Discussion Paper 1210

Decentralization and Collusion

by

Sandeep Baliga
and
Tomas Sjöström

February 15, 1998

Decentralization and Collusion*

Sandeep Baliga
Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Evanston, Illinois 60208

Tomas Sjöström
Department of Economics
Harvard University
Cambridge, Massachusetts 02138

February 15, 1998

Abstract

We consider a model where agents work in sequence on a project, share information not available to the principal, and can collude. Due to limited liability the Coase theorem does not apply. The distribution of surplus among the agents is therefore an important control variable for the principal, which gives us a theory of how to delegate in an organization subject to moral hazard. The optimal distribution of surplus can always be achieved by delegating in the right way (decentralization) without using "message games" (centralization).

1 Introduction

This paper studies the extent to which decentralization is optimal, and how to delegate in a firm. A principal employs two agents to work on a project whose success or

*The detailed comments of Oliver Hart, Eric Maskin and two referees are gratefully acknowledged. We also thank Luis Corchon, the Associate Editor and seminar audiences at Bristol University, Cambridge University, Cornell University, Harvard University, Harvard Business School, I.N.S.E.A.D., L.S.E., University of Minnesota, Kellogg Graduate School of Management (M.E.D.S.), Norwegian School of Management, N.Y.U., Oxford University, Penn State University, Princeton University, Rochester University, Simon School of Management, University of Michigan, University of Pennsylvania, and University of Wisconsin (Madison). Any remaining errors are our own responsibility.

failure is observable and verifiable. The probability of success depends on the agents' effort levels. The agents work in sequence. Agent I's effort is known to both agents while agent II's effort is his private information. The principal does not observe either effort level. Agent I might be a designer such as an architect (or a research and development department in a firm). He delivers a blueprint to agent II who is a builder (or a production department in the firm). Agent II discovers agent I's effort by examining the blueprint, but agent I does not monitor the actual production process and thus cannot observe agent II's effort. The principal's objective is to minimize the cost of getting both agents to work. There is *limited liability*: no wage can be less than zero. Therefore, the agents must receive some form of rent. Our paper studies how the optimal way of decentralizing minimizes that rent.[1]

The *Revelation Principle* suggests the following centralized, two-tier mechanism: agent II announces agent I's effort level; if he says "shirk", agent I is paid nothing: if he says "work", agent I is paid just enough when the project succeeds to give him the incentive to participate; agent II's wage is independent of his announcement. Agent II is paid sufficiently (the "efficiency wage") if the project succeeds to make it worth his while to work. This mechanism has an equilibrium where agent II announces agent I's effort level truthfully and both agents work, but it is vulnerable to collusion. We study if *delegation/decentralization*, where agents' pay is based on output but not on "messages", is optimal if agents can collude.

We allow agents to sign side-contracts, but we impose the limited liability constraint also on transfers among the agents: when an agent has no money he cannot transfer money to the other agent. Thus, the limited liability implies a form of *non-transferable utility* among the agents. The agents can write binding contracts on variables that are observable to them (cf. Tirole [22]). We assume for most of the paper that agent I's effort, all messages and all wages are observable to the agents. However, the agents cannot contract on agent II's effort level, as it is unobservable to agent I. This together with limited liability restricts the agents' ability to side-contract. Agent II may only be able to make a side payment to agent I after the project has been successful so that agent II has received a high wage. However, anticipating this side payment in the success state, agent II may have an insufficient incentive to work hard (i.e. his income in the success state is below his efficiency wage). As the transfer cannot be made independent of the outcome of the project, it may be impossible for agent II to transfer surplus to agent I without violating his own moral hazard constraint. Collusion then does not necessarily lead to the maximization of the sum of the agents' utilities: the Coase theorem does not hold with limited liability among the agents. By altering the distribution of wages among the agents while keeping the total wage payments fixed, the principal affects the set of feasible side-contracts for the agents. Therefore, the distribution of wages among the

[1]Without the limited liability constraint, the principal could achieve first best by "selling the firm" to agent II, who could then monitor and pay agent I.

agents is an important control variable for the principal.

We find that the principal has no need for message games with public messages. In many cases, the optimal contract can be implemented by using a linear organization as follows. The principal hires a single agent, *the General Contractor*, who is responsible for both design and construction of the project, and who is paid a sum of money when the project is successfully completed. The General Contractor is responsible for contracting with, and paying, the other agent. In the building profession this is called the *design/build* process ([2]), and either the builder or architect may be the General Contractor. For example, Kenneth Parry Associates, an architecture firm, was the general contractor in the construction of a duplex. Design Concepts, a construction management firm, was responsible for an elderly housing project ([5]). In other cases, the optimal organization is triangular and the principal pays both agents and lets them side contract with each other. This method of organizing construction is also used in the building trade ([5]).

If agent I (the architect) is the General Contractor, he must pay agent II (the builder) an efficiency wage to get him to work, because he cannot monitor the builder's effort. On the other hand, if the builder is the General Contractor, he can monitor and pay the architect according to the quality of the blueprint. This suggests that the builder should be the General Contractor so that his superior information about the architect's effort can be used to provide good incentives for the architect. In fact, making the builder the General Contractor is the (uniquely) best way to delegate under some parameter values, and in particular if the architect's effort is relatively cheap to induce for the builder. But notice that the principal will not pay the builder the full amount of what a successful project is worth to her: if she did, she would make no profit. Thus, the builder does not internalize the full value of the project, and is tempted to save money by not paying the architect to work hard. This is a bigger problem the more costly is the architect's effort, and the less important it is for the success of the project. For in this case, to give the builder an incentive to sign a contract with the architect which induces the latter to work hard, the principal must promise the builder a very substantial part of the profit of the project. Then it will not be optimal to make the builder the General Contractor.

The architect should be the General Contractor if his own effort is rather costly to induce compared to the builder's effort, and his effort and the builder's effort are complements in the production function. In this case, the most important problem for the principal is to make sure the architect works. However, this problem is mitigated if the architect is the General Contractor, because working hard is then a way for him to transfer rent from the builder to himself. By producing a very good blue print which makes the builder very efficient, the architect relaxes the builder's moral hazard constraint and reduces the builder's efficiency wage. The optimal contract in this case has the property that the *sum* of the agents' payoffs would be maximized if the architect shirked, but the architect works hard anyway to gain a more favorable *distribution* of the payoffs. Thus, by delegating to the right agent, the principal makes sure both agents work *even if this does not maximize the agents' joint surplus.*

Maskin and Tirole [13] argue that if message games are compatible with the assumptions made in the models, then they should not be ruled out *a priori.* Otherwise, the optimality of delegation (Aghion and Tirole [1]) or property rights (Hart [6]) is inconsistent with the possibility that centralized contracts may strictly dominate such organizational modes. In our setup, one may also argue that contracts where messages and wages are *publicly observable* are not the most general contracts possible. Therefore, we consider more sophisticated contracts with *secret* messages and wages in Section 6. Let the principal pay secret randomized wages, and let the wage be zero with some probability. This makes it difficult for an agent to credibly promise to make a side payment, for ex post he can pretend he never got paid. In this case collusion is difficult and message games (centralization) valuable. But a court might find it difficult to enforce randomized wages. An alternative way to eliminate collusion is to keep *messages* secret. If it cannot be verified whether or not an agent "snitched", collusion is impossible. However, secret messages do not work if the principal is a strategic player who can collude, for then he will always convince the agent to send the message which minimizes the wages paid to the other agent. Therefore, our results are robust to the consideration of this more general class of mechanisms if we take the principal's commitment problem and incentives to collude into account.

The literature on multi-agent incentive schemes shows that a non-individualistic situation with monitoring and collusion is at least as good and sometimes strictly better for the principal than a purely individualistic scheme (Itoh [8], Holmström and Milgrom [7], Macho-Stadler and Perez-Castrillo [12], Ramakrishnan and Thakor [20], Tirole [22]). While it is true also in our model that a purely individualistic scheme (which rules out both monitoring and side-contracting) would not be optimal, our focus is different: we obtain a theory of the optimal *distribution* of wages, and predict *how* the principal should delegate. In addition, this literature in general does not allow centralized schemes with monitoring and message games. An exception is Itoh [8] who shows that when there is transferable utility and the agents know the entire effort profile, the principal gains no advantage from a message game: agents anyway always contract on the messages and effort levels that maximize the sum of expected utilities (the Coase theorem holds). [2]

A different strand of the literature looks at the impact of collusion in a principal-supervisor-agent setting (Tirole [21] and [22]). This literature takes the structure of the hierarchy as given, and analyzes how the optimal incentive scheme is modified by the possibility of side-contracting. There is a recent literature that compares decentralized and centralized incentive schemes without collusion in adverse selection models (Melumad, Mookherjee and Reichelstein [15] and [16], Mookherjee and Re-

[2] The idea that collusion can destroy the usefulness of message games is also explored by Baliga [3].

ichelstein [18] and [19], and McAfee and McMillan [14]). These authors have obtained conditions under which decentralization can replicate the *second-best* centralized contract (without collusion).[3] In our model there is also a condition (*Case A*) which implies that decentralization can mimic the second-best contract (without collusion). When this condition is *not* satisfied, decentralization cannot achieve what centralization could achieve in the absence of collusion, but it does as well as centralization if centralized schemes (message games) are subject to collusion. Finally, two recent papers look at the advantages of decentralized versus centralized contracts in the presence of collusion. Macho-Stadler and Perez-Castrillo [11] study the negative effects of different coalitional structures, including those where the principal colludes with one of the agents, in a model of moral hazard. Laffont and Martimort [9] look at an adverse selection model where agents know only their own cost of production and collude under asymmetric information. They show that if in any centralized scheme the principal is restricted to "anonymous" contracts, delegation performs *strictly better* than centralization.[4] The focus of the present paper is somewhat different from these papers, namely, *how* the principal should delegate.

2 The model

Two agents work in sequence. Agent I, who is in charge of design or research and development, delivers a blueprint to agent II, who does the actual production. (Other interpretations are possible, such as that of a production-line where agents work sequentially.) The effort put in by agent i, e_i, is either zero or one, and the cost of one unit of effort is c_i. Let $e = (e_1, e_2)$. When agent I delivers the blueprint to agent II, agent II learns agent I's effort e_1 by inspecting the blueprint. As the principal does not observe e_1, agent II is better informed than the principal. Neither agent I nor the principal can observe e_2.

