WORKING PAPER NO. 12

# COOPERATION IN HIERARCHICAL ORGANIZATIONS: AN INCENTIVE PERSPECTIVE

by

Hideshi Itoh

Faculty of Economics
Kyoto University

Faculty of Economics,
Kyoto University,
Kyoto, 606 JAPAN

WORKING PAPER NO. 12

# COOPERATION IN HIERARCHICAL ORGANIZATIONS: AN INCENTIVE PERSPECTIVE

by

Hideshi Itoh

Faculty of Economics
Kyoto University

March 1991

# COOPERATION IN HIERARCHICAL ORGANIZATIONS: AN INCENTIVE PERSPECTIVE*

BY

HIDESHI ITOH

Faculty of Economics, Kyoto University

Sakyō-ku, Kyoto 606, Japan

March 1991

## 1. Introduction

"Failures of cooperation" is one of the six recurring patterns of weakness in productivity performance in the U. S. pointed out by the MIT Commission on Industrial Productivity (Dertouzos, et al., 1989). The report discusses a lack of cooperation at various levels—within firms, between labor and management, in vertical relationships such as producer-consumer and producer-supplier linkages, and among firms in the same industry segment. For example, the commission ascribes the failure of cooperation within the firm to excessive specialization, multiple layers of bureaucracy, and little lateral flow of information. These features are the major ingredients of what was once considered (and has still been considered by many) as the most effective organizational structure, hierarchy based on extensive division of labor: The production process there is divided into many distinct tasks, each of which is made sole responsibility of a specialist who is only adept to the performance of that task, and coordination among tasks is a specialized job of upper management. The deviation from such an organizational structure seems evident, however. In the field of human resource

* Earlier versions were presented at the Fourth Conference on Game Theory and Mathematical Economics held at Gotenba, Japan, and seminars at INSEAD and Kyoto University. I wish to acknowledge my debt to Bengt Holmstrom and Paul Milgrom whose insightful and tractable linear principal-agent model is used throughout the paper for the illustration of my arguments. I would also like to thank Jean Tirole for helpful discussion and the Center for Economic Policy Research at Stanford University for financial support.

management, for example, movement from individualistic workplace to cooperative team-based organizations has been repeatedly discussed in the press (e.g., Hoerr, 1989) as well as in the academic literature (*e.g.*, Blinder, 1990, Nalbantian, 1988). Furthermore, more extensive, corporate reform in American manufacturing (Piore, 1989) and in world automobile industries (Roos, et al., 1990) has been witnessed.

The purpose of this essay is to ask the following question and to attempt to give some empirically testable predictions from incentive viewpoints: Is "cooperation" among some, but not all, members of an organization (e.g., employees) good for the organization as a whole (e.g., for the employer)? If yes, then when? Some readers might think that obviously the answer would be always yes. It is trivially true to say that cooperation by *all* the members is desirable to the whole organization. Most of the existing theoretical research on organizations has thus focused on how cooperation among self-interested members can be attained (for example, via reputation). However, it is not obvious to answer whether cooperation by a subset of members such as workers in lower tiers of hierarchy is beneficial to top management. In fact, the current trend in incentive theory appears to emphasize "competition" as incentive devices. For example, consider the typical principal-agent model with moral hazard. Cooperation between a principal (an organization designer) and an agent (her subordinate who performs certain tasks for her) is the focus of the model: the principal wants the agent to cooperate with her. However, when there exist many agents, say two agents, the major theoretical result is the optimality of *relative performance evaluation*: if there exist systematic risks so that verifiable noisy performance measures of the agents' unobservable actions are positively correlated with one another, the optimal (second-best) contract pays each agent contingent on his performance relative to the others.[1] If, in addition, higher performance measures signal higher efforts,[2] then under the optimal contract with relative performance evaluation, each agent is paid less the higher the performance measure of the other agent is. This is because the latter's better performance implies favorable environments for the first agent, and hence it should be discounted from his pay. Most readers are probably familiar with a rank-

[1] More precisely, this is true only if (and, with some additional assumption, if) an agent's performance measure $x$ is not a sufficient statistics of $(x, y)$ where $y$ is the performance measure of the other agent (Holmström, 1982, Mookherjee, 1984).

