1990) the principal is able to obtain a first-best payoff (as if there were no agency problem).

As we have proved above, collusion and renegotiation are not necessarily harmful to the principal if he accounts for them in the design of the optimal incentive scheme. In particular, as we clarify in the example presented in Section 8 below, this is obtained by offering a riskier joint remuneration to the supervisor and the agent and eliminating part of this risk at the renegotiation stage.

We conclude this section with two observations. First, Proposition 3 does not hold if the renegotiation is agreed upon before collusion takes place. Indeed, in such a case the supervisor's remuneration after renegotiation is given and *independent* of the bribes chosen at the collusion stage b_0 and b_1 :

$$\frac{\partial \bar{s}(b_0, b_1)}{\partial b_i} = 0 \quad \forall i \in \{0, 1\}.$$

This implies that the supervisor will have a risky remuneration, and therefore the principal would have liked to change the renegotiation offer. We will come back to this point in Section 9, where collusion is not observable to the principal and hence everything is as if the renegotiation contract and the collusion contract were chosen simultaneously. In this case the only equilibria of the collusion and renegotiation subgame compatible with any positive effort exerted by the agent are mixed-strategy equilibria.

Secondly, collusion would become harmful to the principal had we introduced a lower bound (say zero) on the wages offered. In this case, if the supervisor has a sufficiently low risk aversion, the wage schedules for both the agent and supervisor may have to be so steep that the lower bound of w_0 and s_0 is binding. Then the collusion between the agent and the supervisor reduces the principal's payoff. The costs to the principal of collusion are higher for lower risk aversion of the supervisor and for higher risk aversion of the agent.

8. AN EXAMPLE

In order to illustrate the three results presented in Proposition 1, 2, and 3 above, we develop in this section a specific example of our model. The main feature of this example is that it is possible to compute explicit formulae for the agent's and the supervisor's wage schedules. This example is derived from Holmström and Milgrom (1987) and Holmström and Milgrom (1989).

Assume that the agent's technology is random and has as support the real line. More specifically, the outcome x of the agent's effort e is such that $x = e + \varepsilon$, where ε is a random variable, normally distributed with mean zero and standard deviation one. In addition, assume that the agent's preferences over income w and effort e are represented by the exponential function $U(w, e) = -\exp[-r(w - e^2/2)]$, where r is the coefficient of absolute risk aversion.¹⁴ Denote the agent's reservation wage as $U^* \equiv U(w^*, 0)$. In this example we abstract from θ (the agent's type, which is always revealed by the supervisor) in order to simplify the analysis. Including θ in the production technology is however a simple extension of this example.

Similarly, the supervisor's preferences over income s are represented by the function $V(s) = -\exp(-Rs)$, where R is the supervisor's coefficient of absolute risk aversion.

We restrict all contracts to be affine functions of the only verifiable variable in the model: the outcome x of the production technology. In other words, we take the initial contract between the principal and the agent to be $w(x) = ax + d$, and the one between the principal and the supervisor to be $s(x) = hx + m$. Similarly, we take the collusion contract between the agent and the supervisor to be $b(x) = \beta x + \delta$, and the renegotiation contract between the principal and the supervisor to be $\tilde{s}(x) = \tilde{h}x + \tilde{m}$.

We now proceed to solve our example backward, and we start from the agent's effort choice e . This is the outcome of the following problem that the agent solves:

$$\max_e -E_x \left\{ \exp \left[-r \left(ax + d - \beta x - \delta - \frac{e^2}{2} \right) \right] \right\}. \tag{19}$$

Problem (19) can be rewritten using the functional form and the distributional assumptions described above in the following way:

$$\max_e \left((a - \beta)e + (d - \delta) - \frac{r}{2}(a - \beta)^2 - \frac{e^2}{2} \right). \tag{20}$$

From the first order conditions of problem (20) we obtain $e = a - \beta$.

14. Notice that in this formulation the agent's utility function is not separable in income and effort, contrary to what is assumed in the previous sections. This difference does not affect the three results we presented in Section 7 above.