After the agents have worked, the project is revealed to be either a success or failure. This outcome is public information. The probability of a success is $p_{e_1 e_2}$, if agent i's effort is $e_i \in \{0, 1\}$. We assume $0 < p_{00} < p_{11} < 1$ and

$$p_{00} < p_{e_1 e_2} < p_{11} \quad \text{when } e_1 \neq e_2$$

Both the agents and the principal know the parameters $(c_1, c_2, p_{00}, p_{01}, p_{10}, p_{11})$.

The wage cannot directly depend on effort, as it is not observed by the principal. However, it can depend on the outcome of the project, and on messages sent by the

[3] Van Zandt [23] discusses how the equivalence of decentralization and centralization in the adverse selection models depends on whether agents' participation constraints have to be satisfied ex ante or ex post.

[4] Baliga and Sjöström [4] consider an adverse selection model where agents both know the parameters of the model but the principal does not. In that model, when the optimal distribution of surplus is state dependent, centralization can dominate delegation even though agents can collude.

agents in some game designed by the principal. If agent i consumes w_i units of money, and his effort level is e_i, then his payoff is $w_i - e_i c_i$. Agents have zero wealth. All wages must be non-negative due to the limited liability of the agents. Each agent must be offered an expected payoff of at least zero in order to participate.

We assume the principal wants both agents to work hard, $e = (1, 1)$, as the project is sufficiently valuable to her. The issue is at what cost this full effort profile can be achieved. If the effort of both agents were observable to the principal, the "first best" contract would require both agents to work, and would pay agent i the wage c_i. The cost to the principal would be $c_1 + c_2$. However, agent II's effort is unobservable to everybody except himself. Therefore, for agent II to work, a moral hazard constraint must be satisfied. The "second best" contract pays agent I the expected wage c_1, and satisfies agent II's moral hazard constraint at the lowest possible cost. Let w_2 be the wage for agent II if the outcome is a success (it is clearly optimal to pay zero when the project fails). The moral hazard constraint for agent II is $p_{11}w_2 - c_2 \geq p_{10}w_2$. The expected cost to the principal from the second best contract is therefore

$$c_1 + p_{11}w_2 = c_1 + \frac{p_{11}c_2}{p_{11} - p_{10}} > c_1 + c_2 \tag{1}$$

The extra cost to the principal, $c_2 p_{10}/(p_{11} - p_{10})$, is a rent earned by agent II.

If the agents cannot collude, full effort can be implemented at the second best cost by asking agent II to report agent I's effort. If the project is unsuccessful, both agents get zero. If it is successful, agent II gets $w_2 = c_2/(p_{11} - p_{10})$, and I gets $w_1 = c_1/p_{11}$ if II has announced that I's effort was high, otherwise agent I gets zero. This mechanism has an equilibrium where agent II truthfully reports agent I's effort level, and both agents work hard.[5] Moreover, it is clearly *necessary* to include a message game in order to implement full effort at the second best cost, as without messages a moral hazard constraint would have to hold also for agent I, implying a transfer of rent to agent I. But the message game is vulnerable to collusion. If agent I shirks, he is willing to pay a bribe to II in order for him not to "snitch". Since agent II's wage does not depend on his message, he is willing to accept the bribe.

It may be useful to compare our model to a situation where the agents work in complete isolation and neither agent can monitor the other's effort level. In that case, a moral hazard constraint must be satisfied also for agent I, and the principal could not implement full effort at a cost lower than

$$p_{11}\left(\frac{c_1}{p_{11} - p_{01}} + \frac{c_2}{p_{11} - p_{10}}\right)$$

This is greater than (1) and indeed greater than the cost our principal can achieve under the assumption that agent II can monitor agent I and the agents can collude.

[5] In this paper we do not consider the issue of multiple equilibria, but here full effort can be *uniquely* implemented by a mechanism which is more complicated (cf. Ma [10]).

Thus, as Itoh [8] has shown,[6] a non-individualistic context with monitoring and collusion is in general better for the principal than a purely individualistic scheme without monitoring.

3 Collusion and message games

Assume both the principal and the agents know all the parameters of the model. The principal designs a mechanism to elicit information about e_1. The sequence of events is the following.

0. Each agent i sends a message $m_i^0 \in M_i^0$ to the principal, where M_i^0 is the message space. [7]

1. Agent I works ($e_1 = 1$) or shirks ($e_1 = 0$), and the effort is observed by agent II.

2. Each agent i sends a message $m_i \in M_i'$ to the principal, where M_i' is the message space.

3. Agent II works ($e_2 = 1$) or shirks ($e_2 = 0$).

4. The success or failure of the project becomes public information. Conditional on this outcome and on the messages, wages are paid.

All messages (and wages) are announced publicly. The wages can depend on the outcome of the project and on the messages $m = (m_1^0, m_1', m_2^0, m_2')$. Agent i's wage is $w_i^s(m)$ if the project is successful and $w_i^f(m)$ otherwise. Let M denote the message spaces, $M = M_1^0 \times M_2^0 \times M_1' \times M_2'$, and let w denote the wage functions,

$$w = \left(w_1^f(\cdot), w_2^f(\cdot), w_1^s(\cdot), w_2^s(\cdot)\right)$$

The pair (M, w) is a *mechanism*. The mechanism together with the rules given by the stages above induces a multi-stage game with observed actions, denoted $\Gamma(M, w)$. Let $E^{\Gamma(M,w)}$ be the set of subgame perfect Nash equilibrium payoffs of the extensive form game $\Gamma(M, w)$.

Before playing the game, i.e. before time zero, the agents can sign a side contract which is assumed to be enforceable.[8] If some agent refuses to sign a side contract,

[6] See also Holmström and Milgrom [7], Macho-Stadler and Perez-Castrillo [12] and Ramakrishnan and Thakor [20].

[7] The reader may be concerned that in the most general mechanism, messages could be sent sequentially rather than simultaneously. But such sequential communication can be replaced by its normal form. Since we are not concerned about multiple equilibria, this will not be any worse for the principal.

[8] Implicitly, the agents have access to a mechanism for punishing deviators, and the punishment is severe enough to make deviations from a side contract unprofitable. Our results go through if we treat the agents and the principal symmetrically by assuming the punishment has to respect the limited liability, i.e. we could assume the worst the agents can do to each other is to destroy each other's wages.

we suppose they proceed to play a subgame perfect equilibrium of $\Gamma(M, w)$. A side contract $c = (e_1, m, t)$ specifies: (i) agent I's effort level e_1, (ii) the list m of all messages to be sent to the principal, and (iii) a pair of transfers $t = (t^s, t^f)$, where t^s (t^f) is the sum of money to be paid from agent I to agent II if the project is a success (failure). The side contract cannot specify agent II's unobservable effort level as such a contract would not be enforceable.

To be feasible, the side transfers must satisfy:

$$-w_2^s(m) \leq t^s \leq w_1^s(m) \tag{2}$$

$$-w_2^f(m) \leq t^f \leq w_1^f(m) \tag{3}$$

If a side contract c is signed, $\Gamma(M, w)$ is replaced by the following game: at stages 0-2, each agent must behave as specified by c (this simplification is justified by the assumption that violating c is prohibitively costly). At stage 3, agent II decides to work or shirk. At stage 4 wages and transfers (as specified by c) are paid. If the side contract specifies $e_1 \in \{0, 1\}$ and messages m, then player II will work iff the increase in his expected wages cover the cost of his effort, i.e. iff

$$(p_{e_1 1} - p_{e_1 0}) \left((w_2^s(m) + t^s) - \left(w_2^f(m) + t^f \right) \right) \geq c_2 \tag{4}$$

If equality holds in (4) then agent II is indifferent between working or shirking. In this case we assume that he works, and a similar assumption is made for agent I. Thus from now on, $E^{\Gamma(M,w)}$ is the set of subgame perfect Nash equilibrium payoffs where ties are broken in favor of working.

Consider a feasible side contract $c = (e_1, m, t)$. Since it specifies all actions except e_2, and $e_2 = 1$ iff (4) holds, each player i can compute his expected payoff $\pi_i(c)$ from signing this contract c. Then c is an *equilibrium side contract* for (M, w) iff it satisfies: (E1) For each i, there is $x^i = (x_1^i, x_2^i) \in E^{\Gamma(M,w)}$ such that $\pi_i(c) \geq x_i^i$.[9] (E2) There is no other feasible side contract c' satisfying $\pi_1(c') > \pi_1(c)$ and $\pi_2(c') > \pi_2(c)$. Thus, each player i should be better off by signing the contract c than by refusing to sign and instead playing some subgame perfect equilibrium which results in payoffs x^i, and there is no other feasible side contract c' that could be signed which strongly Pareto dominates c.[10] (Our results would also hold for weak Pareto domination, but the present formulation is more convenient).

[9] An alternative, for our purposes equivalent, formulation is that the side contract should dominate some Pareto efficient subgame perfect equilibrium.

[10] In the "Nash demand game", any efficient division of the surplus is a strong Nash equilibrium. The reader may find it useful to imagine that the initial collusion stage to choose from the set of equilibrium side contracts is such a game.

A feasible side contract $c = (e_1, m, t)$ *implements full effort at the cost* C iff $e_1 = 1$, (4) holds, and

$$p_{11}\left(w_1^s(m) + w_2^s(m)\right) + (1 - p_{11})\left(w_1^f(m) + w_2^f(m)\right) = C$$

It could happen that a subgame perfect equilibrium of $\Gamma(M, w)$ is not Pareto dominated by any feasible side contract, so the agents have no (strict) incentive to collude. But even in this case, the agents can just as well sign the side contract that tells them to play according to this subgame perfect equilibrium. This will be an equilibrium side contract, since (E1) is trivially satisfied and (E2) holds by assumption. So without loss of generality, from now on we assume that a side contract is *always* signed before stage zero. A mechanism (M, w) *implements full effort at the cost* C, if there exists *some* equilibrium side contract c for (M, w) which implements full effort at the cost C.[11]

4 Necessary conditions for implementation

In this section we state necessary conditions for implementation by any mechanism.

