JOURNAL OF ECONOMIC THEORY **83**, 196–232 (1998)
ARTICLE NO. ET962462

Decentralization and Collusion*

Sandeep Baliga

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

and

Tomas Sjöström

Department of Economics, Harvard University, Cambridge, Massachusetts 02138

Received July 15, 1996; revised May 29, 1998

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." *Journal of Economic Literature* Classification Numbers: D23, D82, L14, L22. © 1998 Academic Press

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 failure is observable and verifiable. The probability of success depends on the agents' effort levels. The agents work in sequence. Agent 1's effort is known to both agents while agent 2's effort is his private information. The principal does not observe either effort level. Agent 1 might be a designer such as an architect (or a research and

* 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.

0022-0531/98 $25.00

development department in a firm). He delivers a blueprint to agent 2, who is a builder (or a production department in the firm). Agent 2 discovers agent 1's effort by examining the blueprint, but agent 1 does not monitor the actual production process and thus cannot observe agent 2's effort. The principal's objective is to minimize the cost of getting both agents to work hard. 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. After observing agent 1's effort, agent 2 sends a message to the principal: if he says "agent 1 shirked" then agent 1 is paid nothing; if he says "agent 1 worked" then agent 1 is paid just enough when the project succeeds to give him the incentive to participate—agent 2's wage does not depend on his own message. Agent 2 is paid sufficiently (the "efficiency wage") if the project succeeds to make it worth his while to work. This mechanism has a non-cooperative equilibrium where agent 2 always announces agent 1's effort level truthfully and both agents work, but it is vulnerable to collusion among the agents. We study whether *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 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 side contracts on variables that are observable to them (cf. Tirole [21]). We assume for most of the paper that agent 1's effort, all messages, and all wages are observable to the agents. However, the agents cannot contract on agent 2's effort level, as it is unobservable to agent 1. This together with limited liability restricts the agents' ability to side contract. Agent 2 may only be able to make a side payment to agent 1 after the project has been successful so that agent 2 has received a high wage. However, anticipating this side payment in the success state, agent 2 may have an insufficient incentive to work hard (i.e., his income in the success state is below his efficiency wage). Since the transfer cannot be made independent of the outcome of the project, it may be impossible for agent 2 to transfer surplus to agent 1 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

[1] Without the limited liability constraint, the principal could achieve first best by "selling the firm" to agent 2, who could then monitor and pay agent 1.

set of feasible side contracts for the agents. Therefore, the distribution of wages among the agents is an important control variable for the principal.

We find that decentralization is optimal. In many cases, the optimal contract can be implemented by 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, and either the builder or the 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.[2] 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.

If agent 1 (the architect) is the General Contractor, he must pay agent 2 (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 may 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

[2] Information about the building profession was obtained from the American Institute of Architects [1] and the Boston Society of Architects [5].

Contractor, because working hard is then a way for him to transfer rent from the builder to himself. By producing a very good blueprint 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 surplus. Thus, by delegating to the right agent, the principal makes sure both agents work *even if this does not maximize the agents' joint surplus.*

In the main part of the paper, we model *centralization* as a message game in the traditional mechanism design sense. Messages and wages are publicly observable, which makes collusion on messages easy and message games correspondingly ineffective. We show that such message games cannot improve on the optimal decentralized contract. However, as Maskin and Tirole [12] have argued, any mechanism which is compatible with the assumptions made in the models should not be ruled out *a priori*. In that spirit, in Section 6 we expand the class of message games to include games where messages and wages need not be publicly observable. In order to support collusion, the agents must use side payments ("bribes") that are conditional on the outcome of the game. Now 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 pay a bribe, for ex post he can pretend he never received his wage and therefore he cannot pay. In this case collusion is difficult, and we show that such message games (centralization) can do better than decentralization. But a court might find it difficult to enforce randomized wages, so this mechanism may actually not be feasible. An alternative way to eliminate collusion is to keep *messages* secret. If it cannot be verified whether or not an agent "snitched," collusion is again made difficult. However, secret messages do not work if the principal is also a 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, the optimality of delegation is 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. (Baliga, Corchon and Sjöström [3] analyze the principal's commitment problem in an adverse selection context.)

The literature on multi-agent incentive schemes shows that a non-individualistic situation with monitoring and collusion can be strictly better for the principal than a purely individualistic scheme (Itoh [7], Holmström and Milgrom [6], Macho-Stadler and Perez-Castrillo [11], Ramakrishnan and Thakor [19], Tirole [21]). While it is true also in our model that a purely individualistic scheme (with neither monitoring nor side contracting) would not be optimal, our focus is different: we consider 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 [7], 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). This argument is not valid in our model, since utility is not transferable due to limited liability. The idea that collusion can destroy the usefulness of message games is also explored by Baliga [2].

A different strand of the literature looks at the impact of collusion in a principal–supervisor–agent setting (Tirole [20] and [21]). 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. Recent literature compares decentralized and centralized incentive schemes without collusion in adverse selection models (Melumad, Mookherjee and Reichelstein [14] and [15], Mookherjee and Reichelstein [17] and [18], and McAfee and McMillan [13]). 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 (message games) could achieve in the absence of collusion, but it does as well as centralization if centralized schemes 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 [10] 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 [8] 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 our paper is somewhat different; in particular, we study *how* the principal should delegate.

[3] Van Zandt [22] discusses how the equivalence of decentralization and centralization in adverse selection models depends on whether participation constraints apply ex ante or ex post.

[4] Baliga and Sjöström [4] introduced adverse selection in the model of the current paper and showed that if the optimal distribution of surplus is state dependent, centralization can strictly dominate delegation even though agents can collude.

2. THE MODEL

Two agents work in sequence. Agent 1, who is in charge of design or research and development, delivers a blueprint to agent 2, 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 1 delivers the blueprint to agent 2, agent 2 learns agent 1's effort e_1 by inspecting the blueprint. The blueprint is of low quality if agent 1 shirked ($e_1=0$) but of high quality if agent 1 worked hard ($e_1=1$). Since the principal does not observe e_1, agent 2 is better informed than the principal. Neither agent 1 nor the principal can observe e_2.

After the agents have worked (or shirked), the project is revealed to be either a success or failure. This outcome is public information. The probability of a success is $p_{e_1e_2}$, if agent i's effort is $e_i \in \{0, 1\}$. We assume $0<p_{00}<p_{11}<1$ and

$$p_{00}<p_{e_1e_2}<p_{11} \qquad \text{when} \quad e_1 \neq e_2.$$

Both the agents and the principal know the parameters (c_1, c_2, p_{00}, p_{01}, p_{10}, p_{11}), so there is no adverse selection.