We now move to the renegotiation stage. The principal's optimal renegotiation offer solves the following problem:

$$\max_{\tilde{h}, \tilde{m}} E_x [x(1 - a - \tilde{h}) - \tilde{m} - d] \tag{21}$$

$$\begin{aligned} \text{s.t.} \quad & -E_x \{ \exp [-R(\tilde{h}x + \tilde{m} + \beta x + \delta)] \} \\ & \geq -E_x \{ \exp [-R(hx + m + \beta x + \delta)] \}, \end{aligned} \tag{22}$$

which can be transformed using our functional form and distributional assumptions in the equivalent problem.

$$\max_{\tilde{h}, \tilde{m}} (1 - \tilde{h} - a)e - \tilde{m} - d \tag{23}$$

$$\begin{aligned} \text{s.t.} \quad & (\tilde{h} + \beta)e + \tilde{m} + \delta - \frac{R}{2}(\tilde{h} + \beta)^2 \\ & = (h + \beta)e + m + \delta - \frac{R}{2}(h + \beta)^2. \end{aligned} \tag{24}$$

The solution to this problem yields $\tilde{h} = -\beta$, which confirms the fact that the supervisor is fully insured at the renegotiation stage. We also get

$$\tilde{m} = (h + \beta)(a - \beta) + m - \frac{R}{2}(h + \beta)^2.$$

We now move to the collusion stage. Given the assumption that the agent has all the bargaining power, the equilibrium collusion contract solves the following problem:

$$\max_{\beta, \delta} -E_x \left\{ \exp \left[-r \left(ax + d - \beta x - \delta - \frac{e^2}{2} \right) \right] \right\} \tag{25}$$

$$\begin{aligned} \text{s.t.} \quad & -E_x \{ \exp [-R(hx + m + \beta x + \delta)] \} \\ & \geq -E_x \{ \exp [-R(hx + m)] \}, \end{aligned} \tag{26}$$

$$e = a - \beta, \tag{27}$$

which can be transformed into the equivalent problem

$$\max_{\beta, \delta} (a - \beta)^2 + d - \delta - \frac{(a - \beta)^2}{2} - \frac{r}{2}(a - \beta)^2 \tag{28}$$

$$\text{s.t.} \quad (h + \beta)(a - \beta) + m + \delta - \frac{R}{2}(h + \beta)^2 = ha + m - \frac{R}{2}h^2. \tag{29}$$

From the first-order conditions of problem (28) we obtain the formulae for the coefficients of the optimal schedule,

$$\beta = \frac{r}{1 + R + r}a - \frac{1 + R}{1 + R + r}h$$

and

$$\delta = ha - \frac{R}{2}h^2 + \frac{r(a + h)^2}{2(1 + r + R)^2}(rR - 2 - 2R).$$

These formulae allow us to conclude that the slope β of the collusion schedule is increasing in the agent’s risk aversion r and in the slope a of his initial wage schedule. At the same time the slope of the collusion schedule is decreasing in the supervisor’s risk aversion R and in the slope h of the supervisor’s initial wage schedule.

Notice also that a particular choice of the slope of the supervisor initial contract h may lead to no collusion. This choice is

$$h = \frac{r}{1 + R}a.$$

By substitution we also obtain the formula for the effort level as a function of the initial contract:

$$e = \frac{1 + R}{1 + r + R}(a + h).$$

This is increasing in the risk aversion R of the supervisor and in the slopes a and h of the initial wage schedules of the agent and supervisor, and is decreasing in the agent’s risk aversion r .

We can now move to the optimal initial contracts among the principal and both the supervisor and the agent. This is obtained as the solution to the following problem:

$$\max_{a, d, h, m} E_x[x(1 - a - \tilde{h}) - d - \tilde{m}] \tag{30}$$

$$\text{s.t. } -E_x \left\{ \exp \left[-r \left(ax + d - \beta x - \delta - \frac{e^2}{2} \right) \right] \right\} \geq -\exp(-rw^*), \tag{31}$$

$$-E_x \left\{ \exp \left[-R(\tilde{h}x + \beta x + \delta + \tilde{m}) \right] \right\} \geq -\exp(-Rs^*), \tag{32}$$

where e , β , δ , \tilde{h} , and \tilde{m} were defined above as functions of a , d , h , and m .