Proposition 1 *Full effort cannot be implemented by any mechanism at a cost lower than*

$$c_1 + p_{11}\frac{c_2}{p_{11} - p_{10}} \tag{5}$$

Proof. As agent I must be compensated for his effort and player II's moral hazard constraint requires that he gets at least $c_2/(p_{11} - p_{10})$ if the project is a success, the proposition follows. ■

Suppose full effort is implemented, and let w_i denote the wage agent i receives in equilibrium when the project is successful. By Proposition 1, $p_{11}(w_1 + w_2)$ must exceed (5), i.e.

$$w_1 + w_2 \geq \frac{c_1}{p_{11}} + \frac{c_2}{p_{11} - p_{10}} \tag{6}$$

As the agents can collude, there are some additional considerations. Although there is limited liability, it turns out to be useful to look at the *sum* of the payoffs of

[11] In a more general model, we could also allow the agents to sign a side contract at a later stage or, if a side contract already exists, renegotiate it. In this case we would define equilibrium side contracts recursively as follows. At the last time where collusion can occur, equilibrium side contracts are defined analogously to what was done above. Equilibrium side contracts in earlier periods are defined recursively as feasible Pareto-undominated side contracts which give the agents no lower payoff than the worst they could expect by not signing, *taking later negotiations into account*. However, for any equilibrium of this more general model it will again be possible to duplicate the equilibrium path by signing a comprehensive side contract before time zero. Therefore, our model is essentially equivalent to the model with more general collusion/renegotiation possibilities.

the agents. If full effort is implemented, then this sum is $p_{11}(w_1 + w_2) - c_1 - c_2$ in equilibrium. If *both* agents shirk, *without changing their messages*, the sum of their payoffs would be $p_{00}(w_1 + w_2)$. Thus, the sum is greater when both work than when both shirk if and only if

$$w_1 + w_2 \geq \frac{c_1 + c_2}{p_{11} - p_{00}} \tag{7}$$

Our next result shows that this "team moral hazard constraint" must be satisfied in equilibrium.

Proposition 2 *Full effort cannot be implemented by any mechanism at a cost lower than*

$$p_{11}\frac{c_1 + c_2}{p_{11} - p_{00}}$$

Proof. In the appendix.

If agent II works hard but agent I shirks, the sum of their payoffs would be $p_{01}(w_1 + w_2) - c_2$. Thus, this sum is greater when both work than when only agent II works if and only if

$$w_1 + w_2 \geq \frac{c_1}{p_{11} - p_{01}} \tag{8}$$

However, (8) may be violated by the optimal contract. Why would agent I work if it does not maximize the total surplus for the agents ? Even though the total surplus is greater when agent I shirks, more of the surplus may have to be given to agent II to induce him to work with a low quality blueprint. Agent II's moral hazard constraint implies that he needs a rent to work hard, and agent I may be able to reduce this rent by working hard himself. If agent I works hard, then agent II needs an "efficiency wage" equal to $c_2/(p_{11} - p_{10})$ in order to work hard, which will give him a rent equal to $p_{10}c_2/(p_{11} - p_{10})$. If agent I shirks then agent II's efficiency wage is $c_2/(p_{01} - p_{00})$ which implies a rent equal to $p_{00}c_2/(p_{01} - p_{00})$. There are two possibilities. If

$$\frac{p_{00}}{p_{01} - p_{00}} > \frac{p_{10}}{p_{11} - p_{10}} \tag{9}$$

then agent II's rent is reduced if agent I works hard and produces a good blue print. In this case, the fact that agent I's effort relaxes agent II's moral hazard constraint can make agent I work hard even though it reduces the total surplus (i.e. even if (8) is violated). This is the content of Proposition 4.

The second possibility is that (9) is violated. Then it costs less to motivate agent II to exert effort when agent I shirks. In this case, if (8) is violated then there is not only more surplus to share if agent I shirks but this also reduces the rent agent II needs to work. Then agent I would certainly shirk. Therefore, if (9) is violated then (8) must be satisfied. This is the content of Proposition 3.

We have provided an intuitive motivation for why (8) can be violated if when (9) holds. However, in view of Propositions 1 and 2, clearly (8) can only be violated if it is not implied by either (6) or (7), i.e. if (8) is the most difficult constraint to satisfy. This is true if

$$\frac{c_1}{p_{11}-p_{01}} > \max\left\{\frac{c_2}{p_{11}-p_{10}}+\frac{c_1}{p_{11}}, \frac{c_1+c_2}{p_{11}-p_{00}}\right\} \tag{10}$$

For this case we have two results.

Proposition 3 *Suppose (10) holds but (9) does not hold. Then full effort cannot be implemented by any mechanism at a cost lower than*

$$\frac{p_{11}c_1}{p_{11}-p_{01}} \tag{11}$$

Proof. In the appendix.

Proposition 4 *Suppose (9) and (10) hold. Then, full effort cannot be implemented by any mechanism at a cost lower than*

$$\frac{p_{11}}{p_{11}-p_{01}}\left\{c_1-\left[\frac{p_{00}c_2}{p_{01}-p_{00}}-\frac{p_{10}c_2}{p_{11}-p_{10}}\right]\right\} \tag{12}$$

Proof. In the appendix.

The discussion which preceded (9) shows that the positive expression in square brackets in (12) is precisely the reduction in agent II's rent which can be achieved if agent I works hard. It is as if agent I's effective cost of effort is reduced by the amount of rent he can transfer from agent II to himself by producing a good blueprint. This formalizes the intuitive discussion which preceded the propositions.

These results depend on the assumption of limited liability. With unlimited liability the principal can implement full effort at "first best" cost c_1+c_2 by "selling the firm" to agent II, who in effect becomes the new principal who can monitor and side-contract with agent I. This solves the moral hazard problem completely. Formally, consider the following mechanism. Let F be the value to the principal of a failed project, and let $F+\Delta$ be the value of a success. Our assumption that the principal wants both agents to work at the first best implies

$$p_{11}\Delta - c_1 - c_2 > \max\{p_{10}\Delta - c_1, p_{01}\Delta - c_2, p_{00}\Delta\} \tag{13}$$

Pay agent 2 $w_2^f = c_1 + c_2 - p_{11}\Delta < 0$ if the project fails and $w_2^s = w_2^f + \Delta > 0$ if the project succeeds; never pay agent 1 anything. Then agent II suffers the full cost of a failure, and given (13) will certainly have an incentive to side contract with agent I, to monitor him and make him work. He pays agent I the expected wage c_1 and works himself (he is just willing to participate). The principal's profit is first best: $(1-p_{11})F + p_{11}(\Delta + F) - (c_1+c_2)$.

5 Simple contracts

In this section we will show that the lower bounds on the cost of implementing full effort derived in Section 4 can be attained using simple contracts without messages.[12] By definition, a mechanism is a *simple contract* if the principal pays agent i a wage w_i if the project is successful, pays nothing if the project fails, and *there are no messages*: $M_1^0 = M_1' = M_2^0 = M_2' = \emptyset$. The simple contract is then defined by the success wages (w_1, w_2). The extensive form game induced by the simple contract will be denoted $\Gamma(w_1, w_2)$ (or just Γ if there is no chance of confusion). Due to our tie-breaking rule, E^Γ is in fact a singleton. A side contract c specifies I's effort level e_1 and a transfer t to be paid from I to II if the project is a success, but there are no messages, so we write $c = (e_1, t)$.

The results of Section 4 suggest that it is useful to consider various cases separately. As agent II can observe the effort of agent I, one might expect that the optimal simple contract involves delegating the task of monitoring and paying agent I to agent II. We will show that this intuition holds if the right hand side of (6) is greater than the right hand side of (7) and (8), but not necessarily otherwise.

5.1 Case A

This is the case

$$\frac{c_2}{p_{11} - p_{10}} + \frac{c_1}{p_{11}} \geq \max\left\{\frac{c_1 + c_2}{p_{11} - p_{00}}, \frac{c_1}{p_{11} - p_{01}}\right\} \tag{14}$$

We know that full effort cannot be implemented by *any* mechanism at a cost lower than (5). We will show that this lower bound can be attained by a simple contract. Moreover, the *unique* way to implement full effort at this cost is to set $w_2 = c_2/(p_{11} - p_{10}) + c_1/p_{11}$ and $w_1 = 0$. That is, agent II should be the General Contractor.

Proposition 5 *Suppose (14) holds. Full effort can be implemented by a simple contract at the cost*

$$p_{11}\left(\frac{c_2}{p_{11} - p_{10}} + \frac{c_1}{p_{11}}\right)$$

It is necessary that $w_1 = 0$ (i.e. agent II must *be the General Contractor).*

Proof. In the appendix.

In the optimal simple contract for case A, the sum of the agents payoffs is maximized when both work. It is an equilibrium side contract for the agents to agree that agent I should work, and agent II pays c_1/p_{11} to agent I if the project is a success.

[12] The "option mechanism" used to prove Propositions 1- 4 does use messages.

This contract leaves agent I with a zero surplus, and agent II keeps enough money in the good state to precisely satisfy the moral hazard constraint. If the principal were to promise a positive wage to agent I, agent I would never agree to a side contract which gives him zero surplus. So $w_1 = 0$ is necessary.

Case A occurs when $p_{11} - p_{10}$ is small. That is, when if agent I has worked hard, the "blueprint" is so good that agent II's effort does not increase the probability of success by much. In this case, the production department's moral hazard constraint is difficult to satisfy. If the research department were given part of the money ($w_1 > 0$), it would not transfer any of it to the production department; it prefers to have the production department shirk as success is likely anyway and getting the production department to work is so costly. So the principal must pay enough so that the sum of expected payoffs is maximized when both agents work, give the wage packet to the production department, and let it monitor and pay the research department according to the quality of the blueprint. The production department earns a rent as its effort is unobservable. Holmström and Milgrom [7], Itoh [8] and Ramakrishnan and Thakor [20] also suggest that delegation to the agent with superior information (agent II) is optimal. However, in our model there are other cases too.