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 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 $w_i-e_ic_i$. 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 2's effort is unobservable to everybody except himself. Therefore, for agent 2 to work, a moral hazard constraint must be satisfied. The "second-best" contract pays agent 1 the expected wage c_1 and satisfies agent 2's moral hazard constraint at the lowest possible cost. Let w_2 be the wage for agent 2 if the outcome is a success (it is clearly optimal to pay zero when the project fails). The moral hazard constraint for agent 2 is $p_{11}w_2-c_2 \geqslant p_{10}w_2$. 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 2.

If the agents cannot collude, full effort can be implemented at the second best cost by asking agent 2 to report agent 1's effort. If the project is unsuccessful, both agents get zero. If it is successful, agent 2 gets $w_2 = c_2/(p_{11} - p_{10})$, and 1 gets $w_1 = c_1/p_{11}$ if 2 has announced that 1's effort was high, otherwise agent 1 gets zero. This mechanism has an equilibrium where agent 2 truthfully reports agent 1's effort level, and both agents work hard. (There may exist other equilibria: unlike Ma [9], we will not require unique implementation.) 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 1, implying a rent for agent 1. But the message game is vulnerable to collusion. If agent 1 shirks, he is willing to pay a bribe to 2 in order for him not to "snitch." Since agent 2's wage does not depend on his message, he is willing to accept the bribe.

We end this section by comparing our model to two alternative models. The first is 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 1, 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 the principal can achieve in our model under the assumption that agent 2 can monitor agent 1 and the agents can collude. Thus, as in Itoh [7], a non-individualistic context with monitoring and collusion is better for the principal than a purely individualistic scheme without monitoring. See also Holmström and Milgrom [6], Macho-Stadler and Perez-Castrillo [11] and Ramakrishnan and Thakor [19].

The second alternative model would maintain the assumption that agent 2 can monitor and side-contract with agent 1, but would drop the assumption of limited liability. Then the principal can implement full effort at "first best" cost $c_1 + c_2$ by "selling the firm" to agent 2, who in effect becomes the new principal who can monitor and side-contract with agent 1. This solves the moral hazard problem completely. Formally, consider the following contract. 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 hard at the first best implies

$$p_{11}\,\Delta - c_1 - c_2 > \max\{p_{10}\,\Delta - c_1,\; p_{01}\,\Delta - c_2,\; p_{00}\,\Delta\}. \tag{2}$$

Pay agent 2 the amount $w_2^f = c_1 + c_2 - p_{11}\,\Delta < 0$ if the project fails and $w_2^s = w_2^f + \Delta > 0$ if the project succeeds; never pay agent 1 anything. Then agent 2 suffers the full cost of a failure, and given (2) will certainly have an incentive to side contract with agent 1, to monitor him and make sure he works. Agent 2 pays agent 1 the expected wage c_1 and works himself. The principal's profit is first best: $(1 - p_{11})\,F + p_{11}(\Delta + F) - (c_1 + c_2)$. This contract is ruled out in our model because limited liability requires $w_2^f \geqslant 0$.

3. COLLUSION AND MESSAGE GAMES

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.[5]

1. Agent 1 works ($e_1 = 1$) or shirks ($e_1 = 0$), and the effort is observed by agent 2.

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

3. Agent 2 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,

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

The pair (M, w) is a *mechanism*. The mechanism together with the rules given above induces a multi-stage two-player 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 $\Gamma(M, w)$.

[5] 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.

Before playing the game, i.e., before time zero, the agents can sign a side contract which is assumed to be enforceable.[6] If some agent refuses to sign a side contract, 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 1'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 1 to agent 2 if the project is a success (failure). The side contract cannot specify agent 2's unobservable effort level as such a contract would not be enforceable.

To be feasible, the side transfers must satisfy

$$-w_2^s(m) \leqslant t^s \leqslant w_1^s(m) \tag{3}$$

$$-w_2^f(m) \leqslant t^f \leqslant w_1^f(m). \tag{4}$$

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 2 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 2 will work iff the increase in his expected wages cover the cost of his effort, i.e., iff

$$(p_{e_1 1} - p_{e_1 0})((w_2^s(m) + t^s) - (w_2^f(m) + t^f)) \geqslant c_2. \tag{5}$$

We refer to (5) as agent 2's moral hazard constraint. If equality holds in (5) then agent 2 is indifferent between working or shirking. In this case we assume that he works. 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 (5) 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 $\pi_i(c) \geqslant x_i^i$.[7] (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

[6] 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.

[7] An alternative, for our purposes equivalent, formulation is that the side contract should dominate some *Pareto efficient* subgame perfect equilibrium of $\Gamma(M, w)$.

payoffs x^i, and there is no other feasible side contract c' that could be signed which strongly Pareto dominates c.[8] (Our results would also hold for *weak* Pareto domination.)

A feasible side contract $c=(e_1, m, t)$ *implements full effort at the cost* C iff $e_1=1$, (5) holds, and

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

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, 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.[9]

If there are many side contracts that satisfy E1 and E2 then the Pareto frontier for the agents is non-trivial, and we *assume* the agents will choose that side contract which maximizes the principal's payoff. In order to implement full effort *uniquely*, some sort of message game would be needed to reduce the Pareto frontier to one point. An example is provided by the "option mechanism" in the appendix. However, we shall not require unique implementation in this paper.

4. NECESSARY CONDITIONS FOR IMPLEMENTATION

In this section we state necessary conditions for implementation by any mechanism, including message games.

[8] 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.

[9] 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.

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

$$c_1 + p_{11}\frac{c_2}{p_{11}-p_{10}}. \tag{6}$$

Proof. Since agent 1 must be compensated for his effort and player 2'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. There is no reason to pay anything in case of failure (see the appendix). By Proposition 1, the expected wage payments $p_{11}(w_1+w_2)$ must exceed (6), i.e.,

$$w_1 + w_2 \geqslant \frac{c_1}{p_{11}} + \frac{c_2}{p_{11}-p_{10}}. \tag{7}$$

Since 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 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

$$w_1 + w_2 \geqslant \frac{c_1+c_2}{p_{11}-p_{00}}. \tag{8}$$

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 2 works but agent 1 shirks, the sum of their payoffs would be $p_{01}(w_1+w_2)-c_2$. The sum is greater when both work than when only agent 2 works if and only if

$$w_1 + w_2 \geqslant \frac{c_1}{p_{11}-p_{01}}. \tag{9}$$

However, (9) may be violated by the optimal contract. Why would agent 1 work if it does not maximize the total surplus for the agents, and the agents can collude? Even though the total surplus is increased when agent 1 shirks (while agent 2 works), a greater share of the surplus may have to be given to agent 2 to induce him to work with a low quality blueprint. This means a smaller share for agent 1, which can make him worse off. Agent 1 may be able to reduce the rent agent 2 must be given (to satisfy his moral hazard constraint) by working hard himself. If agent 1 works hard, then agent 2 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 1 shirks then agent 2'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

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

then agent 2's rent is reduced if agent 1 works hard and produces a good blueprint. In this case, the fact that agent 1's effort relaxes agent 2's moral hazard constraint can make agent 1 work hard even though it reduces the total surplus (i.e., even if (9) is violated). The necessary condition for implementation (Proposition 3) turns out to be that the total surplus when both work hard must be greater than the total surplus when agent 1 shirks *minus* the amount of rent that agent 1 can transfer to himself by working hard. This condition is

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

where the expression is square brackets is the reduction in agent 2's rent if agent 1 works hard.