Using our functional form and distributional assumptions, this problem is equivalent to

$$\max_{a, d, h, m} \left(1 - \frac{1+R}{1+R+r}(a+h) \right) \frac{1+R}{1+R+r}(a+h) - d - m + \frac{r(a+h)^2}{2(1+R+r)^2}(rR-2-2R) \quad (33)$$

$$\text{s.t. } \frac{(1+R)^2}{2(1+R+r)^2}(a+h)^2 + d - s^* + m - \frac{r(a+h)^2}{2(1+R+r)^2} \times (rR-2-2R) - \frac{r(1+R)^2}{2(1+R+r)^2}(a+h)^2 = w^*, \quad (34)$$

$$\frac{1+R}{1+R+r}h(a+h) + m - \frac{R}{2}h^2 = s^*. \quad (35)$$

From the first-order conditions of problem (33) and few algebra steps, we obtain the following slope of the joint remuneration schedules of the agent and the supervisor:

$$a+h = \frac{1+R+r}{(1+R)(1+r)}. \quad (36)$$

Equation (36) clearly shows that the principal has a degree of freedom in determining how much risk to introduce in the individual remunerations of both the agent and the supervisor. This degree of freedom is exactly what allows the principal to offer the agent a collusion-proof contract, as we proved in Proposition 3 above. It also identifies a whole continuum of contracts that achieve second-best efficiency (as Proposition 3 shows). All of them, with the exception of the unique collusion-proof one, will induce collusion in equilibrium.

Notice that the slope of the joint remuneration schedule of both the agent and supervisor, $a+h$, is decreasing in the risk aversion of either employee. In particular, it decreases in the risk aversion of the supervisor, R , because with a greater risk aversion, the supervisor is less willing to offer insurance to the agent. At the same time, it decreases in the risk aversion of the agent, r , because with a greater risk aversion, the principal chooses to offer less incentives to exert effort: these incentives become more expensive. The equilibrium effort level is then

$$e = \frac{1}{1+r}. \quad (37)$$

The effort level in (37) is the same that the agent would exert in a collusion-free environment. Hence, the principal does not incur any loss in terms of the level of surplus produced in the presence of both collusion and renegotiation.

Consider now the optimal collusion-proof contract. This contract is characterized by no collusion, ($\beta = \delta = 0$), of course, and by a flat renegotiated wage schedule for the supervisor ($h = 0$ and $\tilde{m} = s^*$). The slope of the initial contract is $a = 1/(1+r)$ for the agent and $h = r/[(1+R)(1+R)]$ for the supervisor. In other words, the extra degree of freedom is used to guarantee the collusion-proof feature of the initial contract. Finally, the intercepts of the remuneration schedules of the initial contracts for the agent and the supervisor are

$$d = w^* - \frac{1-r}{2(1+r)^2}, \quad m = s^* - \frac{r(2+2R-rR)}{2(1+R)^2(1+r)^2}.$$

The overall payoff to the principal under the optimal incentive scheme is

$$\frac{1}{2(1+r)} - w^* - s^*. \tag{38}$$

We compare now the agent's effort choice in (37) and the principal's payoff in (38) with the ones obtained in the case in which the principal and the supervisor cannot renegotiate. In the latter case, the slope of the joint remuneration schedule of the agent and supervisor is:¹⁵

$$\bar{a} + \bar{h} = \frac{(1+r+R)(1+R)}{r^2R + (1+R)(1+r+R+rR)}, \tag{39}$$

while the equilibrium effort level is

$$\bar{e} = \frac{(1+R)^2}{r^2R + (1+R)(1+r+R+rR)}. \tag{40}$$

The payoff to the principal is then

$$\frac{\bar{e}}{2} - w^* - s^*. \tag{41}$$

15. We denote by \bar{e} , \bar{h} , and \bar{a} the effort level and the slopes of the agent's and supervisor's remuneration in the case in which no renegotiation between the principal and the supervisor is allowed.

We can see now that, for any positive R , the effort level and the payoff to the principal in the absence of renegotiation, (40) and (41), are strictly smaller than the effort level and payoff for the principal when renegotiation is feasible, (37) and (38). These effort levels and payoffs coincide only if the supervisor is risk-neutral ($R = 0$) or if the supervisor is infinitely risk-averse ($R = \infty$).

Indeed, if the supervisor is risk-neutral, everything is as if the principal has sold the firm to the supervisor; the supervisor would then be the residual claimant of the agent's effort choice and would provide him with the right incentives to exert effort. If the supervisor is infinitely risk-averse, then she does not accept any risky insurance agreement offer at the collusion stage, and the problem is effectively as if collusion were not an issue.

In our example, the difference in effort level and principal's payoff in the two cases (renegotiation and no renegotiation) is largest when $R = 1$.