[2] That is, the performance signals satisfy MLRP (monotone likelihood ratio property). See Milgrom (1981).

order tournament, which is the extreme form of relative performance evaluation. Rank-order tournaments have been extensively discussed by economists and applied to various situations. (See, for example, Mookherjee (1988) for a survey.)

The typical model from which the optimality of relative performance evaluation is derived assumes that there is no production externality among agents. However, it is clear that once their interaction in production is permitted, they are only interested, under relative performance evaluation, in reducing the probability that the others get good performance measures. In fact, Lazear (1989) analyzes such a model under tournament schemes and shows that pay compression between the winner and the loser may be preferable in order to reduce "sabotage" by the agents.

If production processes are significantly interrelated as above, then why not modify the contract so that each agent appreciates high performance by other agents, rather than stick to relative performance evaluation? A team contract is such an example under which each agent is paid contingent on team performance rather than individual performance. And if positive production externalities like "help" are desirable from the technological point of view, the principal may want to provide the agents with monetary incentives to engage in helping each other on some tasks. We may call this sort of cooperation *induced cooperation*—cooperation induced by a grand contract designed by the principal. However, economists are generally not excited about the team-based or any other pay system resulting in induced cooperation because of the problem of individual motivation. In order for a self-interested agent to allocate some of his efforts for another agent, the principal must assign the former agent to *joint responsibility* for the latter's task: the risk associated with that task is borne by both agents. For example, under such a scheme, verifiable output from a machine is affected by the actions of the workers who hold joint responsibility for the output, rather than the action of a worker who is solely responsible to that machine under the extensive division of labor with individual-based or competitive pay schemes. Then the former regime appears to weaken the connection between pay and the effort of each individual agent, and hence is expected to give the agents greater incentives to shirk. As another illustrative example, consider two salespersons who are assigned to the same territory. Their cooperation may increase the sales and benefit the sales division, while encouraging cooperation will make the sales volume of each salesperson less informative as a performance measure for his effort. Because of this agency cost, the suggestion by incentive theorists in this example is again relative performance evaluation (assuming that the common territory has significant uncertain factors that affect the performance of both salespersons).

Such a motivational or "free-rider" problem in "teams" has been studied by economists in the framework of principal-agent problems (Alchian and Demsetz, 1972, and Holmström, 1982). The literature suggests the importance of monitoring by the principal: When the principal can only observe aggregate performance measures of the agents who are risk averse and have distastes for work, it is usually valuable for her to utilize individual measures for each agent's performance.[3] Of course, this argument does not directly apply to the situation given in the previous paragraph because cooperation there alters signals for individual actions under the unambiguous division of labor to signals for joint actions. What changes is *not* the availability of additional signals *but* the nature of existing signals. Thus, whether the net incentive effect of induced cooperation is in fact negative is not as clear as one might expect.

I show that the principal sometimes wants to induce cooperation among agents, even though the free-rider problem exists. The illustrative model used is the one developed by Holmstrom and Milgrom (1987, 1990b). Their 1990 paper presents a simple model of task allocation in which it is never optimal for two agents to be jointly responsible for any task: Each task is performed by just one agent, which is an important principle underlying hierarchy. With some modifications, I obtain an optimality of induced cooperation: it is optimal for an agent to help the other agent, and hence to be jointly responsible for each task despite the free-rider problem. The benefit from such induced cooperation is motivational: The first agent will reduce his effort on the task when he acquires some help, while the lower effort is less costly to induce when his cost of effort exhibits decreasing returns (e.g., monotonous tasks). This can be interpreted as an economic rationale of job enlargement and enrichment.