Proposition 3. *Suppose* (10) *holds. 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.

Notice that (11) is equivalent to $p_{11}(w_1+w_2)$ being greater than the expression in Eq. (12). The previous discussion showed that the positive expression in square brackets in (12) is the reduction in agent 2's rent which can be achieved if agent 1 works hard and produces a good blueprint. It is as if the principal can reduce agent 1's effective cost of effort by the amount of rent agent 1 can transfer from agent 2 to himself by working hard. In the next section, we show how this can be achieved by making agent 1 a General Contractor.[10]

The remaining possibility is that (10) is violated. Then it is cheaper to motivate agent 2 to work when agent 1 shirks. In this case, if (9) is violated then there is not only more surplus to share if agent 1 shirks but this also reduces the rent agent 2 needs to work. Then agent 1 would certainly shirk, so full effort would not be implemented. Therefore, if (10) is violated then (9) must be satisfied. This is the content of Proposition 4.

PROPOSITION 4. *Suppose* (10) *does not hold. Then full effort cannot be implemented by any mechanism at a cost lower than*

$$\frac{p_{11}c_1}{p_{11}-p_{01}}. \tag{13}$$

Proof. In the appendix.

5. SIMPLE CONTRACTS

In Section 4, we found lower bounds for the cost of implementing full effort, using any mechanism (including message games). Only one of the lower bounds derived in Propositions 1–4 will actually be binding: which one depends on the parameters. (The binding constraint is of course the one with the highest cost). We will consider the different cases separately, and we show that in each case, there is a simple contract without messages which implements full effort at a cost equal to the greatest of the lower bounds. Thus, such simple contracts are always optimal.

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*: $M_1^0 = M'_1 = M_2^0 = M'_2 = \varnothing$. 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 1's effort level e_1 and a transfer t to be paid from 1 to 2 if the project is a success, but there are no messages, so we write $c = (e_1, t)$.

[10] We have given an intuitive motivation for why (9) can be violated if agent 1 is a General Contractor and (10) holds. However, in view of Propositions 1 and 2, (9) can only be violated if it is not *implied* by either (7) or (8), i.e., if (9) is the most difficult constraint to satisfy. This is true in case C discussed below.

Since agent 2 can observe the effort of agent 1, one might expect that the optimal simple contract involves delegating the task of monitoring and paying agent 1 to agent 2. We will show that this intuition holds if the right hand side of (7) is greater than the right hand side of (8) and (9), but not necessarily otherwise.

5.1. *Case A*

This is the case

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

We know that full effort cannot be implemented by *any* mechanism at a cost lower than (6). We will show that if (14) holds then this lower bound can be attained by a simple contract. Thus, such contracts are optimal. Moreover, the *unique* way to implement full effort at this cost using a simple contract is to set $w_2=c_2/(p_{11}-p_{10})+c_1/p_{11}$ and $w_1=0$. That is, agent 1 should not be paid anything by the principal. Agent 2 should be a *General Contractor* who receives the whole wage packet from the principal in case of success, and who side contracts with agent 1 to get a good blue print.

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* 2 *must be a 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 1 should work, and agent 2 pays c_1/p_{11} to agent 1 if the project is a success. This contract leaves agent 1 with a zero surplus, and agent 2 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 1, agent 1 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. Then, when agent 1 has worked hard, the "blueprint" is so good that agent 2'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 [6], Itoh [7] and Ramakrishnan and Thakor [19] also suggest that delegation to the agent with superior information (agent 2) is optimal. However, in our model there are other cases too.

5.2. *Case B*

This is the case

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

Like in case A, the sum of the agents' payoffs is maximized when both work, and if agent 2 is the General Contractor then he will sign a contract with agent 1 that makes both agents work. Thus, delegating to agent 2 is optimal ($w_1=0$). However, in contrast to Case A, the most difficult constraint is the "team moral hazard constraint" (8), 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 2 keeps enough of the surplus that he is willing to work. If agent 2 gets an insufficient share of the surplus, he would need a transfer from agent 1 in order to work. Since agent 1 cannot observe agent 2's effort, the transfer has to be big enough to satisfy agent 2's moral hazard constraint. But this gives agent 2 a rent, and agent 1 might then prefer to have agent 2 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 2 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 2 can be a General Contractor).

Proof. In the appendix.

5.3. *Case C*

This is the remaining case, where

$$\frac{c_1}{p_{11} - p_{01}} > \max\left\{\frac{c_2}{p_{11} - p_{10}} + \frac{c_1}{p_{11}}, \frac{c_1 + c_2}{p_{11} - p_{00}}\right\}. \tag{16}$$

Agent 1's effort costs him c_1, and increases the probability of success by $p_{11} - p_{01}$ (assuming 2 works). When (16) holds, it is relatively costly for agent 1 to improve the probability of success, and hence it is tempting for the agents to allow 1 to shirk.

Suppose (10) holds, as in Proposition 3. When (16) and (10) hold, it turns out to be optimal for the principal to pay wages in such a way that (9) is violated, and the sum of the agents' expected payoffs when both work is *smaller* than the sum of the expected payoffs *when agent 1 shirks and agent 2 works.*[11] That is,

$$p_{11}(w_1 + w_2) - c_1 - c_2 < p_{01}(w_1 + w_2) - c_2. \tag{17}$$

Just as explained in Section 4, for this configuration of parameters the principal can exploit the fact that agent 1 is willing to work hard to transfer rent from agent 2 to himself. Suppose agent 1 is the General Contractor who is given all of the wages in case of success, i.e., $w_2 = 0$. For agent 2 to work, he must receive a transfer from agent 1 if the project is successful. Now, (10) implies that the expected transfer that must be given to agent 2 to make him work is greater when agent 1 has shirked than when agent 1 has worked hard: agent 2's marginal productivity is reduced by working with a low quality blueprint. Consider the following two side contracts. Under *contract* α both agents work, and agent 1 transfers $t^\alpha = c_2/(p_{11} - p_{10})$ to agent 2 if the project succeeds. This is the smallest transfer that will make agent 2 work (his efficiency wage), and it will give him a rent equal to $p_{10}c_2/(p_{11} - p_{10})$. Under *contract* β agent 1 shirks, and pays agent 2 transfer $t^\beta = c_2/(p_{01} - p_{00})$ if the project succeeds. This is the smallest transfer that will make agent 2 work given that 1 has shirked, and it will give agent 2 a rent equal to $p_{00}c_2/(p_{01} - p_{00})$. Since (10) holds, agent 2 prefers