9. NONOBSERVABLE COLLUSION

The previous sections assume that the collusion contract, even though not verifiable, is observable to all members of the organization, principal included. In this section we consider the case in which the collusion contract is observable to the agent and the supervisor but not to the principal.

Then, the renegotiation stage becomes a bargaining game under asymmetric information: the supervisor knows the collusion contract she accepted from the agent, while the principal does not know the exact value of the promised bribes. We model this bargaining under asymmetric information in the same way as in the previous sections. We assume that the principal, the uninformed party, makes a take-it-or-leave-it offer to the supervisor. We further assume that the principal is restricted to make a unique renegotiation offer that consists of a pair of remunerations contingent on the two realizations of the outcome x . Notice that this implies that we do not allow the principal to offer a menu of renegotiation contracts that screen the supervisor's types [as, for example, in Fudenberg and Tirole (1990)].¹⁶ These types are represented by the collusion agreements the supervisor might have accepted at the collusion stage.

16. This is potentially a strong assumption in that, if the principal can separate the various types of supervisor, we can be in a similar situation to the one we analyzed in Section 7 above. We leave the analysis of this case for future research. However, in such a case—because of the arguments presented in this section—we would also expect the equilibrium to involve mixed strategies and both collusion and renegotiation to occur on the equilibrium path with strictly positive probability.

For the same reasons as in Fudenberg and Tirole (1990), in our environment, if the agent is provided with any incentives to exert effort, then the equilibrium of the collusion-and-renegotiation subgame also cannot be in pure strategies. The intuition for this result is simple to describe. At the renegotiation stage, the principal provides the supervisor with full insurance. Since this renegotiation contract is given when the collusion contract is agreed upon, the agent, if restricted to pure strategies, will make an offer at the collusion stage to the supervisor that transfers all the risks to the supervisor and through her to the principal. In other words, agent and supervisor behave as if they no longer internalized at the collusion stage the principal's renegotiation offer and, with it, the overall size of the surplus. Therefore full insurance could be a pure-strategy equilibrium with no incentives for the agent to exert any effort.

In more formal terms, assume, by way of contradiction, that the equilibrium of the collusion-and-renegotiation subgame is in pure strategies and that at the collusion stage the bribes being exchanged are (b_1, b_0) such that $w_1 - b_1 > w_0 - b_0$, so that the agent has incentives to exert effort. At the renegotiation stage the principal offers the supervisor a remuneration schedule $(\tilde{s}_1, \tilde{s}_0)$ such that $\tilde{s}_1 + b_1 = \tilde{s}_0 + b_0$. Given the timing and the information structure of the renegotiation subgame, the pair $(\tilde{s}_1, \tilde{s}_0)$ does not depend on the bribes (b_1, b_0) agreed upon at the collusion stage. This implies that the agent takes the offers $(\tilde{s}_1, \tilde{s}_0)$ as given at the collusion stage, and offers a pair of bribes (b_1, b_0) satisfying, from the first-order conditions of Problem (9),

$$\frac{V'(\tilde{s}_1 + b_1)}{U'(w_1 - b_1)} = \frac{V'(\tilde{s}_0 + b_0)}{U'(w_0 - b_0)}. \quad (42)$$

Condition (42) and $\tilde{s}_1 + b_1 = \tilde{s}_0 + b_0$ imply that $w_1 - b_1 = w_0 - b_0$, a contradiction to the hypothesis $w_1 - b_1 > w_0 - b_0$.

We summarize this result in the following proposition.

PROPOSITION 4: *If collusion is not observable, there is no equilibrium in pure strategies of the collusion-and-renegotiation subgame that provides any positive incentives for effort.*

A natural question to ask at this point is whether for any optimal choice of the original contract between the principal and both the supervisor and the agent the continuation equilibrium of the collusion-and-renegotiation subgame is such that the agent is left with no incentives to exert effort. We show below that this is not true. In particular there exist parameter values such that the optimal original contract (w_1, w_0, s_1, s_0) yields a continuation equilibrium such that the agent is left with enough incentives to exert effort. Of course,

given Proposition 4, in this case the continuation equilibrium of the collusion-and-renegotiation subgame is a mixed-strategy equilibrium.