I then examine the relation of the optimality of this induced cooperation with stochastic correlation between task-specific performance measures. It is shown that there exists a threshold level of correlation coefficient such that induced cooperation is optimal if and only if the correlation coefficient is lower than that level. If the correlation coefficient is higher than the threshold level, the benefit from relative performance evaluation (filtering out systematic risks) dominates any benefit from induced cooperation, even if uncooperative behavior analyzed by Lazear appears and has negative effects on the principal's welfare. I also conduct comparative

[3] The precise condition for this to hold is similar to the one in footnote 1.

statics exercises, showing that induced cooperation is more likely to be optimal as the agents are less risk averse, or the tasks are similar in terms of performance measurability and costs of actions.

In our discussion so far, an agent "cooperates" with other agents if and only if he is induced to do so by the principal through design of an appropriate initial contract, and hence we have used the term induced cooperation.[4] This view may be called the *grand contracting approach*: "All members of the organization are linked by a grand contract, and their interaction is limited to procedures specified by this contract." (Tirole, 1988, p. 461). This approach, common in most of the literature on the principal-agent relationship, seems to be extreme: As argued by organization theorists for a long time and recently reemphasized by Tirole (1986, 1988), organization members often behave as a group, maximizing group welfare via some forms of side contracting, rather than behave independently as individuals. For example, most of the recent reviews of economic theories of organizations *by sociologists* (Baron, 1988, Granovetter, 1985, Perrow, 1986) criticize economists' emphasis on formal properties of organizations. They argue that informal aspects of organizations such as work norms and social relations among organizational members are no less important.

Tirole (1986), motivated by the sociological studies of organizations, considers the other polar case of *comprehensive contracting*, in which it is assumed that all side contracts among members are feasible. He assumes that a group of members can costlessly write any side contract *based on information commonly observable among them*. The original analysis of a three-tier organization of principal/supervisor/agent by Tirole demonstrates that the possibility of group behavior at a nexus of information (between the supervisor and the agent in his model) reduces the net payoff to the principal because of an opportunity for the supervisor and the agent to collude to manipulate their private information. Thus, in his example, the principal wishes to prohibit side trades if feasible. For example, she could do this by closing the communication channel between her and the supervisor, and by using rigid rules instead.

Side trading activities of agents do not necessarily lead to collusion. Some of the recent literature on labor and human resource management argues that the free-rider problem under work teams can be resolved via mutual monitoring and peer pressure. For example, Levine

[4] Agents may enjoy helping others, without any reward, up to some limit. If this is true, then this statement should read, "an agent cooperates beyond that limit if and only if ..."

and Tyson (1990, p. 187) writes, "Suppose workers are divided into work groups or teams on the basis of the interdependence of their work, pay is based on team output, and the teams help organize their work. By working together, team members recognize their mutual interests and observe how shirking by one can hurt the group. Shirking or free riding now imposes an observable cost directly on all co-workers, so that social sanctions may be rationally applied against workers who deviate from the cooperative work norm." That is, the principal may benefit by delegating to the agents the arrangement of cooperation among them. We may call this sort of cooperation *delegated cooperation* in contrast to induced cooperation discussed before.

Based on the recent work by Holmstrom and Milgrom (1990a), Itoh (1990), and Ramakrishnan and Thakor (1991), we explain when and how side trading leads to delegated cooperation, using the same illustrative model of the principal/two agents relationship with moral hazard as above. These papers, as well as Varian (1990), show that the most important factor is the agents' monitoring capabilities. When they do not share any private information, or in other words, what they can commonly observe is also observable to the principal, delegated cooperation has no value to the principal. On the other hand, when they can monitor each other's actions perfectly, delegated cooperation turns out to be valuable even in the case where no agent can affect the other agents' performance measures (via help as in the previous discussion). Itoh (1990) shows this in a general Grossman-Hart (1983) type model under the assumption that the error terms are independent. Holmstrom and Milgrom show in the same model as the one used here that in such a case, there again exists a cut-off level of the correlation coefficient between performance signals of two tasks such that the principal prefers side trades by the agents to no side trade if and only if the correlation coefficient is lower than that level.[5] Delegated cooperation under perfect mutual monitoring enables the principal to obtain appropriate efforts from the agents with less risks imposed on them than no cooperation. To do this, however, the principal must make the agents responsible to each other's outcomes (as in the case of induced cooperation discussed before) in spite of technological independence, which feature prevents the use of relative performance evaluation. The latter method is more valuable the higher the degree of correlation between the outcomes, and hence the result follows.