[11] Notice that if the right hand side of (9) is smaller than the right hand side expressions of (7) and (8), then (9) must in fact hold because we have shown that (7) and (8) must hold. The interesting case is therefore case C where $c_1/(p_{11} - p_{01})$ is big.

contract β with the greater rent, but agent 1 prefers contract α if $p_{11}(w_1 - t^\alpha) - c_1$ is greater than $p_{01}(w_1 - t^\beta)$. This is true if

$$p_{11} w_1 - c_1 - \frac{p_{10}}{p_{11} - p_{10}} c_2 \geqslant p_{01} w_1 - \frac{p_{00}}{p_{01} - p_{00}} c_2$$

or equivalently

$$w_1 \geqslant \frac{c_1}{p_{11} - p_{01}} + \frac{c_2}{p_{11} - p_{01}} \left[\frac{p_{10}}{p_{11} - p_{10}} - \frac{p_{00}}{p_{01} - p_{00}} \right]. \tag{18}$$

Moreover, if (18) holds there is no way of making contract β more attractive for agent 1, as any reduction in t^β will cause agent 2 to shirk. Since contract β is joint surplus maximizing, agent 2 would like to make an ex ante lump sum transfer to agent 1 in exchange for incentive contract β, but lump sum transfers are ruled out by the limited liability (agent 2 cannot pay in the failure state). If agent 1 is a General Contractor, he will prefer contract α.

The crucial issue is the degree of complementarity between the two agents' inputs. If (10) holds then agent 1's effort makes agent 2 more productive. If agent 1 is the General Contractor then he works hard to relax agent 2's moral hazard constraint (even if it reduces *total* surplus). Thus, while in case A, the principal should not pay agent 1 to assure implementation of an effort profile (1, 1) that maximizes the sum of the agents' expected payoffs, now she must pay agent 1 to *prevent* implementation of an effort profile (0, 1) that maximizes the sum of expected payoffs. If agent 2 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 2 will never pay agent 1 to work in this case. So it is *not* optimal to make agent 2 a General Contractor.

Define

$$w^* \equiv \max \left\{ \frac{c_1}{p_{11} - p_{01}} + \frac{c_2}{p_{11} - p_{01}} \left[\frac{p_{10}}{p_{11} - p_{10}} - \frac{p_{00}}{p_{01} - p_{00}} \right], \frac{c_1}{p_{11}} + \frac{c_2}{p_{11} - p_{10}}, \frac{c_1 + c_2}{p_{11} - p_{00}} \right\}. \tag{19}$$

If (10) and (16) hold then

$$w^* < c_1 / (p_{11} - p_{01}). \tag{20}$$

It follows from Propositions 1, 2, and 3 that full effort cannot be implemented by any mechanism at a cost below $p_{11} w^*$.

PROPOSITION 7. *Suppose* (10) *and* (16) *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* 2 *cannot be a General Contractor*).

Proof. In the appendix.

The final case is when (16) holds but (10) does not hold. Proposition 4 shows that no contract can have a cost lower than $p_{11}c_1/(p_{11} - p_{01})$. Again, this lower bound can be achieved by a simple contract. Since (10) does not hold, $w_1 = 0$ is again optimal.

PROPOSITION 8. *Suppose* (16) *holds but* (10) *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* 2 *can be a 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 2 the General Contractor, *except* when (10) and (16) hold. The only remaining issue is, if (10) and (16) hold, is it optimal to make *agent 1* 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 1 wants agent 2 to work, he prefers to also work himself to minimize agent 2's rent. Does agent 1 want to induce agent 2 to work? By working and paying agent 2 his efficiency wage $c_2/(p_{11} - p_{10})$, agent 1'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 2, agent 1 gets $p_{10}w^* - c_1$, and if both agents shirk, agent 1 gets $p_{00}w^*$. This implies that agent 1 is willing to pay agent 2 to work if and only if

$$w^* \geqslant \max\left\{\frac{p_{11}c_2}{(p_{11} - p_{10})^2}, \frac{c_1}{p_{11} - p_{00}} + \frac{p_{11}c_2}{(p_{11} - p_{10})(p_{11} - p_{00})}\right\}. \tag{21}$$

Thus, if (10) and (16) hold, then agent 1 should be the General Contractor only if (21) holds,[12] but *otherwise the agents must split*[13] *the wage packet* ($w_1 > 0$ and $w_2 > 0$), that is, a "triangular" organization is optimal for the principal. Notice that (21) 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 2 is vital for the success of the project. Then if agent 1 is a General Contractor he will have the incentive to pay agent 2 an efficiency wage to work, and he will work hard himself to reduce the efficiency wage.

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. 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 randomized 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 1's effort, and will punish an agent who cheats. But by assumption, it cannot monitor secret wage payments (or agent 2's effort). Let the cost of being punished by the union be $h > 0$, 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

[12] It is easy to check that neither (21) nor its negation are implied by (16) and (10).

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

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* sent by the agents to the principal can be kept 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 agent 2 has sent the "right" message to the principal about agent 1's effort. However, secret messages open the possibility for the principal to collude with agent 2. Suppose the principal and agent 2 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 2 cannot be ruled out. First, the principal has no budget restriction, so he cannot (as agent 1 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 2.[14] We now show how this can make secret messages useless.

Consider a mechanism where at stage 2 agent 2 reports agent 1's effort: $M'_2 = \{work, shirk\}$. Let $(w_1^s(m), w_1^f(m))$ denote the expected wage payments to agent 1 in the success and failure states respectively, conditional on message $m \in \{work, shirk\}$. Suppose at stage 2, the principal and agent 2 can collude. In equilibrium, agent 1 works, and agent 2 tells the

[14] Agent 2 sees e_1 and the principal does not, so collusion between agent 2 and the principal could potentially involve asymmetric information as in Laffont and Martimort [8]. 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 1's wages are not minimized along the equilibrium path, there will be a collusive contract which the principal can propose to agent 2, and which agent 2 will accept, involving changing the messages so that agent 1's wages are minimized. Such a proposal has no "signalling" effects because the principal has no private information.

truth about agent 1's effort. Agent 1's expected income in equilibrium is, therefore,

$$p_{11} w_1^s(work) + (1 - p_{11})\, w_1^f(work). \tag{22}$$

The moral hazard constraint for agent 1 is

$$p_{11} w_1^s(work) + (1 - p_{11})\, w_1^f(work) - c_1 \geqslant p_{01} w_1^s(shirk) + (1 - p_{01})\, w_1^f(shirk). \tag{23}$$