5.2 Case B

This is the case

$$\frac{c_1 + c_2}{p_{11} - p_{00}} \geq \max\left\{\frac{c_2}{p_{11} - p_{10}} + \frac{c_1}{p_{11}}, \frac{c_1}{p_{11} - p_{01}}\right\} \tag{15}$$

As in case A, the sum of the agents' payoffs is maximized when both work, and if agent II is the General Contractor then he will sign a contract with agent I that makes both agents work. Thus, delegating to agent II is optimal ($w_1 = 0$). However, in contrast to Case A, the most difficult constraint is the "team moral hazard constraint" (7), and the main concern is for $w_1 + w_2$ to be sufficiently large that it pays for both agents to work rather than shirk. It is therefore possible to set $w_1 > 0$, as long as agent II keeps enough of the surplus that he is willing to work. If agent II gets an insufficient share of the surplus, he would need a transfer from agent I in order to work. As agent I cannot observe agent II's effort, the transfer has to be big enough to satisfy agent II's moral hazard constraint. But this gives agent II a rent, and agent I might then prefer to have agent II shirk rather than transferring this rent. Thus, although $w_1 = 0$ is not the only possibility, w_1 cannot be too big.

We know from Section 4 that full effort cannot be implemented by *any* mechanism at a cost lower than

$$p_{11}\frac{c_1 + c_2}{p_{11} - p_{00}}$$

Full effort can be implemented at this cost by making agent II the General Contractor.

Proposition 6 *Suppose (15) holds. Full effort can be implemented by a simple contract at the cost*

$$p_{11}\frac{c_1 + c_2}{p_{11} - p_{00}}$$

It is possible (but not necessary) to set $w_1 = 0$ (i.e. agent II can be the General Contractor).

Proof. In the appendix.

5.3 Case C

This is the case where the most difficult constraint to satisfy is that agent I should not shirk, i.e. (10) holds. Agent I's effort costs him c_1, and increases the probability of success by $p_{11} - p_{01}$ (assuming II works). When (10) holds, it is relatively costly for agent I to improve the probability of success, and hence it is tempting for the agents to allow I to shirk.

First suppose that (9) holds so the configuration of parameters corresponds to Proposition 4. When (10) and (9) hold, it turns out to be optimal for the principal to pay wages in such a way that (8) is violated, and the sum of the agents' expected payoffs when both work is *smaller* than the sum of the expected payoffs *when agent I shirks and agent II works.*[13] That is:

$$p_{11}(w_1 + w_2) - c_1 - c_2 < p_{01}(w_1 + w_2) - c_2 \tag{16}$$

As explained in Section 4, for this configuration of parameters the principal can exploit the fact that agent I is willing to work hard to transfer rent from agent II to himself. The way to do this is by *not* making agent II the General Contractor. Suppose agent I is given all of the wages, i.e. $w_2 = 0$. For agent II to work, he must receive a transfer from agent I if the project is successful. Now, (9) implies that the expected transfer that must be given to agent II to make him work is greater when agent I has shirked than when agent I has worked hard: agent II's marginal productivity is reduced by working with a low quality blueprint. Consider the following two side contracts. Under *contract A* both agents work, and agent I transfers $t = c_2/(p_{11} - p_{10})$ to agent II if the project succeeds. This is the smallest transfer that will make agent II work (his efficiency wage), and it will give him a rent equal to $p_{10}c_2/(p_{11} - p_{10})$. Under *contract B* agent I shirks, and pays agent II $t = c_2/(p_{01} - p_{00})$ if the project succeeds. This is the smallest transfer that will make agent II work

[13] Notice that if the right hand side of (8) is smaller than the right hand side expressions of (6) and (7), then (8) must in fact hold because we have shown that (6) and (7) must hold. The interesting case is therefore case C where $c_1/(p_{11} - p_{01})$ is big.

given that I has shirked, and it will give agent II a rent equal to $p_{00}c_2/(p_{01}-p_{00})$. As (9) holds, agent II prefers (B) with the greater rent, but agent I prefers (A) if

$$p_{11}(w_1 - c_2/(p_{11}-p_{10})) - c_1 \geq p_{01}(w_1 - c_2/(p_{01}-p_{00}))$$

or equivalently

$$w_1 \geq \frac{c_1}{p_{11}-p_{01}} + \frac{c_2}{p_{11}-p_{01}}[\frac{p_{10}}{p_{11}-p_{10}} - \frac{p_{00}}{p_{01}-p_{00}}]. \tag{17}$$

Moreover, if (17) holds there is no way of making (B) more attractive for agent I, as any reduction in the transfer will cause agent II to shirk. As (B) is joint surplus maximizing, if agent II could make an ex ante lump sum transfer to agent I in exchange for incentive contract (B), both could be made better off, but such transfers are ruled out by the limited liability (agent II cannot pay in the failure state). If agent I is the General Contractor, he will prefer contract (A).

The crucial issue is the degree of complementarity between the two agents' inputs. If (9) holds then agent I's effort makes agent II more productive. If agent I is the General Contractor then he works hard to relax agent II's moral hazard constraint (even if $e = (0, 1)$ implies a higher *total* surplus). Thus, while in case A, the principal should not pay agent I to assure implementation of an effort profile $(1, 1)$ that maximizes the sum of the agents' expected payoffs, now she must pay agent I to *prevent* implementation of an effort profile $(0, 1)$ that maximizes the sum of expected payoffs. Indeed, if agent II receives a large share of the wage packet and $e = (0, 1)$ maximizes the sum of the payoffs, then the outcome will be $e = (0, 1)$ because agent II will never pay agent I to work in this case. So it is *not* optimal to make agent II a General Contractor and give him the whole wage packet.

If (9) does not hold then of course the above argument does not go through and $w_1 = 0$ is again optimal.

Define

$$w^* \equiv \max\{\frac{c_1}{p_{11}-p_{01}} + \frac{c_2}{p_{11}-p_{01}}[\frac{p_{10}}{p_{11}-p_{10}} - \frac{p_{00}}{p_{01}-p_{00}}], \frac{c_1}{p_{11}} + \frac{c_2}{p_{11}-p_{10}}, \frac{c_1+c_2}{p_{11}-p_{00}}\} \tag{18}$$

If (10) and (9) hold then

$$w^* < c_1/(p_{11}-p_{01}) \tag{19}$$

It follows from Propositions 1, 2 and 4 that full effort cannot be implemented by any mechanism at a cost below $p_{11}w^*$.

Proposition 7 *Suppose (10) and (9) hold. Then, full effort can be implemented by a simple contract at the cost* $p_{11}w^*$. *It is necessary that* $w_2 < w^*$ *(i.e. agent II cannot be General Contractor).*

Proof. In the appendix.

Proposition 8 *Suppose (10) holds but (9) does not hold. Then full effort can be implemented by a simple contract at the cost*

$$p_{11}\frac{c_1}{p_{11}-p_{01}}$$

It is possible (but not necessary) to set $w_1 = 0$ (i.e. agent II can be General Contractor).

Proof. In the appendix.

To summarize the discussion of cases A, B and C, it is optimal (within the class of *all* mechanisms) to make agent II the General Contractor, *except* when (10) and (9) hold. The only remaining issue is, if (10) and (9) hold, is it optimal to make *agent I* the General Contractor, or must the principal pay *both* agents? Suppose we set $w_1 = w^*$ and $w_2 = 0$. Then as shown in the proof of Proposition 7, *if* agent I wants agent II to work, he prefers to also work himself to minimize agent II's rent. Does agent I want to induce agent II to work ? By working and paying agent II his efficiency wage $c_2/(p_{11} - p_{10})$, agent I's payoff is

$$p_{11}\left(w^* - \frac{c_2}{p_{11}-p_{10}}\right) - c_1$$

By working alone and paying nothing to agent II, agent I gets $p_{10}w^* - c_1$, and if both agents shirk, agent I gets $p_{00}w^*$. This implies that agent I is willing to pay agent II to work if and only if

$$w^* \geq \max\left\{\frac{p_{11}c_2}{(p_{11}-p_{10})^2}\,;\ \frac{c_1}{p_{11}-p_{00}} + \frac{p_{11}c_2}{(p_{11}-p_{10})(p_{11}-p_{00})}\right\} \tag{20}$$

Thus, if (10) and (9) hold, then agent I should be the General Contractor only if (20) holds,[14] but *otherwise the agents must split*[15] *the wage packet* ($w_1 > 0$ and $w_2 > 0$) and a triangular organization is optimal for the principal. Notice that (20) holds if and only if $p_{10}w^* - c_1$ and $p_{00}w^*$ are both low, which happens if p_{10} and p_{00} are both low, i.e. if agent II is vital for the success of the project. Then agent I will have the incentive to pay agent II to work, and he will work hard himself to reduce the efficiency wage.

[14] It is easy to check that neither (20) nor its negation are implied by (10) and (9) .

[15] *The exact way in which it can be done is shown in the proof of Proposition 7.*

6 Eliminating Collusion by Limiting Public Information

In this section we consider two ways of eliminating collusion: *secret (randomized) wages* and *secret messages*. Once collusion is made impossible, message games can be used to implement full effort at the second best cost.

First, suppose the principal can design a message game where at stage 4 each agent can observe his own wage but not the other agent's wage. Otherwise the situation is the same as before. Section 4 established the minimum cost of getting full effort for different parameters with public wage payments, and this cost was, in general, strictly higher than the second best cost. We now show how secret wages reduce this cost.

We need to be specific about how side contracts are enforced. We suppose the agents have access to a third party, called a "union", which punishes deviations from a side contract. The union can inspect a collusive agreement, it observes side-payments, messages, and agent I's effort, and will punish an agent who cheats. But by assumption, it cannot monitor secret wage payments. Let the cost of being punished by the union be $h \geq 0$, where possibly $h = +\infty$. Suppose that the principal pays *randomized* secret wages. Neither the agents nor the union can observe the randomizations. To make sure that the principal uses the right probabilities and does not cheat, we can suppose the principal keeps a record of wage payments and randomizations. These are not made available to the agents or to the union. However, an impartial "judge" can inspect the documents and make sure that the principal does not cheat.