[5] A related result is found in Ramakrishnan and Thakor (1991) in a different model.

Itoh (1990) also examines the benefit of delegated cooperation under production externalities and team production in the general model. In this paper, I obtain related, but more transparent results from the special linear model. It is shown that the benefit of delegated cooperation can be realized if the correlation between the error terms are so small that the principal prefers induced cooperation to relative performance evaluation, and if the agents are sufficiently homogeneous in their risk attitudes and costs of actions. Thus, when the principal wishes to induce cooperation among agents, she would also like to encourage mutual monitoring and coordination of effort among them. Furthermore, I provide a theoretical justification of the argument by Levine and Tyson (1990) cited above concerning the role of mutual monitoring and sanctioning under team production. In addition, it is shown that under team production the principal may not need to hire a supervisor who can observe and report individual performance, if the agents side trade. The problem of the supervisor's "hidden gaming" analyzed by Laffont (1990) therefore does not arise under some conditions.

There is one important caveat on the analysis of side trading in this paper as well as in the literature mentioned above. The benchmark case of no side trade is the standard second-best solution attained by the optimal incentive contract. That is, it is assumed that the principal does not design more complex communication mechanisms for the implementation of actions. When actions are not mutually observable among agents, this restriction is without loss of generality (except for the issues of multiple equilibria as in Ma (1988)). However, when actions are observable among them, if agents do not side trade, the principal can attain the first-best outcome by an appropriate communication mechanism à la Ma (1988). Thus, some readers may argue that the benchmark in this case should be the first-best, and hence side trades do not improve the principal's welfare. I exclude this possibility by turning to the literature on incomplete contracts (e.g., Grossman and Hart (1986)): It is often reasonable to assume that the principal cannot write explicit contracts contingent on actions because it is hard to specify exactly what the actions are in a contract. On the other hand, since actions are mutually observable by agents, they can more easily contract on actions implicitly in their side trades. With this perspective, side trades may enlarge the set of feasible contracts.[6]

Note that such communication mechanisms would play no role *in the world of side trading* even if explicit contracts on the agents' reports about their actions were feasible. There,

[6] I am grateful to Jean Tirole for suggesting this.

once the communication channels opened, the agents would collude in communication stages. Thus, there is no collusion-proof mechanism improving the principal's welfare from delegated cooperation with no communication (Itoh, 1990). The principal thus would not attempt to centralize the information about the agents' actions, and in this respect, the regime considered here is truly delegation.

The rest of the paper is organized as follows. Sections 2 and 3 focus on induced cooperation and delegated cooperation, respectively, and present the results mentioned above. Section 4 are concluding remarks. Appendix contains all the proofs.

## 2. Induced Cooperation

In this section, I explain two things: (i) the principal's incentive to induce her agents to cooperate on some tasks, in the sense of productive interaction such as mutual help, when there exists a free-rider problem; (ii) a tradeoff between the induced cooperation and relative performance evaluation. It is assumed in this section that only the principal designs grand contracts and no party engages in side contracting.

*Illustrative model*: I use the following illustrative model of a principal-multiagent problem with moral hazard, due to Holmstrom and Milgrom (1987, 1990b). There are two agents labeled A and B. They select, respectively, inputs ("efforts") $a$ and $b$ for production, with monetary costs $C^A(a)$ and $C^B(b)$. The efforts are unobservable to the principal. The agents are (strictly) risk averse with preferences represented by the exponential utility functions with the coefficients of absolute risk aversion denoted by $r_A$ and $r_B$.[7] The principal is risk neutral.