At stage 2, the principal can propose that agent 2 reports $m_2 = shirk$ if this minimizes the wage-payments to agent 1. This makes the principal and agent 2 jointly better off, and there is always a bribe from the principal to agent 2 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 2 has an incentive to work. Agent 1's expected wage becomes

$$p_{11} w_1^s(shirk) + (1 - p_{11})\, w_1^f(shirk). \tag{24}$$

For agent 2 and the principal not to make this deal, it must be the case that $m_2 = work$ actually minimizes agent 1's expected wage:

$$p_{11} w_1^s(shirk) + (1 - p_{11})\, w_1^f(shirk) \geqslant p_{11} w_1^s(work) + (1 - p_{11})\, w_1^f(work). \tag{25}$$

Now consider the smallest expected wage payment (22) agent 1 can receive, subject to (23) and (25). We may set $w_1^f(shirk) = 0$, for if $w_1^f(shirk) > 0$ then we can reduce $w_1^f(shirk)$ and increase $w_1^s(shirk)$ while keeping

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

constant. Then, (23) will still be satisfied, while (25) now holds with strict inequality because $p_{11} > p_{01}$. Thus, set $w_1^f(shirk) = 0$. Then (25) implies

$$w_1^s(shirk) \geqslant w_1^s(work) + \frac{1 - p_{11}}{p_{11}}\, w_1^f(work). \tag{26}$$

Substituting (26) in (23) yields

$$\begin{aligned} p_{11} w_1^s(work) + (1-p_{11})\, w_1^f(work) - c_1 &\\ \geqslant p_{01} w_1^s(shirk) &\\ \geqslant p_{01} \left(w_1^s(work) + \frac{1-p_{11}}{p_{11}}\, w_1^f(work) \right). & \end{aligned}$$

Rearranging, we find that the expected wage payment to agent 1 satisfies

$$p_{11} w_1^s(work) + (1-p_{11})\, w_1^f(work) \geqslant p_{11} \frac{c_1}{p_{11} - p_{01}}.$$

Agent 2's moral hazard constraint must also be satisfied, which implies

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

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

$$w_1^s = \frac{c_1}{p_{11} - p_{01}}, \qquad w_1^f = 0$$

$$w_2^s = \frac{c_2}{p_{11} - p_{10}}, \qquad w_2^f = 0$$

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

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

This result, which shows that again decentralization is optimal, is intuitive. For messages to be useful, agent 2's message must be used to punish agent 1 if he shirks. Punishment means lower wages for agent 1, but then the principal can convince agent 2 to always claim agent 1 shirked. If the principal cannot commit not to collude with agent 2 in this way, then 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

Since 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 receive a certain guaranteed payoff, the principal makes sure that the distribution of surplus is the right one. In fact, 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 this result may be sensitive to the precise assumptions we made in the paper.

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

$$(w_i^f(m), w_i^s(m)) = \begin{cases} \left(0, \dfrac{\pi_i + c_i}{p_{11}}\right) & \text{if} \quad m_1 = m_2 = stay \\ (\pi_i, \pi_i) & \text{otherwise.} \end{cases} \tag{27}$$

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.[15]

LEMMA 9. *Suppose a mechanism* (M, w) *has a side contract* $\hat{c}$ *which implements full effort. Then, the option mechanism where wages are given by*

[15] We show later that in a simple contract the distribution of the wage packet between the two agents can perform a similar role as the outside option. Therefore, using our definition of implementation, delegation without any messages is optimal. However, notice that in the optimal option mechanism, the relative bargaining position of the agents is completely determined: the Pareto frontier for the agents contains *only* the payoff pair (π_1, π_2). With delegation, the Pareto frontier may contain several payoff pairs, and so full effort may not be the *unique* equilibrium side contract. However, unique implementation is not the focus of this paper.

(27), *with* $\pi_i = \pi_i(\hat{c})$, *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 no side payments and the expected wage payments are the same as in the equilibrium side contract* $\hat{c}$ *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

$$\pi_1(\hat{c}) = p_{11}(w_1^s(\hat{m}) - \hat{t}^s) + (1 - p_{11})(w_1^f(\hat{m}) - \hat{t}^f) - c_1 \tag{28}$$

$$\pi_2(\hat{c}) = p_{11}(w_2^s(\hat{m}) + \hat{t}^s) + (1 - p_{11})(w_2^f(\hat{m}) + \hat{t}^f) - c_2 \tag{29}$$

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

$$(\bar{w}_1^f(m), \bar{w}_1^s(m)) = \begin{cases} (w_1^f(\hat{m}) - \hat{t}^f, w_1^s(\hat{m}) - \hat{t}^s) & \text{if } m_1 = m_2 = stay \\ (\pi_1(\hat{c}), \pi_1(\hat{c})) & \text{otherwise} \end{cases} \tag{30}$$

$$(\bar{w}_2^f(m), \bar{w}_2^s(m)) = \begin{cases} (w_2^f(\hat{m}) + \hat{t}^f, w_2^s(\hat{m}) + \hat{t}^s) & \text{if } m_1 = m_2 = stay \\ (\pi_2(\hat{c}), \pi_2(\hat{c})) & \text{otherwise.} \end{cases} \tag{31}$$

Notice that expected wage payments in the new mechanism if both agents announce *stay* are the same as in the equilibrium side contract $\hat{c}$ of (M, w). We claim that $\bar{c} = (\bar{e}_1, \bar{m}, \bar{t})$, where $\bar{e}_1 = 1$, $\bar{m}_1 = \bar{m}_2 = stay$ 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 2 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 (28) and (29). Agent 1, of course, must work once he signs $\bar{c}$, since side contracts are binding. We only need to verify the conditions E1 and E2 of the definition of equilibrium side contract. 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. Condition 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 $m' = (stay, stay)$. 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 1 agrees to e_1', the agents send the original messages $\hat{m}$ (which gives the same wages as $m' = (stay, stay)$ 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 (i.e., it contradicts condition E2). Therefore, $(\bar{M}, \bar{w})$ implements full effort.