If the principal actually pays zero with some probability, the agents cannot make credible promises of monetary transfers. An agent can always claim to have received a zero wage and refuse to pay, and it will be impossible for the union to know if he is lying. If the union only punishes an agent who it knows has surely broken a side contract, this clearly renders side payments (and hence collusion) impossible. In fact, even if the union would be willing to punish an agent on the mere suspicion that he may have been cheating his co-worker, collusion can still be ruled out. This is shown formally in Appendix 2. Without collusion, message games are valuable and full effort can be obtained the second best cost $c_1 + p_{11}c_2/(p_{11} - p_{10})$. We conclude that the combination of public messages and secret random wages leads to an improvement compared to the results derived in Section 4.

Secret wages are common in the real world, but a judge might find enforcing a *randomized* scheme problematic. Suppose the principal can only pay secret *non-randomized* wages, but also *messages* can be made secret. For example, academic tenure decisions may involve senior faculty members sending secret messages to the dean. The dean will never reveal the content of messages to a junior professor. In the case of a law suit involving a tenure decision, a judge can decide the case after inspecting the relevant documents, which are kept on file by the dean, but the judge will never tell a junior professor the precise content of the secret messages.

Secret messages destroy collusion opportunities among the agents, for the union will not be able to verify if an agent has sent the right message. However, secret messages open the possibility for the principal to collude with agent II. Suppose the principal and agent II have access to a third party which will enforce collusion, similar to the previously described "union". Even if messages are secret, collusion between the principal and agent II cannot be ruled out. First, the principal has no budget restriction, so he cannot (as agent I above) refuse to pay a side payment for lack of money. Second, the principal can always store the secret messages and, if necessary, show them to the third party. Therefore, it will be possible for a third party to enforce collusion between the principal and agent II.[16] We now show how this can make secret messages useless.

Consider a mechanism where at stage 2 agent II reports agent I's effort: $M_2' = \{\texttt{work},\texttt{shirk}\}$. Let $(w_1^s(m), w_1^f(m))$ denote the expected wage payments to agent I in the success and failure states respectively, conditional on message $m \in \{\texttt{work},\texttt{shirk}\}$. Suppose at stage 3, the principal and agent II can collude against agent I. In equilibrium, agent I must work, and agent II is supposed to tell the truth about agent I's effort. Agent I's expected income in equilibrium is, therefore,

$$p_{11} w_1^s(\texttt{work}) + (1 - p_{11}) w_1^f(\texttt{work}) \tag{21}$$

The moral hazard constraint for agent I is

$$p_{11} w_1^s(\texttt{work}) + (1 - p_{11}) w_1^f(\texttt{work}) - c_1 \geq p_{01} w_1^s(\texttt{shirk}) + (1 - p_{01}) w_1^f(\texttt{shirk}) \tag{22}$$

Now suppose agent I works in equilibrium. If collusion between the principal and agent II is possible, the principal can propose that agent II reports $m_2 = \texttt{shirk}$ if this minimizes the wage-payments to agent I. This makes the principal and agent II jointly better off, and there is always a bribe from the principal to agent II that would make this acceptable. The principal can always pay the bribe as he is not cash-constrained. Moreover, as we always assume side payments can be made contingent on the outcome of the project, the principal will pay the bribe only in case of success to guarantee that agent II has an incentive to work. Agent I's expected wage becomes

$$p_{11} w_1^s(\texttt{shirk}) + (1 - p_{11}) w_1^f(\texttt{shirk}) \tag{23}$$

[16] Agent II sees e_1 and the principal does not, so collusion between agent II and the principal could potentially involve asymmetric information as in Laffont and Martimort [9]. However, in an equilibrium which implements full effort, the principal knows that $e_1 = 1$ with probability one so along the equilibrium path there is no asymmetric information. Thus, if agent I's wages are not minimized along the equilibrium path, there will be a collusive contract which the principal (or the third party) can propose to agent II, and which agent II will accept, involving changing the message so that agent I's wages are minimized. (Such a proposal has no "signalling" effects because the principal has no private information).

For agent II and the principal not to make this deal, it must be the case that $m_2 =$work actually minimizes agent I's expected wage:

$$p_{11}w_1^s(\texttt{shirk}) + (1-p_{11})w_1^f(\texttt{shirk}) \geq p_{11}w_1^s(\texttt{work}) + (1-p_{11})w_1^f(\texttt{work}) \qquad (24)$$

Now consider the smallest expected wage payment (21) agent I can receive, subject to (22) and (24). We may set $w_1^f(\texttt{shirk}) = 0$, for if $w_1^f(\texttt{shirk}) > 0$ then we can reduce $w_1^f(\texttt{shirk})$ and increase $w_1^s(\texttt{shirk})$ while keeping

$$p_{01}w_1^s(\texttt{shirk}) + (1-p_{01})w_1^f(\texttt{shirk})$$

constant. Then, (22) will still be satisfied, while (24) now holds with strict inequality because $p_{11} > p_{01}$. Then (24) implies

$$w_1^s(\texttt{shirk}) \geq w_1^s(\texttt{work}) + \frac{1-p_{11}}{p_{11}}w_1^f(\texttt{work}) \qquad (25)$$

Substituting (25) in (22) yields

$$p_{11}w_1^s(\texttt{work}) + (1-p_{11})w_1^f(\texttt{work}) - c_1 \geq p_{01}w_1^s(\texttt{shirk}) \geq p_{01}\left(w_1^s(\texttt{work}) + \frac{1-p_{11}}{p_{11}}w_1^f(\texttt{work})\right)$$

Rearranging, we find that the expected wage payment to agent I satisfies

$$p_{11}w_1^s(\texttt{work}) + (1-p_{11})w_1^f(\texttt{work}) \geq p_{11}\frac{c_1}{p_{11}-p_{01}}$$

Agent II's moral hazard constraint must also be satisfied, which implies

$$w_2^s(\texttt{work}) \geq \frac{c_2}{p_{11}-p_{10}}$$

Therefore, the total expected wage payment is at least

$$p_{11}\left(\frac{c_1}{p_{11}-p_{01}} + \frac{c_2}{p_{11}-p_{10}}\right)$$

However, this is not cheaper than a simple contract. The principal can always implement full effort with a message-free contract with

$$\begin{array}{ll} w_1^s = \frac{c_1}{p_{11}-p_{01}} \;, & w_1^f = 0 \\ w_2^s = \frac{c_2}{p_{11}-p_{10}} \;, & w_2^f = 0 \end{array}$$

because in this case both agents' moral hazard constraints are satisfied, together with the "group constraint"

$$(p_{11}-p_{00})(w_1^s + w_2^s) \geq c_1 + c_2$$

This is intuitive. For messages to be useful, agent II's message must be used to punish agent I if he shirks. Punishment means lower wages, but then the principal can convince agent II to always claim agent I shirked. If the principal cannot commit not to collude with agent II, the combination of secret messages and secret non-randomized wages has no value. To get implementation at the second best cost, the principal *must* use (public) messages and secret *randomized* wages (Appendix 2 shows that in this case collusion on the part of the principal can be ruled out). If secret randomizations are impossible to enforce for a third party, then the contracts analyzed in the main part of this paper are the best available to the principal.

7 Appendix 1: Proofs

7.1 Proof of Proposition 2

As the Coase theorem does not hold, a crucial decision variable for the principal is the distribution of surplus among the agents. By giving each agent the "option" to leave the game and receive a certain guaranteed payoff, the principal makes sure that the distribution of surplus is the right one. Therefore, it is quite straightforward to show that a message game called an *option mechanism*, where each agent says "stay" or "leave", will always be optimal. We will prove this as a preliminary result in Lemma 9. In fact, as long as the agents can collude on messages, this result seems to be more general than our particular model. The next step, to show that *no message at all* is necessary is more difficult, and may depend on the precise assumptions.

A mechanism is an option mechanism if and only if $M_1^0 = M_2^0 = \{\texttt{stay}, \texttt{leave}\}$, $M_1' = M_2' = \emptyset$, and there exists numbers (π_1, π_2) such that:

$$(w_i^f(m), w_i^s(m)) = \begin{cases} (0, \frac{\pi_i + c_i}{p_{11}}) & \text{if } m_1 = m_2 = \texttt{stay} \\ (\pi_i, \pi_i) & \text{otherwise} \end{cases} \tag{26}$$

The option mechanism is *not* collusion-proof: in equilibrium, agents may sign a side contract. However, agent i will never accept a payoff less than π_i, which he can guarantee himself by saying `leave` and refusing to work. Thus, by choosing the numbers (π_1, π_2) the principal influences the collusive contracts the agents may sign.[17]

Lemma 9 *Suppose a mechanism (M, w) has a side contract c which implements full effort. Then, the option mechanism where wages are given by (26), with $\pi_i = \pi_i(c)$, has an equilibrium side contract $\bar{c}$ which implements full effort. In the equilibrium side contract $\bar{c}$ of the option mechanism, both agents work hard, announce* `stay`, *there are*

[17] We show later that in a simple contract the distribution of the wage packet between the two agents can perform the same role as the outside option. Therefore, delegation without any messages is optimal.

no side payments and the expected wage payments are the same as in the equilibrium side contract c of the mechanism (M, w).