Throughout the paper, I assume there are two tasks labeled 1 and 2. Task $i$ yields an uncertain payoff $x_i$, which is, unless otherwise noted, publicly observable and hence used as a performance measure in contracts the principal offers. The principal's total payoff before payments to the agents is thus $x_1 + x_2$. Each agent possibly provides inputs to both tasks: Let $a = (a_1, a_2)$ and $b = (b_1, b_2)$ where $a_i$ and $b_i$, the inputs to task $i$, are real numbers for $i = 1, 2$. They are normalized such that the least costly efforts are zero vector for each agent. It is assumed that the payoff from task $i$ ($i = 1, 2$) depends on the agents' inputs and an exogenous random factor as follows:

$$x_i = \mu_i(a_i, b_i) + \epsilon_i$$

[7] Agent A's utility is thus $-\exp[-r_A(w - C^A(a))]$ where $w$ is his income.

where $\epsilon_i$ is Normally distributed with mean zero and covariance matrix $\Sigma$. Let $\sigma_i^2 > 0$ be the variance of $\epsilon_i$ and $\rho$ be the correlation coefficient ($0 \leq \rho \leq 1$). The expected payoff $\mu_i$ is assumed to be twice differentiable, strictly increasing, and (weakly) concave.

The principal designs incentive contracts $(a, b; w_A, w_B)$ which specify the agents' actions and performance-contingent payments to them. It is assumed that payment schemes are linear in the performance measures:[8]

$$w_A(x_1, x_2) = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_0 \qquad \text{and} \qquad w_B(x_1, x_2) = \beta_1 x_1 + \beta_2 x_2 + \beta_0.$$

The optimal contract maximizes the certainty equivalent of the joint surplus of the three parties subject to the *participation constraints* which guarantee some minimum levels of utility for the agents, and the *incentive compatibility constraints* which ensure that the agents, behaving independently, follow the instructions by the principal, and hence their choice forms a Nash equilibrium.[9] Since the fixed salary components $\alpha_0$ and $\beta_0$ simply play a role of surplus transfer among them, I can ignore them and focus on the choice of the share parameters $\alpha = (\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$.

*An optimality of joint responsibility*: The share parameters are determined to balance among risk allocation, effort incentives, and effort allocation. The detailed study of the tradeoff among these facets is found in Holmstrom and Milgrom (1990b). Most of their analyses are conducted under the assumption that the efforts are perfect substitutes in the agent's costs and the payoffs from the tasks.[10] The assumption that the costs depend only on the total effort implies that an increase in an agent's input to one task, *however small*, induces reduction in his input to the other task. Holmstrom and Milgrom show under this assumption that when tasks are "small," in the sense of a continuum of tasks, and the error terms of task-specific performance measures are independent, it is never optimal to assign both agents to *joint responsibility* of the same task: $a_i b_i = 0$ and $\alpha_i \beta_i = 0$ for $i = 1, 2$ should hold. The idea is simple. Joint responsibility on a task requires that the principal impose risk from that task

[8] See Holmstrom and Milgrom (1987) for justification of this assumption. The model used here is regarded as a reduced form of their dynamic model.

[9] For example, the certainty equivalent of agent A is written as $\alpha_0 + \alpha_1 \mu_1 + \alpha_2 \mu_2 - C^A(a_1, a_2) - \frac{1}{2} r_A \alpha \Sigma \alpha$ where the last term is the agent's risk premium which is equal to the variance of his income. The formal optimization problem is presented in Appendix.

[10] That is, the cost functions and the expected payoffs are written as $C^A(a_1 + a_2)$, $C^B(b_1 + b_2)$, and $\mu_i(a_i + b_i)$ for $i = 1, 2$.

on both agents, which duplication is wasteful. The principal can reallocate the agents' efforts only to reduce the risk that the agents must bear, by making an agent *solely responsible* for each task.