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

$$(w_i^f(m), w_i^s(m)) = \begin{cases} \left(0, \dfrac{\pi_i(\hat{c}) + c_i}{p_{11}}\right) & \text{if} \quad m_1 = m_2 = stay \\ (\pi_i(\hat{c}), \pi_i(\hat{c})) & \text{otherwise.} \end{cases} \tag{32}$$

Comparing (30)–(31) and (32), and using (28) and (29), we find that the option mechanism is identical to $(\bar{M}, \bar{w})$ if in the original mechanism (M, w) it is true that $w_1^f(\hat{m}) = w_2^f(\hat{m}) = 0$. Otherwise, the option mechanism differs by shifting all the wage payments to the success state, but the expected wages conditional on both working and announcing *stay* is the same as in $(\bar{M}, \bar{w})$ (which in turn is the same as in (M, w)). Clearly the option mechanism makes shirking less desirable for each individual agent. Since 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: $\bar{c} = (\bar{e}_1 = 1, \bar{m} = (stay, stay), \bar{t} = (0, 0))$. Then this side contract is feasible and 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

$$w_i \equiv w_i^s(stay, stay).$$

By the Lemma, we can focus on an equilibrium side contract of the form $\bar{c} = (\bar{e}_1, \bar{m}, \bar{t})$ with $\bar{e}_1 = 1$, $\bar{m} = (stay, stay)$ and $\bar{t} = (0, 0)$. The payoffs for the agents are (π_1, π_2) where

$$\pi_i = p_{11} w_i - c_i \geqslant 0.$$

The contract $\bar{c}$ must satisfy the conditions E1 and E2 of the definition of equilibrium side contract.

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

$$p_{11}(w_1 + w_2) < p_{11} \frac{c_1 + c_2}{p_{11} - p_{00}}. \tag{33}$$

We shall construct a new side contract $\tilde{c}$ which both agents strictly prefer to $\bar{c}$. To this end, define $\tilde{c} = (\tilde{e}_1, \tilde{m}, \tilde{t})$ where $\tilde{e}_1 = 0$, $\tilde{m} = (stay, stay)$, $\tilde{t}^f = 0$,

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

and $\varepsilon > 0$. Suppose under $\tilde{c}$ agent 2 shirks. Then

$$\pi_1(\tilde{c}) = p_{00}(w_1 - \tilde{t}^s) = p_{11} w_1 - c_1 + p_{00}\varepsilon > \pi_1 \tag{34}$$

and

$$\pi_2(\tilde{c}) = p_{00}(w_2 + \tilde{t}^s) = p_{00}(w_2 + w_1) - p_{11} w_1 + c_1 - p_{00}\varepsilon \tag{35}$$

Using (33),

$$\begin{aligned} \pi_2(\tilde{c}) - \pi_2 &= p_{00}(w_1 + w_2) - p_{11} w_1 + c_1 - p_{00}\varepsilon - (p_{11} w_2 - c_2) \\ &= c_1 + c_2 - (p_{11} - p_{00})(w_1 + w_2) - p_{00}\varepsilon > 0 \end{aligned} \tag{36}$$

for sufficiently small $\varepsilon > 0$. Thus, under the assumption that 2 shirks under $\tilde{c}$, both agents are strictly better off under $\tilde{c}$ than under $\bar{c}$. If agent 2 actually prefers to work under $\tilde{c}$, by revealed preference it must give him an even higher payoff than (35). Agent 2 working cannot hurt agent 1 either, as agent 1 makes no money if the project fails. Thus, in any case, both agents are strictly better off with $\tilde{c}$ than with $\bar{c}$. Finally, (34) and (36) imply that $w_1 - \tilde{t}^s \geqslant 0$ and $w_2 + \tilde{t}^s \geqslant 0$, so $\tilde{c}$ is feasible. Then $\bar{c}$ does not satisfy E2, a contradiction. ▮

7.2. *Proof of Proposition* 3

Proof. Suppose

$$\frac{p_{10}}{p_{11} - p_{10}} < \frac{p_{00}}{p_{01} - p_{00}}. \tag{37}$$

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 restrict attention to an option mechanism which implements full effort without side payments. Let

$$w_i \equiv w_i^s(stay, stay)$$

for $i \in \{1, 2\}$. Again we may focus on an equilibrium side contract $\bar{c} = (\bar{e}_1 = 1, \bar{m} = (stay, stay), \bar{t} = (0, 0))$ which satisfies conditions E1 and E2. Since $\bar{c}$ implements full effort, agent 2's moral hazard constraint (5) must be satisfied. Since there are no side payments and no wages in the failure state, this is equivalent to

$$w_2 \geqslant \frac{c_2}{p_{11} - p_{10}}. \tag{38}$$

Just as before, $\pi_i = p_{11} w_i - c_i \geqslant 0$ is agent i's equilibrium payoff in the equilibrium of the option mechanism.

Suppose in order to obtain a contradiction that the cost to the principal is

$$p_{11}(w_1 + w_2) < \frac{p_{11} c_1}{p_{11} - p_{01}} + \frac{p_{11} c_2}{p_{11} - p_{01}} \left[\frac{p_{10}}{p_{11} - p_{10}} - \frac{p_{00}}{p_{01} - p_{00}} \right]. \tag{39}$$

Consider the side contract $\tilde{c} = (\tilde{e}_1, \tilde{m}, \tilde{t}^f, \tilde{t}^s)$ with $\tilde{e}_1 = 0$, $\tilde{m} = (stay, stay)$, $\tilde{t}^f = 0$ and

$$\tilde{t}^s = w_1 - \frac{p_{11}}{p_{01}} \left(w_1 - \frac{c_1}{p_{11}} \right). \tag{40}$$

Note that (39) implies

$$w_1 + w_2 > \frac{p_{11}}{p_{01}} (w_1 + w_2) - \frac{c_1}{p_{01}} - \frac{c_2}{p_{01}} \left[\frac{p_{10}}{p_{11} - p_{10}} - \frac{p_{00}}{p_{01} - p_{00}} \right]. \tag{41}$$

It now follows that agent 2's moral hazard constraint is satisfied under $\tilde{c}$. Indeed, (41) and (38) imply

$$\begin{aligned} w_2 + \tilde{t}^s &= w_1 + w_2 - \frac{p_{11}}{p_{01}} \left(w_1 - \frac{c_1}{p_{11}} \right) \\ &> \frac{p_{11}}{p_{01}} w_2 - \frac{c_2}{p_{01}} \left[\frac{p_{10}}{p_{11} - p_{10}} - \frac{p_{00}}{p_{01} - p_{00}} \right] \geqslant \frac{c_2}{p_{01} - p_{00}} \end{aligned} \tag{42}$$

so that

$$(p_{01} - p_{00})(w_2 + \tilde{t}^s) > c_2, \tag{43}$$

which is precisely Eq. (5), since there are no payments in the failure state. Therefore, agent 2 works under $\tilde{c}$, while agent 1 shirks under $\tilde{c}$ by construction. Agent 1's payoff under $\tilde{c}$ is the same as in the original contract:

$$p_{01}(w_1 - \tilde{t}^s) = p_{11} w_1 - c_1 \geqslant 0. \tag{44}$$

Combining (39) and (37) we get

$$p_{11}(w_1+w_2)-c_1<p_{01}(w_1+w_2). \tag{45}$$

By (45), agent 2's payoff under $\tilde{c}$ is

$$p_{01}(w_2+\tilde{t}^s)-c_2>p_{11}w_2-c_2\geqslant 0. \tag{46}$$

Now (44), (46), and (43) imply that $\tilde{c}$ is an equilibrium side contract which makes both agents weakly better off than the contract $\bar{c}$. Also, (44) and (46) imply the proposed transfer under $\tilde{c}$ is feasible. Moreover, the inequalities in (46) and (43) are strict. Therefore, by reducing the transfer to $\tilde{t}^s-\epsilon$ for $\epsilon>0$ small, we can make both agents strictly better off than at $\bar{c}$ (Eq. (46) implies $\tilde{t}^s>0$). Therefore, $\bar{c}$ does not satisfy E2, which is a contradiction. ▮