Consider an initial contract with $s_1 = s_0$ and $w_1 - w_0$ sufficiently large, and a very risk-averse supervisor. Assume, by way of contradiction, that for all parameter values there exists a pure-strategy equilibrium pair of optimally chosen bribes (b_1, b_0) such that the agent is fully insured ($w_1 - b_1 = w_0 - b_0$) and the agent does not exert any effort. Because the equilibrium is in pure strategies, we also know that the supervisor will be fully insured at the renegotiation stage ($\tilde{s}_1 + b_1 = \tilde{s}_0 + b_0 = z$). Since at the renegotiation stage the principal is assumed to have all the bargaining power, we necessarily have that $V(z) = \theta V(s_1 + b_1) + (1 - \theta)V(s_0 + b_0)$. Recall that we assumed that the supervisor is very risk-averse. Then, because $s_1 = s_0$ and $b_1 > b_0$ (because $w_1 > w_0$), we conclude that z is close to $s_0 + b_0$.

Consider now the collusion stage. The supervisor has to decide whether to accept the offer (b_1, b_0) . If the supervisor does not accept the collusion offer, she can get either $[\theta + e(0, 0)]V(s_1) + [1 - \theta - e(0, 0)]V(s_0)$ if she also does not accept the principal's renegotiation offer—where $e(0, 0)$ represents the agent's effort level if no bribes are accepted—or $[\theta + e(0, 0)]V(z - b_1) + [1 - \theta - e(0, 0)]V(z - b_0)$ if she accepts the principal's renegotiation offer. Because z is close to $s_0 + b_0$, this last option yields the supervisor an expected utility close to $[\theta + e(0, 0)]V(s_0 - w_1 + w_0) + [1 - \theta - e(0, 0)]V(s_0)$, which is smaller than what she would obtain rejecting both the collusion and renegotiation offers, $V(s_0)$ (recall that $s_1 = s_0$). Now, since at the collusion stage the agent is assumed to have all the bargaining power, we must have that $V(z) = [\theta + e(0, 0)]V(s_1) + [1 - \theta - e(0, 0)]V(s_0)$, which means $z = s_0 = s_1$. This implies that b_0 is close to zero.

In this case then the agent would rather offer (b_1, b_0) and get $U(w_0 - b_0)$ (because $w_1 - b_1 = w_0 - b_0$) than make no offer and get $[\theta + e(0, 0)]U(w_1) + [1 - \theta - e(0, 0)]U(w_0) - G(e(0, 0))$. But if b_0 is close to zero we have $[\theta + e(0, 0)]U(w_1) + [1 - \theta - e(0, 0)]U(w_0) - G(e(0, 0)) > \theta U(w_1) + (1 - \theta)U(w_0) > U(w_0 - b_0)$, a contradiction.

We can now state this result in the following proposition.

PROPOSITION 5: *There exist parameter values of the model with nonobservable collusion such that the principal can design an optimal initial contract such that the continuation equilibrium of the collusion and renegotiation subgame, if it exists, is a mixed-strategy equilibrium that leaves the agent with incentives to exert effort.*

The two propositions above tell us that there is no equilibrium in pure strategies of the collusion-and-renegotiation subgame where effort is exerted in equilibrium. Moreover, under certain conditions the only equilibrium of the collusion-and-renegotiation subgame, if it exists, is in mixed strategies and induces the agent to exert effort.

In what follows we show that a mixed-strategy equilibrium of the collusion-and-renegotiation subgame does exist. In this case both collusion and renegotiation occur on the equilibrium path with strictly positive probability.

In order to prove existence, we assume that the collusion offers (b_1, b_0) are restricted to a (fine) grid $B_1^n \times B_0^n$ where the highest and lowest elements in B_1^n and B_0^n are finite (and with large absolute values).¹⁷ Similarly, we assume that the renegotiation offers $(\tilde{s}_1, \tilde{s}_0)$ are also restricted to a (fine) grid $\tilde{S}_1^n \times \tilde{S}_0^n$ where the highest and lowest elements in \tilde{S}_1^n and \tilde{S}_0^n are finite (and with large absolute values).