Proof. Suppose $\hat{c} = (\hat{e}_1, \hat{m}, \hat{t})$ is an equilibrium side contract of a mechanism (M, w) which implements full effort. By definition, the equilibrium payoffs are

$$\pi_1(\hat{c}) = p_{11}\left(w_1^s(\hat{m}) - \hat{t}^s\right) + (1 - p_{11})\left(w_1^f(\hat{m}) - \hat{t}^f\right) - c_1 \tag{27}$$

$$\pi_2(\hat{c}) = p_{11}\left(w_2^s(\hat{m}) + \hat{t}^s\right) + (1 - p_{11})\left(w_2^f(\hat{m}) + \hat{t}^f\right) - c_2 \tag{28}$$

Consider a mechanism $(\bar{M}, \bar{w})$ such that: $\bar{M}_1^0 = \bar{M}_2^0 = \{\texttt{stay}, \texttt{leave}\}$, $\bar{M}_1' = \bar{M}_2' = \emptyset$, and:

$$(\bar{w}_1^f(m), \bar{w}_1^s(m)) = \begin{cases} (w_1^f(\hat{m}) - \hat{t}^f, w_1^s(\hat{m}) - \hat{t}^s) & \text{if } m_1 = m_2 = \texttt{stay} \\ (\pi_1(\hat{c}), \pi_1(\hat{c})) & \text{otherwise} \end{cases} \tag{29}$$

$$(\bar{w}_2^f(m), \bar{w}_2^s(m)) = \begin{cases} (w_2^f(\hat{m}) + \hat{t}^f, w_2^s(\hat{m}) + \hat{t}^s) & \text{if } m_1 = m_2 = \texttt{stay} \\ (\pi_2(\hat{c}), \pi_2(\hat{c})) & \text{otherwise} \end{cases} \tag{30}$$

Notice that expected wage payments in this option mechanism if both agents announce stay are the same as in the equilibrium side contract of (M, w). We claim that $\bar{c} = (\bar{e}_1, \{\texttt{stay}, \texttt{stay}\}, \bar{t})$, where $\bar{e}_1 = 1$ and $\bar{t} = (0, 0)$, is an equilibrium side contract of $(\bar{M}, \bar{w})$ which implements full effort. To check this, first notice that signing $\bar{c}$ in $(\bar{M}, \bar{w})$ gives agent II an incentive to work, and results in payoffs $\pi_i(\bar{c}) = \pi_i(\hat{c})$ where the $\pi_i(\hat{c})$ are given by (27) and (28). Agent I, of course, must work once he signs $\bar{c}$, since side contracts are binding. We only need to verify (E1) and (E2). Now, (E1) holds for the contract $\bar{c}$ because if some agent refuses to accept $\bar{c}$, it is a subgame perfect equilibrium for both to announce leave which gives agent i the wage $\pi_i(\hat{c})$ for sure, which is no improvement. Secondly, (E2) holds because if there exists a feasible side contract $c' = (e_1', m', t')$ for $(\bar{M}, \bar{w})$ which satisfies $\pi_1(c') > \pi_1(\bar{c})$ and $\pi_2(c') > \pi_2(\bar{c})$, then it must be that $m' = \{\texttt{stay}, \texttt{stay}\}$. But then the agents would be able to improve on $\hat{c}$ in the original mechanism (M, w) too, via the following feasible side contract: agent I agrees to e_1', the agents send the original messages $\hat{m}$ (which gives the same wages as $m' = \{\texttt{stay}, \texttt{stay}\}$ in $(\bar{M}, \bar{w})$), and the transfers are $\hat{t} + t'$. This precisely duplicates the side contract c'. However, such an improvement contradicts $\hat{c}$ being an equilibrium side contract. Therefore, $(\bar{M}, \bar{w})$ implements full effort.

Finally, given that $(\bar{M}, \bar{w})$ implements full effort, consider the option mechanism with

$$(w_i^f(m), w_i^s(m)) = \begin{cases} (0, \frac{\pi_i(\hat{c}) + c_i}{p_{11}}) & \text{if } m_1 = m_2 = \texttt{stay} \\ (\pi_i(\hat{c}), \pi_i(\hat{c})) & \text{otherwise} \end{cases} \tag{31}$$

Comparing (29)-(30) and (31), and using (27) and (28), we find that the option mechanism is identical to $(\bar{M}, \bar{w})$ if $w_1^f(\hat{m}) = w_2^f(\hat{m}) = 0$. Otherwise, the option mechanism differs by shifting all the wage payments to the success state, but the expected wages conditional on both working is the same as in $(\bar{M}, \bar{w})$. Clearly the option mechanism makes shirking less desirable for each individual agent. As it reduces the sum of expected wage payments for all effort profiles where at least one agent shirks, it makes collusion to shirk less desirable too. So, if the agents agree to work in $(\bar{M}, \bar{w})$ (as we assume) then they should certainly agree to work in the option mechanism. Indeed, consider the side contract for the option mechanism: $c = (e_1, \{\texttt{stay}, \texttt{stay}\}, t)$ with $e_1 = 1$, $t = (0, 0)$. Then this side contract is feasible and it is easy to check that it implements full effort. Therefore, the option mechanism implements full effort. ■

Proof of Proposition 2. We claim effort cannot be implemented by any mechanism at a cost lower than

$$p_{11} \frac{c_1 + c_2}{p_{11} - p_{00}}$$

By Lemma 9 we can without loss of generality consider an option mechanism which implements full effort without side payments. Let

$$w_i \equiv w_i^s(\texttt{stay}, \texttt{stay})$$

In the equilibrium side contract, both agents announce stay, there are no side-payments and the payoffs for the agents are (π_1, π_2) where $\pi_i = p_{11} w_i - c_i \geq 0$.

Suppose, in order to derive a contradiction, that the cost to the principal is

$$p_{11} (w_1 + w_2) < p_{11} \frac{c_1 + c_2}{p_{11} - p_{00}} \tag{32}$$

Consider the feasible side contract $\bar{c} = (\bar{e}_1, \{\texttt{stay}, \texttt{stay}\}, \bar{t})$ where $\bar{e}_1 = 0$, $\bar{t}^f = 0$.

$$\bar{t}^s = w_1 - \frac{p_{11}}{p_{00}} w_1 + \frac{c_1}{p_{00}} - \varepsilon$$

and $\varepsilon > 0$. Suppose under $\bar{c}$ agent II shirks. Then

$$\pi_1(\bar{c}) = p_{00} (w_1 - \bar{t}^s) = p_{11} w_1 - c_1 + p_{00} \varepsilon > \pi_1 \tag{33}$$

and

$$\pi_2(\bar{c}) = p_{00} (w_2 + \bar{t}^s) = p_{00} (w_2 + w_1) - p_{11} w_1 + c_1 - p_{00} \varepsilon \tag{34}$$

Using (32),

$$\begin{aligned} \pi_2(\bar{c}) - \pi_2 &= p_{00} (w_1 + w_2) - p_{11} w_1 + c_1 - p_{00} \varepsilon - (p_{11} w_2 - c_2) \\ &= c_1 + c_2 - (p_{11} - p_{00}) (w_1 + w_2) - p_{00} \varepsilon > 0 \end{aligned} \tag{35}$$

for sufficiently small $\varepsilon > 0$. Thus, under the assumption that II shirks, both agents are strictly better off under $\bar{c}$. If agent II actually prefers to work under $\bar{c}$, by revealed preference it must give him an even higher payoff than (34). Agent II working cannot hurt agent I either, as agent I makes no money if the project fails. Thus, in any case, both agents are strictly better off with $\bar{c}$. Finally, (33) and (35) imply that $w_1 - \tilde{t}^s$ and $w_2 + \tilde{t}^s$ are both non-negative, so $\bar{c}$ is a feasible side contract. This contradicts c being an equilibrium side contract. ■

7.2 Proof of Proposition 3

Suppose

$$\frac{c_1}{p_{11} - p_{01}} > \max \left\{ \frac{c_1 + c_2}{p_{11} - p_{00}}, \frac{c_1}{p_{11}} + \frac{c_2}{p_{11} - p_{10}} \right\}$$

and

$$\frac{p_{10}}{p_{11} - p_{10}} \geq \frac{p_{00}}{p_{01} - p_{00}} \tag{36}$$

We claim full effort cannot be implemented by any mechanism at a cost lower than

$$p_{11} \frac{c_1}{p_{11} - p_{01}}$$

By Lemma 9 we may again consider an option mechanism which implements full effort without side payments. Let

$$w_i \equiv w_i^s(\texttt{stay}, \texttt{stay})$$

Again, $\pi_i = p_{11} w_i - c_i$ is agent i's equilibrium payoff. As the option mechanism implements full effort,

$$w_2 \geq \frac{c_2}{p_{11} - p_{10}} \tag{37}$$

Suppose in order to obtain a contradiction that

$$p_{11} (w_1 + w_2) < \frac{p_{11} c_1}{p_{11} - p_{01}} \tag{38}$$

Then,

$$p_{11} (w_1 + w_2) - c_1 - c_2 < p_{01} (w_1 + w_2) - c_2 \tag{39}$$

That is, if agent I shirks (and the agents stick to the messages (stay,stay)), the sum of the agents expected payoffs is strictly increased. Let

$$\tilde{t}^s = \frac{p_{11} w_2}{p_{01}} - w_2 \tag{40}$$

We claim that this positive transfer fulfills the moral hazard condition for agent II, conditional on $e_1 = 0$:

$$(p_{01} - p_{00})(w_2 + \tilde{t}^s) \geq c_2 \tag{41}$$

Indeed, if (41) does not hold, then using (37) and the definition of $\tilde{t}^s$,

$$\begin{aligned}\frac{p_{10}c_2}{p_{11}-p_{10}} = \frac{p_{11}c_2}{p_{11}-p_{10}} - c_2 \le p_{11}w_2 - c_2 = p_{01}(w_2+\tilde{t}^s) - c_2 \\ < \frac{p_{01}c_2}{p_{01}-p_{00}} - c_2 = \frac{p_{00}c_2}{p_{01}-p_{00}}.\end{aligned} \tag{42}$$

But this contradicts (36). Thus (41) holds. But then, for small enough $\epsilon > 0$, we have violated condition (E2) with respect to the feasible side contract $\bar{c} = (\bar{e}_1, \bar{m}, \bar{t})$ where $\bar{e}_1 = 0$, $\bar{m} = (\texttt{stay}, \texttt{stay})$, $\bar{t}^f = 0$ and $\bar{t}^s = \tilde{t}^s + \epsilon/p_{01}$. For we have:

$$w_2 + \tilde{t}^s + \epsilon/p_{01} > \frac{c_2}{p_{01}-p_{00}}$$

by (41), and

$$\pi_2(\bar{c}) = p_{01}(w_2 + \tilde{t}^s + \epsilon/p_{01}) - c_2 = p_{11}w_2 - c_2 + \epsilon > p_{11}w_2 - c_2 = \pi_2$$

and

$$\pi_1(\bar{c}) = p_{01}(w_1 - \tilde{t}^s - \epsilon/p_{01}) > p_{11}w_1 - c_1 = \pi_1 \tag{43}$$

by (39). By (43), $\tilde{t}^s < w_1$ so $\bar{c}$ is feasible. But this improvement contradicts the assumption that the option mechanism implements full effort. ■