7.3. *Proof of Proposition* 4

Proof. Suppose

$$\frac{p_{10}}{p_{11}-p_{10}}\geqslant\frac{p_{00}}{p_{01}-p_{00}}. \tag{47}$$

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

$$w_i\equiv w_i^s(stay, stay).$$

Again we may focus on an equilibrium side contract $\bar{c}=(\bar{e}_1=1, \bar{m}=(stay, stay), \bar{t}=(0,0))$ which satisfies conditions E1 and E2, together with agent 2's moral hazard constraint (38). Again, $\pi_i=p_{11}w_i-c_i\geqslant 0$ is agent i's equilibrium payoff.

Suppose in order to obtain a contradiction that

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

Then,

$$p_{11}(w_1+w_2)-c_1-c_2<p_{01}(w_1+w_2)-c_2. \tag{49}$$

That is, if agent 1 shirks (and the agents stick to the messages (*stay*, *stay*)), the sum of the agents expected payoffs is strictly increased. Let

$$\hat{t}^s = \frac{p_{11} w_2}{p_{01}} - w_2. \tag{50}$$

We claim that this positive transfer fulfills the moral hazard condition for agent 2, Eq. (5), conditional on $e_1 = 0$. Since there are no payments in the failure state, this condition is

$$(p_{01} - p_{00})(w_2 + \hat{t}^s) \geqslant c_2. \tag{51}$$

To see that this holds, notice that if (51) does not hold, then using (38) and the definition of $\hat{t}^s$,

$$\frac{p_{10} c_2}{p_{11} - p_{10}} = \frac{p_{11} c_2}{p_{11} - p_{10}} - c_2 \leqslant p_{11} w_2 - c_2 = p_{01}(w_2 + \hat{t}^s) - c_2$$
$$< \frac{p_{01} c_2}{p_{01} - p_{00}} - c_2 = \frac{p_{00} c_2}{p_{01} - p_{00}}. \tag{52}$$

But this contradicts (47). Thus (51) holds.

Now for small $\epsilon > 0$ consider the feasible side contract $\tilde{c} = (\tilde{e}_1, \tilde{m}, \tilde{t})$ where $\tilde{e}_1 = 0$, $\tilde{m} = (stay, stay)$, $\tilde{t}^f = 0$ and $\tilde{t}^s = \hat{t}^s + \epsilon / p_{01}$. We have

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

by (51) so agent 2 will work. Moreover,

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

and for small ϵ,

$$\pi_1(\tilde{c}) = p_{01}(w_1 - \hat{t}^s - \epsilon / p_{01}) > p_{11} w_1 - c_1 = \pi_1 \tag{53}$$

by (49). Therefore, both agents strictly prefer $\tilde{c}$ to $\bar{c}$. By (53), $\tilde{t}^s < w_1$ so $\tilde{c}$ is feasible. Therefore, $\bar{c}$ cannot satisfy E2, 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* (54) *holds*:

$$x \leqslant p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1. \tag{54}$$

A sufficient condition for implementation of full effort is that (54) *holds and* $e = (1, 1)$ *maximizes the sum of the agents expected payoffs* $\pi_1 + \pi_2$ *subject to* $\pi_i \geqslant 0$ *for* $i = 1, 2$.

Proof. First, suppose full effort is implemented by some equilibrium side contract $c = (e_1, t)$, but (54) does not hold. Then by E1,

$$\pi_1(c) = p_{11}(w_1 - t) - c_1 \geqslant x > p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1. \tag{55}$$

But (55) implies

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

Then 2's moral hazard constraint is not satisfied so c will not induce him to work, a contradiction. Thus, (54) is necessary.

Now suppose (54) holds and $e = (1, 1)$ maximizes the sum of the agents expected payoffs $\pi_1 + \pi_2$ subject to $\pi_i \geqslant 0$ 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 2 gets

$$\begin{aligned} w_2 + t &= w_1 + w_2 - \frac{c_1 + x}{p_{11}} \\ &\geqslant w_1 + w_2 - \frac{c_1}{p_{11}} - \frac{1}{p_{11}}\left(p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1\right) = \frac{c_2}{p_{11} - p_{10}} \end{aligned}$$

so that 2'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 $\pi_2(c) \geqslant y$, 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

$$p_{11}(w_1+w_2)-c_1-c_2 \geqslant p_{01}(w_1+w_2)-c_2 \tag{56}$$

$$p_{11}(w_1+w_2)-c_1-c_2 \geqslant p_{10}(w_1+w_2)-c_1 \tag{57}$$

$$p_{11}(w_1+w_2)-c_1-c_2 \geqslant p_{00}(w_1+w_2), \tag{58}$$

where (56) and (58) use (14). Thus, the sum of the agents payoffs is greater for $e=(1, 1)$ than for any other effort levels. We have

$$p_{11}\left(w_1+w_2-\frac{c_2}{p_{11}-p_{10}}\right)-c_1=0 \tag{59}$$

If $w_1=0$ then the most agent 1 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 1 at least $p_{00}w_1>0$. Since (59) holds, this contradicts Lemma 10. ∎

7.5. *Proof of Proposition* 6

Proof. 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

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

Just 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

$$p_{11}\left(\frac{c_1+c_2}{p_{11}-p_{00}}-\frac{c_2}{p_{11}-p_{10}}\right)-c_1\geqslant 0 \tag{60}$$

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

If $w_1=0$, then the most agent 1 can get in a subgame perfect equilibrium of $\Gamma(w_1, w_2)$ is zero. By (60) and Lemma 10, full effort is implemented.