Their result provides an incentive-based rationale of the extensive division of labor which is an important feature of hierarchical organizations. In contrast, however, it is more and more frequently argued that job enlargement and enrichment through regular job rotation among different tasks make worker motivation keep high even though each of the tasks would be monotonous and boring. Here I show this statement formally. I assume that there are two "big" tasks, the task-specific performance measures have independent error terms ($\rho = 0$), the expected payoff from each task depends only on the total input ($\mu_i(a_i, b_i) = \mu_i(a_i + b_i)$), and it is strictly increasing and strictly concave. This last assumption implies that an agent, who chooses a positive effort on a task, reduces the effort when the other agent increases his effort on the same task.[11] In addition, throughout the paper, I assume the following:

$$C^A(a) = c_{A1}(a_1) + c_{A2}(a_2);$$
$$C^B(b) = c_{B1}(b_1) + c_{B2}(b_2)$$

where each term is twice differentiable, strictly convex, and has the first derivative equal to zero at zero effort.[12] The cost functions reflect decreasing returns to each activity. What is important here is that making an agent work a little bit for a task does not require effort away from the other task. This is in contrast to the attention allocation case considered by Holmstrom and Milgrom. I adopt this assumption in order to focus on effects of an agent's effort for a task on the other agent's effort for the same task. Then the following holds.

PROPOSITION 1. *Under the assumptions given above, it is optimal for the principal to make the two agents jointly responsible for each of the two tasks.*

The proposition asserts that when tasks exhibit decreasing returns, joint responsibility may be desirable, even though the agents free ride on each other, because it reduces incentive costs. To clarify this, consider a situation in which agent B has sole responsibility for task 2.

[11] For example, agent B's reaction function has the slope the sign of which is equal to that of $\beta_2 \mu_2''(a_2 + b_2)$. It is negative by the assumptions if $\beta_2 > 0$.

[12] These assumptions are much more special than are needed. In particular, the additive separability is not essential. See Itoh (1991a) for a general case in which I use a Grossman-Hart (1983) type model of the principal/two agents relationship.

Suppose that the principal assigns agent A to small responsibility for task 2 by increasing the share parameter $\alpha_2$ from zero. Agent A then raises his effort $a_2$ from zero. This process costs very small by assumption. The only effect to be examined is how the increase in agent A's effort on task 2 affects agent B's effort on the same task. As mentioned above, agent B *reduces* his effort $b_2$ when $a_2$ increases. However, the decrease in $b_2$ has the other effect of reducing costs of inducing agent B to select $b_2$ (by making the incentive compatibility constraints less tight), which effect dominates the free-rider effect. The principal thus prefers inducing the agents to cooperate on each of the tasks, which regime is called *induced cooperation*.

The stark difference between this result and Holmstrom and Milgrom (1990b)'s mainly comes from the difference in cost structures as mentioned above. In the example after Proposition 1, if the cost functions depend only on the total efforts, inducing agent A to work on task 2 requires a fixed cost of risk-bearing. Unless this fixed cost is sufficiently small, the result does not directly apply in such a case.[13]

*Induce cooperation and relative performance evaluation*: In the result presented above, the error terms in the performance measures are assumed to be independent. When there exist systematic risks so that the error terms are positively correlated, the principal can filter out systematic risks by comparing agents with each other. In particular, when she can prevent them from engaging in unproductive interaction, the optimal contract in fact has such a feature (e.g., Holmström, 1982). To see this point in the model analyzed here, following Holmstrom and Milgrom (1990a), suppose $\mu_1(a_1, b_1) = a_1$ and $\mu_2(a_2, b_2) = b_2$. Then the optimal share rates for agent A are given by

$$\alpha_1 = (1 + r_A \sigma_1^2 (1 - \rho^2) c''_{A1})^{-1} \qquad \text{and} \qquad \alpha_2 = -\alpha_1 \rho \frac{\sigma_1}{\sigma_2}. \tag{$*$}$$

Note that $\alpha_1$ is always positive and $\alpha_2$ is negative with positive correlation ($\rho > 0$). This is because higher performance measure for task 2 implies more favorable environments for agent A, and hence it should be discounted from his pay. Parameter $\alpha_2$ is simply determined to minimize the risk premium of agent A given incentives provided by $\alpha_1$.