7.3 Proof of Proposition 4

Suppose

$$\frac{c_1}{p_{11}-p_{01}} > \max\left\{\frac{c_1+c_2}{p_{11}-p_{00}}, \frac{c_1}{p_{11}} + \frac{c_2}{p_{11}-p_{10}}\right\}$$

and

$$\frac{p_{10}}{p_{11}-p_{10}} < \frac{p_{00}}{p_{01}-p_{00}} \tag{44}$$

We claim full effort cannot be implemented by any mechanism at a cost lower than

$$\frac{p_{11}c_1}{p_{11}-p_{01}} + \frac{p_{11}c_2}{p_{11}-p_{01}}\left[\frac{p_{10}}{p_{11}-p_{10}} - \frac{p_{00}}{p_{01}-p_{00}}\right]$$

By Lemma 9 we may again consider an option mechanism which implements full effort without side payments. Let

$$w_i \equiv w_i^s(\texttt{stay}, \texttt{stay})$$

As agent II works in the equilibrium, (37) holds. Suppose in order to obtain a contradiction that the cost to the principal is

$$p_{11}(w_1+w_2) < \frac{p_{11}c_1}{p_{11}-p_{01}} + \frac{p_{11}c_2}{p_{11}-p_{01}}\left[\frac{p_{10}}{p_{11}-p_{10}} - \frac{p_{00}}{p_{01}-p_{00}}\right] \tag{45}$$

We shall construct a feasible contract which is strictly preferred by the agents. Consider $\tilde{c} = (\tilde{e}_1, \tilde{m}, \tilde{t}^f, \tilde{t}^s)$ with $\tilde{e}_1 = 0$, $\tilde{m} = (\texttt{stay, stay})$, $\tilde{t}^f = 0$ and

$$\tilde{t}^s = w_1 - \frac{p_{11}}{p_{01}}\left(w_1 - \frac{c_1}{p_{11}}\right) \tag{46}$$

Then

$$p_{01}(w_1 - \tilde{t}^s) = p_{11}w_1 - c_1 \geq 0 \tag{47}$$

Combining (45) and (44) we get

$$p_{11}(w_1 + w_2) - c_1 < p_{01}(w_1 + w_2) \tag{48}$$

By (48),

$$p_{01}(w_2 + \tilde{t}^s) - c_2 > p_{11}w_2 - c_2 \geq 0 \tag{49}$$

Note that (45) implies

$$w_1 + w_2 > \frac{p_{11}}{p_{01}}(w_1 + w_2) - \frac{c_1}{p_{01}} - \frac{c_2}{p_{01}}\left[\frac{p_{10}}{p_{11} - p_{10}} - \frac{p_{00}}{p_{01} - p_{00}}\right] \tag{50}$$

¿From (50), and as (37) must hold, agent II's moral hazard constraint is satisfied under $\tilde{c}$:

$$\begin{aligned} w_2 + \tilde{t}^s &= w_1 + w_2 - \tfrac{p_{11}}{p_{01}}\left(w_1 - \tfrac{c_1}{p_{11}}\right) \\ &> \tfrac{p_{11}}{p_{01}}w_2 - \tfrac{c_2}{p_{01}}\left[\tfrac{p_{10}}{p_{11}-p_{10}} - \tfrac{p_{00}}{p_{01}-p_{00}}\right] \geq \tfrac{c_2}{p_{01}-p_{00}} \end{aligned} \tag{51}$$

Now (47), (49) and (51) implies that $\tilde{c}$ is an equilibrium side contract which makes both agents weakly better off, and only agent II works under $\tilde{c}$. Also, (47) and (49) imply the transfer is feasible. The inequalities in (49) and (51) are strict so we can make both agents strictly better off by reducing the transfer to $\tilde{t}^s - \epsilon$ for ϵ small and positive. Thus (E2) is violated, a contradiction. ■

7.4 Proof of Proposition 5

We first state a preliminary result which will be useful in several subsequent proofs.

Lemma 10 *Let (w_1, w_2) be a simple contract and $E^{\Gamma(w_1,w_2)} = (x, y)$. A necessary condition for implementation of full effort is that (52) holds:*

$$x \leq p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1 \tag{52}$$

A sufficient condition for implementation of full effort is that (52) holds and $e = (1, 1)$ maximizes the sum of the agents expected payoffs $\pi_1 + \pi_2$ subject to $\pi_i \geq 0$ for $i = 1, 2$.

Proof. First, suppose full effort is implemented by some equilibrium side contract $c = (e_1, t)$, but (52) does not hold. Then by (E1),

$$\pi_1(c) = p_{11}(w_1 - t) - c_1 \geq x > p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1 \tag{53}$$

But (53) implies

$$w_2 + t < \frac{c_2}{p_{11} - p_{10}}$$

Then II's moral hazard constraint is not satisfied so c will not induce him to work, a contradiction. Thus, (52) is necessary.

Now suppose (52) holds and $e = (1, 1)$ maximizes the sum of the agents expected payoffs $\pi_1 + \pi_2$ subject to $\pi_i \geq 0$ for $i = 1, 2$. Consider the feasible side contract $c = (e_1, t)$ with $e_1 = 1$ and

$$p_{11}(w_1 - t) - c_1 = x$$

In case of success II gets

$$\begin{gathered} w_2 + t = w_1 + w_2 - \tfrac{c_1 + x}{p_{11}} \\ \geq w_1 + w_2 - \tfrac{c_1}{p_{11}} - \tfrac{1}{p_{11}}\left(p_{11}\left(w_1 + w_2 - \tfrac{c_2}{p_{11} - p_{10}}\right) - c_1\right) = \tfrac{c_2}{p_{11} - p_{10}} \end{gathered}$$

so that II's moral hazard constraint is satisfied. Since effort levels $e = (1, 1)$ maximize the sum of the payoffs in the positive quadrant, $\pi_1(c) = x$ and $\pi_2(c) \geq y$, c is an equilibrium side contract which implements full effort. ■

Proof of Proposition 5. Suppose (14) holds. We claim full effort can be implemented by a simple contract at the cost

$$p_{11}\left(\frac{c_2}{p_{11} - p_{10}} + \frac{c_1}{p_{11}}\right)$$

and to implement full effort at this cost, it is necessary that $w_1 = 0$.

Let (w_1, w_2) be such that

$$w_1 + w_2 = \frac{c_2}{p_{11} - p_{10}} + \frac{c_1}{p_{11}}$$

Note first that

$$p_{11}(w_1 + w_2) - c_1 - c_2 \geq p_{01}(w_1 + w_2) - c_2 \tag{54}$$

$$p_{11}(w_1 + w_2) - c_1 - c_2 \geq p_{10}(w_1 + w_2) - c_1 \tag{55}$$

$$p_{11}(w_1 + w_2) - c_1 - c_2 \geq p_{00}(w_1 + w_2) \tag{56}$$

where (54) and (56) use (14). Thus, the sum of the agents payoffs is greater for $e = (1, 1)$ than for any other effort levels. We have

$$p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1 = 0 \tag{57}$$

If $w_1 = 0$ then the most agent I can get in a subgame perfect equilibrium of $\Gamma(w_1, w_2)$ is clearly zero. By Lemma 10, full effort is implemented. Now suppose $w_1 > 0$. Then, a subgame perfect equilibrium of $\Gamma(w_1, w_2)$ must give agent I at least $p_{00}w_1 > 0$. As (57) holds, this contradicts Lemma 10. ■

7.5 Proof of Proposition 6

Suppose (15) holds. We claim that full effort can be implemented by a simple contract at the cost

$$p_{11}\frac{c_1 + c_2}{p_{11} - p_{00}}$$

and to implement full effort at this cost, it is possible (but not necessary) to set $w_1 = 0$.

Let (w_1, w_2) be a simple contract with

$$w_1 + w_2 = \frac{c_1 + c_2}{p_{11} - p_{00}}$$

As before, one can check that the sum of the agents payoffs is greater for $e = (1, 1)$ than for any other effort levels. Moreover, (15) implies

$$p_{11}\left(\frac{c_1 + c_2}{p_{11} - p_{00}} - \frac{c_2}{p_{11} - p_{10}}\right) - c_1 \geq 0 \tag{58}$$

with a strict inequality if there is a strict inequality in (15).

If $w_1 = 0$, then the most agent I can get in a subgame perfect equilibrium of $\Gamma(w_1, w_2)$ is zero. By (58) and Lemma 10, full effort is implemented.

Now suppose there is strict inequality in (15), and hence in (58). Then by Lemma 10, it is possible to set $w_1 > 0$, as long as player I's payoff in subgame perfect equilibrium of $\Gamma(w_1, w_2)$ does not exceed the left hand side of the expression in (58). (It is possible to use Lemma 10 to compute the precise upper bound for w_1, but it is not very informative). ■

7.6 Proof of Proposition 7

Suppose (10) and (9) hold. We claim that full effort can be implemented by a simple contract at the cost $p_{11}w^*$.

Case 1. Suppose

$$w^* < \frac{c_1}{p_{11} - p_{00}} + \frac{p_{10}}{p_{00}}\frac{c_2}{p_{11} - p_{10}}$$

Set

$$w_2 = \frac{p_{10}}{p_{00}} \frac{c_2}{p_{11} - p_{10}} < \frac{c_2}{p_{01} - p_{00}}$$

where the inequality uses (9). Notice that $w_2 < w^*$, for if not then $w^* < \frac{c_2}{p_{01}-p_{00}}$ so that $p_{01}w^* - c_2 < p_{00}w^*$. But (19) implies $p_{11}w^{**} - c_1 - c_2 < p_{01}w^* - c_2$, so that

$$p_{11}w^* - c_1 - c_2 < p_{00}w^*$$

But this contradicts the definition of w^*. Thus, $w_2 < w^*$. Set

$$w_1 = w^* - w_2 < \frac{c_1}{p_{11} - p_{00}}$$

In the subgame perfect equilibrium of Γ, agent II will work iff agent I has worked, but agent I will not work (because $p_{00}w_1 > p_{11}w_1 - c_1$). Therefore, $E^\Gamma = (p_{00}w_1, p_{00}w_2)$. To implement full effort, it is necessary and sufficient that there exists a transfer t such that (59) - (62) hold:

$$p_{11}(w_1 - t) - c_1 \geq p_{00}w_1 \tag{59}$$

$$p_{11}(w_2 + t) - c_2 \geq p_{00}w_2 \tag{60}$$

$$(p_{11} - p_{10})(w_2 + t) \geq c_2 \tag{61}$$

$$p_{11}(w_1 - t) - c_1 \geq p_{01}\left(w_1 + w_2 - \frac{c_2}{p_{01} - p_{00}}\right) \tag{62}$$

Equation (62) is the condition which guarantees that the contract where both agents work and agent I pays t to agent II in case of success is not Pareto-dominated by some side contract where only agent II works. Indeed, to make agent II willing to work alone, he needs the efficiency wage $\frac{c_2}{p_{01}-p_{00}}$, but then (62) implies that agent I would be made worse off.