Define then a strategy for the agent as a function $p_A : B_1^n \times B_0^n \leftarrow [0, 1]$, where $p_A(b_1, b_0)$ represents the probability with which the agent offers bribes (b_1, b_0) . Then $\sum_{b_1} \sum_{b_0} p_A(b_1, b_0) = 1$. Similarly, define a strategy for the principal (at the renegotiation stage) as a function $p_P : \tilde{S}_1^n \times \tilde{S}_0^n \leftarrow [0, 1]$, where $p_P(\tilde{s}_1, \tilde{s}_0)$ represents the probability with which the principal makes renegotiation offers $(\tilde{s}_1, \tilde{s}_0)$. Then $\sum_{\tilde{s}_1} \sum_{\tilde{s}_0} p_P(\tilde{s}_1, \tilde{s}_0) = 1$. Finally, define a strategy for the supervisor as a pair of functions $p_{Sc} : B_1^n \times B_0^n \leftarrow [0, 1]$ and $p_{Sr} : B_1^n \times B_0^n \times \tilde{S}_1^n \times \tilde{S}_0^n \times \{1, 0\} \leftarrow [0, 1]$, where $p_{Sc}(b_1, b_0)$ represents the probability with which the supervisor accepts the bribes (b_1, b_0) at the collusion stage, and $p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, t)$ represents the probability with which a supervisor accepts the renegotiation offer $(\tilde{s}_1, \tilde{s}_0)$ when she receives an offer (b_1, b_0) at the collusion stage and her decision to accept the collusion offer is denoted by the binary variable t ($t = 1$ represents acceptance of the collusion offer, and $t = 0$ represents rejection).

We can now construct the best-response correspondences by defining the best response of each player: p_A^* , p_P^* , and (p_{Sc}^*, p_{Sr}^*) . These best responses are constructed in the Appendix. The best responses as in (A.1), (A.2), (A.3), (A.4), and (A.5) in the Appendix define a correspondence in the Cartesian product of the simplex on $B_1^n \times B_0^n$, of the simplex on $\tilde{S}_1^n \times \tilde{S}_0^n$, and of $[0, 1]^y$, where $y = \#(B_1^n \times B_0^n) + 2\#(B_1^n \times B_0^n \times \tilde{S}_1^n \times \tilde{S}_0^n)$.¹⁸

We have now all the elements to prove the following proposition.

PROPOSITION 6: *When collusion is not observable and collusion and renegotiation offers can only be made from a grid, there exists a perfect Bayesian equilibrium of the collusion-and-renegotiation subgame for any choice of the initial contract.*

17. The results in Dasgupta and Maskin (1986) cannot be applied directly here, because in the continuous case, the strategy of the supervisor is infinite-dimensional.

18. As is customary, we denote by the symbol $\#A$ the cardinality of the set A .

Proof. From Nash (1950) we know that the correspondence defined by (A.1), (A.2), (A.3), (A.4), and (A.5) satisfies the conditions of Kakutani's fixed-point theorem. Thus, this correspondence has a fixed point. Every fixed point of this correspondence is a perfect Bayesian equilibrium of the collusion-and-renegotiation subgame that follows the principal's choice of the original contract. \square

The last three propositions show that if collusion is not observable, the supervisor is sufficiently risk-averse, and the cost of exerting effort is not too high, then the principal designs an optimal incentive scheme for the organization that induces a nondegenerate mixed-strategy equilibrium of the collusion and renegotiation subgame. By definition this implies that in this case collusion and renegotiation occur in equilibrium with strictly positive probability.¹⁹ Notice that renegotiation in this case of nonobservable collusion is also playing a role similar to the one in the case of observable collusion: that of reducing the inefficiency resulting from the supervisor having a risky contract. However, because here the renegotiation offers cannot be a function of the collusion side payments, the supervisor is not fully insured with probability one.

10. CONCLUDING REMARKS

When the parties to an organization have the opportunity to collude against the optimal incentive scheme that the principal imposes on them, the organization's surplus decreases. We have seen that if the parties, following collusion, have the opportunity to renegotiate their optimal incentive scheme, then they will internalize, at least in part, the externality imposed by the collusive agreement, reducing these inefficiencies considerably.

We have also seen that it is not possible to achieve this result with a collusion-proof and renegotiation-proof optimal incentive scheme. In other words, if collusion is observable to the principal, the optimal incentive scheme may be collusion-proof but it will always be renegotiated in equilibrium. This is because collusion is observable but not verifiable and therefore the optimal incentive scheme cannot be made contingent on the collusion contract the supervisor and the agent will agree upon. Conversely, if collusion is not even observable to the principal, then there exist conditions such that the only continuation equilibrium of the collusion-and-renegotiation subgame is such that both collusion and renegotiation will occur in equilibrium

19. Notice that the results of Propositions 4 and 5 hold if the collusion and renegotiation offers are restricted to a grid and this grid is sufficiently fine.

with positive probability. This is because the only equilibrium, of this continuation game is a mixed-strategy equilibrium in which parties randomize between offering and accepting both collusion and renegotiation. Therefore there exist realizations of the optimal strategies such that both collusion and renegotiation are observed in equilibrium.