Now suppose there is strict inequality in (15), and hence in (60). Then by Lemma 10, it is possible to set $w_1>0$, as long as player 1's payoff in subgame perfect equilibrium of $\Gamma(w_1, w_2)$ does not exceed the left hand side of the expression in (60). (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

Proof. Suppose (16) and (10) 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 (10). Notice that $w_2<w^*$, for if not then $w^*<c_2/(p_{01}-p_{00})$ so that $p_{01}w^*-c_2<p_{00}w^*$. But (20) 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 2 will work iff agent 1 has worked, but agent 1 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 (61)–(64) hold:

$$p_{11}(w_1 - t) - c_1 \geqslant p_{00} w_1 \tag{61}$$

$$p_{11}(w_2 + t) - c_2 \geqslant p_{00} w_2 \tag{62}$$

$$(p_{11} - p_{10})(w_2 + t) \geqslant c_2 \tag{63}$$

$$p_{11}(w_1 - t) - c_1 \geqslant p_{01}\left(w_1 + w_2 - \frac{c_2}{p_{01} - p_{00}}\right). \tag{64}$$

Equation (64) is the condition which guarantees that the contract where both agents work and agent 1 pays t to agent 2 in case of success is not Pareto-dominated by some side contract where only agent 2 works. Indeed, to make agent 2 willing to work alone, he needs the efficiency wage $c_2/(p_{01} - p_{00})$, but then (64) implies that agent 1 would be made worse off.

Using the definition of w^*, one can check that (61)–(64) 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

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

Set

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

and

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

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 (64) with $t = 0$, which can be written as

$$w_1 = w^* - \frac{c_2}{p_{11} - p_{10}} \geqslant \frac{p_{01}}{p_{11}}\left(w^* - \frac{c_2}{p_{01} - p_{00}}\right) + \frac{c_1}{p_{11}}. \tag{65}$$

But (65) holds by definition of w^*. Thus full effort is implemented.

Finally, agent 2 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 (20) implies that agent 1's contribution $(p_{11} - p_{01})\, w^*$ is smaller than his cost of effort c_1, so agent 2 will never work and pay agent 1 to work. ▌

7.7. *Proof of Proposition* 8

Proof. Suppose (16) holds but (10) 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 (16) that

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

and

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

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

$$p_{11}\left(w_1 + w_2 - \frac{c_2}{p_{11} - p_{10}}\right) - c_1 = p_{11}\left(\frac{c_1}{p_{11} - p_{01}} - \frac{c_2}{p_{11} - p_{10}} - \frac{c_1}{p_{11}}\right) > 0 \quad (66)$$

from (16). Lemma 10 implies that full effort is implemented if $w_1 = 0$. But as there is strict inequality in (66), it is possible to set $w_1 > 0$. (Just 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 wages, 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 2 reports on agent 1's effort level: $M_1^0 = M_2^0 = M_1' = \varnothing$, $M_2' = \{work, shirk\}$. For a fixed $\varepsilon > 0$, wages are given by

$$w_1^f(work) = w_1^s(shirk) = w_1^f(shirk) = w_2^f(shirk) = w_2^f(work) = 0$$

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

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

Since player 2'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 allows agent 1 to shirk and makes both agents better off, agent 1 must pay an *expected* transfer to agent 2 of at least

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

A transfer can only occur if agent 1 gets a non-zero wage, so $\bar{T} \leqslant \varepsilon t^s$. That is, agent 1 promises to pay $t^s \geqslant \bar{T}/\varepsilon$ whenever he has any money, in exchange for the right to shirk. To give agent 1 the proper incentive to pay whenever possible, agent 1 must suffer a cost $h \geqslant \bar{T}/\varepsilon$ 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. Suppose the agents sign the following side-contract: agent 1 sets $e_1 = 1$, agent 2 announces *work* and there are no transfers. If any agent does not conform to this contract, i.e., if agent 1 does not work or if agent 2 does not announce *work* even though agent 1 worked, the union punishes the cheating agent at a cost of $h \geqslant c_1$ (it is possible because the union can observe agent 1's effort level 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 2 and give him the incentive to announce *shirk* after agent 1 has in fact worked, she must pay 2 at least c_1 to counterbalance the punishment the workers' union will impose on agent 2. But then there is no incentive for the principal to collude with agent 2 as her total payments do not go down.

REFERENCES

1. American Institute of Architects, Building relationships, mimeo, 1990.
2. S. Baliga, Universally collusion-proof contracts, mimeo, Cambridge University, 1995.
3. S. Baliga, L. Corchon, and T. Sjöström, The theory of implementation when the planner is a player, *J. Econ. Theory* **77** (1997), 15–33.
4. S. Baliga and T. Sjöström, Decentralization and collusion, working paper, Harvard Institute of Economic Research, 1995.
5. Boston Society of Architects, The professional team concept, mimeo, 1973.
6. B. Holmström and P. Milgrom, Regulating trade among agents, *J. Inst. Theoret. Econ.* **146** (1990), 85–105.
7. H. Itoh, Coalitions, incentives and risk-sharing, *J. Econ. Theory* **60** (1993), 410–418.
8. J.-J. Laffont and D. Martimort, Collusion and delegation, mimeo, University of Toulouse, 1996.
9. C. Ma, Unique implementation of incentive contracts with many agents, *Rev. Econ. Stud.* **55** (1988), 555–572.
10. I. Macho-Stadler and J. Perez-Castrillo, Centralized and decentralized contracts in a moral hazard environment, working paper, University Autonoma de Barcelona, 1994.
11. I. Macho-Stadler and J. Perez-Castrillo, Moral hazard with several agents: The gains from cooperation, *Int. J. Ind. Organ.* **11** (1993), 73–100.
12. E. Maskin and J. Tirole, Dynamic programming, unforeseen contingencies and incomplete contracts, mimeo, Harvard University, 1996.
13. P. McAfee and J. McMillan, Organizational diseconomies of scale, mimeo, University of Texas at Austin, 1995.
14. N. Melumad, D. Mookherjee, and S. Reichelstein, A theory of responsibility centers, *J. Acc. Econ.* **15** (1992), 445–484.
15. N. Melumad, D. Mookherjee, and S. Reichelstein, Hierarchial decentralization of incentive contracts, *Rand J. Econ.* **26** (1995), 654–672.
16. D. Mookherjee, Optimal incentive schemes with many agents, *Rev. Econ. Stud.* **51** (1984), 433–446.
17. D. Mookherjee and S. Reichelstein, Budgeting and hierarchical control, mimeo, Boston University, 1995.
18. D. Mookherjee and S. Reichelstein, Incentives and coordination in hierarchies, mimeo, Boston University, 1995.
19. R. Ramakrishnan and A. Thakor, Cooperation vs. competition in agency, *J. Law Econ. Organ.* **7** (1991), 248–283.

20. J. Tirole, Hierarchies and bureaucracies: on the role of collusion in organizations, *J. Law Econ. Organ.* **2** (1986), 181–214.
21. J. Tirole, Collusion and the theory of organizations, *in* “Advances in Economic Theory: Sixth World Congress Vol. II” (J.-J. Laffont, Ed.), Cambridge Univ. Press, Cambridge, UK, 1992.
22. T. Van Zandt, Lecture notes on hierarchical incentive contracting, mimeo, Princeton University, 1995.