Using the definition of w^*, one can check that (59) - (62) hold if

$$t = \frac{c_2}{p_{11} - p_{10}} - w_2$$

Therefore, $c = (1, t)$ is an equilibrium side contract which implements full effort.

Case 2. Suppose

$$w^* \geq \frac{c_1}{p_{11} - p_{00}} + \frac{c_2}{p_{11} - p_{10}}$$

Set

$$w_2 = \frac{c_2}{p_{11} - p_{10}}$$

$$w_1 = w^* - w_2 \geq c_1/(p_{11} - p_{00})$$

If no side agreement is signed, both agents will work, so $E^\Gamma = (p_{11}w_1 - c_1, p_{11}w_2 - c_2)$. Therefore, full effort is implemented if and only if this equilibrium (without transfers) is a Pareto efficient outcome for the agents. The condition for this is (62) with $t = 0$, which can be written as:

$$w_1 = w^* - \frac{c_2}{p_{11} - p_{10}} \geq \frac{p_{01}}{p_{11}} \left(w^* - \frac{c_2}{p_{01} - p_{00}} \right) + \frac{c_1}{p_{11}} \tag{63}$$

But (63) holds by definition of w^*. Thus full effort is implemented.

Finally, agent II cannot be General Contractor if full effort is implemented at the lowest cost. For if we set $w_1 = 0$ and $w_2 = w^*$, then (19) implies that agent I's contribution $(p_{11} - p_{10})w^*$ is smaller than his cost of effort c_1, so agent II will never work and pay agent I to work, a contradiction. ■

7.7 Proof of Proposition 8

Suppose (10) holds but (9) does not hold. We claim full effort can be implemented by a simple contract at the cost

$$p_{11} \frac{c_1}{p_{11} - p_{01}}$$

and it is possible, but not necessary, to set $w_1 = 0$.

Let (w_1, w_2) be a simple contract with

$$w_1 + w_2 = \frac{c_1}{p_{11} - p_{01}}$$

Then

$$p_{11}(w_1 + w_2) - c_1 - c_2 = p_{01}(w_1 + w_2) - c_2$$

and it follows from (10) that

$$p_{11}(w_1 + w_2) - c_1 - c_2 \geq p_{10}(w_1 + w_2) - c_1$$

and

$$p_{11}(w_1 + w_2) - c_1 - c_2 \geq p_{00}(w_1 + w_2)$$

Thus, the sum of the agents's payoffs is equally great for $e = (0, 1)$ and $e = (1, 1)$, and smaller for other effort levels. Moreover

$$p_{11} \left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}} \right) - c_1 = p_{11} \left(\frac{c_1}{p_{11} - p_{01}} - \frac{c_2}{p_{11} - p_{10}} - \frac{c_1}{p_{11}} \right) > 0 \tag{64}$$

from (10). Lemma 10 implies that full effort is implemented if $w_1 = 0$. But as there is strict inequality in (64), it is possible to set $w_1 > 0$. (As before, we could use Lemma 10 to compute the precise upper bound for w_1). ■

8 Appendix 2: Collusion Proofness with Random Wages and Secret Messages

In the public messages and secret random wages model of Section 6, even if the union is willing to punish an agent who does not pay a side payment on the mere suspicion that he may be cheating, collusion can be still ruled out. For the punishment will be carried out every time the principal pays zero, as in this case the agent cannot possibly pay, and if this happens often enough collusion is not worthwhile. Formally, the following message game implements full effort at the second best cost $c_1 + p_{11}c_2/(p_{11} - p_{10})$. At stage 2, agent II reports on agent I's effort level: $M_1^0 = M_2^0 = M_1' = \emptyset$, $M_2' = \{\texttt{work}, \texttt{shirk}\}$. For a fixed $\varepsilon > 0$, wages are given by:

$$w_1^f(\texttt{work}) = w_1^s(\texttt{shirk}) = w_1^f(\texttt{shirk}) = w_2^f(\texttt{shirk}) = w_2^f(\texttt{work}) = 0$$

$$w_2^s(\texttt{work}) = w_2^s(\texttt{shirk}) = \frac{c_2}{p_{11} - p_{10}}$$

$$w_1^s(\texttt{work}) = \begin{cases} \frac{c_1}{\varepsilon p_{11}} & \text{with probability } \varepsilon \\ 0 & \text{with probability } 1 - \varepsilon \end{cases}$$

As player II's wage is independent of his message, there is a subgame perfect equilibrium where he tells the truth, and both agents work. In any side contract that makes both agents better off, agent I must pay an *expected* transfer to agent II of at least

$$\begin{aligned} \min\{(p_{11}w_2^s(\texttt{work}) - c_2) - p_{00}w_2^s(\texttt{work}), (p_{11}w_2^s(\texttt{work}) - c_2) - (p_{01}w_2^s(\texttt{work}) - c_2)\} \\ = \min\left\{\tfrac{p_{10}-p_{00}}{p_{11}-p_{10}}c_2, \tfrac{p_{11}-p_{01}}{p_{11}-p_{10}}c_2\right\} \equiv \bar{T} > 0 \end{aligned}$$

A transfer t^s can only occur if agent I gets a non-zero wage, so $\bar{T} \leq \varepsilon t^s$. That is, agent I promises to pay $t^s \geq \bar{T}/\varepsilon$ whenever he has any money, in exchange for the right to shirk. To give agent I the proper incentive to pay whenever possible, agent I must suffer a cost $h \geq \bar{T}/\varepsilon$ if the project is successful but he does not pay. Because with probability $1 - \varepsilon$ he cannot pay, the expected cost of this punishment is at least $p_{00}(1 - \varepsilon)\bar{T}/\varepsilon$. Then, for sufficiently small ε collusion is clearly not worthwhile.

In this model, collusion on the part of the principal can be avoided if the workers' union can impose sufficiently strong penalties on agents who collude with the principal. Suppose the agents sign the following side-contract: agent I sets $e_1 = 1$, agent II announces `work` and there are no transfers. If any agent does not conform to this contract, i.e. if agent I does not work or if agent II does not announce `work` even though agent I worked, the union punishes the cheating agent at a cost of $h \geq c_1$ (it is possible because the union can observe effort levels and public messages by assumption). This contract is an equilibrium side-contract as it certainly satisfies (E1) and also satisfies (E2) as by the above argument other collusive contracts cannot be enforced. If the principal is to collude successfully with agent II and give him the incentive to announce `shirk` after agent I has in fact worked, she must pay II at least c_1 to counterbalance the punishment the workers' union will impose on agent II. But then there is no incentive for the principal to collude with agent II as her total payments do not go down.

References

[1] Aghion, P. and J. Tirole (1997), Formal and Real Authority in Organizations, *Journal of Political Economy*, 105: 1-29.

[2] The American Institute of Architects (1990), *Building Relationships.*

[3] Baliga, S. (1995), Universally Collusion-Proof Contracts, Mimeo, Cambridge University.

[4] Baliga, S. and T. Sjöström (1995), Decentralization and Collusion, Harvard Institute of Economic Research Working Paper

[5] The Boston Society of Architects (1973), *The Professional Team Concept.*

[6] Hart, O. (1995), *Firms, Contracts and Financial Structure.* Oxford: Oxford University Press.

[7] Holmström B. and P. Milgrom (1990), Regulating Trade Among Agents, *Journal of Institutional and Theoretical Economics*, 146: 85-105.

[8] Itoh H. (1993), Coalitions, Incentives and Risk- Sharing, *Journal of Economic Theory*, 60: 410-418.

[9] Laffont J.-J. and D. Martimort, Collusion and Delegation, mimeo, Toulouse.

[10] Ma, C. (1988), Unique Implementation of Incentive Contracts with Many Agents. *Review of Economic Studies* 55: 555-572.

[11] Macho-Stadler I. and Perez-Castrillo J. (1994), Centralized and Decentralized Contracts in a Moral Hazard Environment, Working Paper, University Autonoma de Barcelona.

[12] Macho-Stadler I. and Perez-Castrillo J. (1993), Moral Hazard with Several Agents: The gains from Cooperation", *International Journal of Industrial Organization,* 11: 73-100.

[13] Maskin E., and J. Tirole (1996), Dynamic Programming, Unforeseen Contingencies and Incomplete Contracts, mimeo, Harvard University.

[14] McAfee R.P. and J. McMillan (1995), Organizational Diseconomies of Scale, mimeo. University of Texas at Austin.

[15] Melumad, N., Mookherjee D., and S. Reichelstein (1992), A Theory of Responsibility Centers, *Journal of Accounting and Economics*, 15, 445-484.

[16] Melumad, N., Mookherjee D., and S. Reichelstein (1995), Hierarchial Decentralization of Incentive Contracts, *RAND Journal of Economics*, 26: 654-672.

[17] Mookherjee, D. (1984), Optimal Incentive Schemes with Many Agents, *Review of Economic Studies*, 51: 433-446.

[18] Mookherjee, D. and S. Reichelstein (1995a), Budgeting and Hierarchical Control, mimeo. Boston University.

[19] Mookherjee, D. and S. Reichelstein (1995b), Incentives and Coordination in Hierarchies, mimeo, Boston University.

[20] Ramakrishnan R., and A. Thakor (1991), Cooperation vs. Competition in Agency, *Journal of Law, Economics and Organization*, 7: 248-283.

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

[22] Tirole, J. (1992), Collusion and the Theory of Organizations in *Advances in Economic Theory: Sixth World Congress (Vol. II)* edited by J.-J. Laffont. Cambridge: Cambridge University Press.

[23] Timothy Van Zandt (1995), Lecture Notes on Hierarchical Incentive Contracting, mimeo, Princeton University.