We interpret our analysis as saying that the occurrence of both collusion and renegotiation should not be taken as evidence of the inefficiency of an organization.

APPENDIX

This appendix presents the best-response correspondences for the agent, principal, and supervisor in the case in which collusion is nonobservable.

The best response of the agent, p_A^* , can be defined as

$$\begin{aligned}
 p_A^* = \operatorname{argmax}_{p_A} \sum_{b_1} \sum_{b_0} p_A(b_1, b_0) & (p_{Sc}(b_1, b_0) \{ [\theta + e(b_1, b_0)] U(w_1 - b_1) \\
 & + [1 - \theta - e(b_1, b_0)] U(w_0 - b_0) G(e(b_1, b_0)) \} \\
 & + [1 - p_{Sc}(b_1, b_0)] + \{ [\theta + e(0, 0)] U(w_1) \\
 & + [1 - \theta - e(0, 0)] U(w_0) - G(e(0, 0)) \}),
 \end{aligned} \tag{A.1}$$

where $e(b_1, b_0)$ is defined by $G'(e(b_1, b_0)) = U(w_1 - b_1) - U(w_0 - b_0)$.

The best response of the principal, p_P^* , can be defined as

$$\begin{aligned}
 p_P^* = \operatorname{argmin}_{p_P} \sum_{\tilde{s}_1} \sum_{\tilde{s}_0} p_P(\tilde{s}_1, \tilde{s}_0) \sum_{b_1} \sum_{b_0} p_A(b_1, b_0) & [p_{Sc}(b_1, b_0) \\
 \times (p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1) \{ \theta + e(b_1, b_0) \tilde{s}_1 & + [1 - \theta - e(b_1, b_0)] \tilde{s}_0 \} \\
 + [1 - p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1)] \{ [\theta + e(b_1, b_0)] s_1 & \\
 + [1 - \theta - e(b_1, b_0)] s_0 \}) + [1 - p_{Sc}(b_1, b_0)] & (p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0) \\
 \times \{ \theta + e(0, 0) \tilde{s}_1 + [1 - \theta - e(0, 0)] \tilde{s}_0 \} & + (1 - p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0) \\
 \times [(\theta + e(0, 0)] s_1 + [1 - \theta - e(0, 0)] s_0 \}) &].
 \end{aligned} \tag{A.2}$$

Finally, the supervisor's best response (p_{Sc}^*, p_{Sr}^*), for every b_1, b_0, \tilde{s}_1 , and \tilde{s}_0 , can be defined as

$$\begin{aligned}
 p_{Sr}^*(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1) = \operatorname{argmax}_{p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1)} & p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1) \\
 \times \{ [\theta + e(b_1, b_0)] V(\tilde{s}_1 + b_1) & + [1 - \theta - e(b_1, b_0)] V(\tilde{s}_0) + b_0 \}
 \end{aligned}$$

$$\begin{aligned}
& + [1 - p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1)] \{ [\theta + e(b_1, b_0)] \\
& \times V(s_1 + b_1) + [1 - \theta - e(b_1, b_0)] \\
& \times V(s_0 + b_0) \},
\end{aligned} \tag{A.3}$$

$$\begin{aligned}
p_{Sr}^*(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0) = & \operatorname{argmax}_{p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0)} p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0) \\
& \times \{ [\theta + e(0, 0)]V(\tilde{s}_1) + [1 - \theta - e(0, 0)] \\
& \times V(\tilde{s}_0) \} + [1 - p_{Sr}(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0)] \\
& \times \{ [\theta + e(0, 0)]V(s_1) \\
& + [1 - \theta - e(0, 0)]V(s_0) \},
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
p_{Sc}^*(b_1, b_0) = & \operatorname{argmax}_{p_{Sc}(b_1, b_0)} p_{Sc}(b_1, b_0) \sum_{\tilde{s}_1} \sum_{\tilde{s}_0} (p_{Sr}^*(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1) \\
& \times \{ [\theta + e(b_1, b_0)]V(\tilde{s}_1 + b_1) + [1 - \theta - e(b_1, b_0)] \\
& \times V(\tilde{s}_0 + b_0) \} + [1 - p_{Sr}^*(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 1)] \\
& \times \{ [\theta + e(b_1, b_0)]V(s_1 + b_1) + [1 - \theta - e(b_1, b_0)] \\
& \times V(s_0 + b_0) \}) + [1 - p_{Sr}(b_1, b_0)] \\
& \times \sum_{\tilde{s}_1} \sum_{\tilde{s}_0} (p_{Sr}^*(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0) \{ [\theta + e(0, 0)]V(\tilde{s}_1) \\
& + [1 - \theta - e(0, 0)]V(\tilde{s}_0) \} + [1 - p_{Sr}^*(b_1, b_0, \tilde{s}_1, \tilde{s}_0, 0)] \\
& \times \{ [\theta + e(0, 0)]V(s_1) + [1 - \theta - e(0, 0)]V(s_0) \}).
\end{aligned} \tag{A.5}$$

REFERENCES

- Acemoglu, D., 1996, "Career Concerns and Cover-ups: A Dynamic Mode of Collusion," Mimeo, M.I.T.
- Dasgupta, P. and E. Maskin, 1986, "The Existence of Equilibrium in Discontinuous Economic Games, 1: Theory," *Review of Economic Studies*, 53, 1-26.
- Dewatripont, M., 1989, "Renegotiation and Information Revelation over Time: The Case of Optimal Labor Contracts," *Quarterly Journal of Economics*, 104, 589-619.
- Dewatripont, M. and E. Maskin, 1995, "Contractual Contingencies and Renegotiation," *RAND Journal of Economics*, 26, 704-719.
- Felli, L., 1996, "Preventing Collusion through Discretion," Theoretical Economics Discussion Paper TE/96/303, STICERD, London School of Economics.
- Felli, L. and M. Villas-Boas, 1996, "Friendships in Vertical Relations," Theoretical Economics Discussion Paper TE/96/314, STICERD, London School of Economics.

- Fudenberg, D. and J. Tirole, 1990, "Moral Hazard and Renegotiation in Agency Contracts," *Econometrica*, 58, 1279–1319.
- Hermalin, B. and M. Katz, 1990, "Moral Hazard and Verifiability: The Effects of Renegotiation in Agency," *Econometrica*, 59, 1735–1753.
- Holmström, B. and P. Milgrom, 1987, "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 303–328.
- and —, 1989, "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics*, 146, 85–105.
- Itoh, H., 1993, "Coalitions, Incentives, and Risk Sharing," *Journal of Economic Theory*, 60, 410–427.
- Kofman, F. and J. Lawarrée, 1996, "On the Optimality of Allowing Collusion," *Journal of Public Economics*, 61, 383–407.
- Laffont, J.-J. and D. Martimort, 2000, "Mechanism Design with Collusion and Correlation," *Econometrica*, 68, 309–342.
- Ma, C.-A., 1994, "Renegotiation and Optimality in Agency Contracts," *Review of Economic Studies*, 61, 109–129.
- Martimort, D., 1997, "The Life Cycle of Regulatory Agencies: Dynamic Capture and Transaction Costs," *Review of Economic Studies*, 66, 929–947.
- Maskin, E. and J. Tirole, 1992, "The Principal Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 60, 1–42.
- and —, 1999, "Unforeseen Contingencies and Incomplete Contracts," *Review of Economic Studies*, 66, 83–114.
- Matthews, S.A., 1995, "Renegotiation of Sales Contracts," *Econometrica*, 63, 567–590.
- Nash, J., 1950, "Equilibrium Points in n -Person Games," *Proceedings of the National Academy of Sciences*, 36, 48–49.
- Reiche, S., 1999, "Foundations of Incomplete Contracts under Asymmetric Information," Mimeo, London School of Economics.
- Tirole, J., 1986, "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economics and Organizations*, 2, 181–214.
- , 1992, "Collusion and the Theory of Organizations," in J.-J. Laffont, ed., *Advances in Economic Theory: Sixth World Congress*, Vol. II, Cambridge, UK: Cambridge University Press.
- Varian, H.R., 1989, "Monitoring Agents with Other Agents," *Journal of Institutional and Theoretical Economics*, 146, 153–174.