This method and induced cooperation are clearly incompatible: if interaction between the agents is allowed under relative performance evaluation, each agent is only interested in

[13] See Itoh (1991a) for a formal derivation of this result.

reducing the other's performance measure (Lazear, 1989). What determine the principal's preferences between the two methods?

I examine this comparison under a simplifying assumption of separable technology as follows: $\mu_i(a_i, b_i) = a_i + b_i$ for $i = 1, 2$. The share parameters of each agent are then determined independently of those of the other agent. Actions $a_i$ and $b_i$ can be positive or negative: Positive efforts are "help" as before and negative efforts are "sabotage" as analyzed by Lazear (1989). Both types of actions are assumed to be costly: $c_{Ai}(e)$ and $c_{Bi}(e)$ increase as $e$ moves away from zero to either direction. For simplicity, $c''_{Ai}$ and $c''_{Bi}$ are assumed to be constant for $i = 1, 2$. To unify Lazear's analysis of relative performance evaluation with sabotage and my analysis of induced cooperation, I first assume that the principal cannot restrict interaction between agents. Use of relative performance evaluation then accompanies sabotage. Later I consider the case where she can restrict their uncooperative behavior completely.

In the following proposition, $m_{Ai}$ is defined by $m_{Ai} = \sigma_i c''_{Ai}$ and is called agent A's *efficiency-loss measure* on task $i$. It measures the difficulty of providing agent A with incentives to be productive on task $i$ due to the moral hazard problem: $m_{Ai}$ is higher the more difficult to measure the performance at task $i$ is or the more costly it is to induce agent A to increase his effort on the task.

PROPOSITION 2. *Suppose $m_{A2} \geq m_{A1}$. Then the following hold for agent A:* (a) *it is always optimal to induce agent A to be productive ($a_1 > 0$) on task 1;* (b) *it is always optimal to induce agent A to be productive ($a_2 > 0$) on task 2 as well if $m_{A2} - m_{A1} < (r_A\sigma_1)^{-1}$ holds;* (c) *if $m_{A2} - m_{A1} > (r_A\sigma_1)^{-1}$, there exists a threshold level of the correlation coefficient, denoted by $\rho^* < 1$, such that induced cooperation is optimal ($\alpha_2 > 0$) if and only if $\rho < \rho^*$, and relative performance evaluation is optimal ($\alpha_2 < 0$) if and only if $\rho > \rho^*$. Similar results hold for agent B.*

It has been shown in Proposition 1 that induced cooperation is optimal when there is no correlation in performance measures. Proposition 2 says that the optimality of induced cooperation continues to hold for any level of the correlation coefficient if two tasks are sufficiently similar in terms of moral hazard measures. When they differ sufficiently with regard to the difficulty of providing incentives, the principal prefers agent A to perform both tasks only if the correlation coefficient is lower than the cut-off level. If the correlation is higher, the principal only induces agent A to work on the task which is easier to provide incentives (task 1 in the proposition), and uses the performance measure of the other task to reduce his exposure to risk, even though such a regime introduces his unproductive behavior on the latter task.

The next proposition presents results from comparative statics exercises on the threshold value $\rho^*$.

PROPOSITION 3. *Suppose $m_{A2} - m_{A1} > (r_A\sigma_1)^{-1}$. The threshold level of the correlation coefficient $\rho^*$ for agent A is decreasing in $(r_A, \sigma_2^2, c''_{A2})$ and increasing in $c''_{A1}$. It is first decreasing and then increasing in $\sigma_1^2$. Similar results hold for agent B.*

The proposition implies that if rewarding agent A by relative performance is optimal for some fixed parameters $(r_A, m_{A2})$, then it is also optimal for higher values of these parameters. Similarly, if inducing agent A to cooperate on both tasks is optimal, then it is also optimal for lower values of these parameters. These are intuitive results: Relative performance is more valuable the more risk averse the agent is or the more difficult it is to provide incentives to work on the less efficient task (task 2). The effect of the efficiency-loss measure of the more efficient task (task 1) is different. Induced cooperation is more likely to be optimal the more costly it is to induce agent A to be productive on task 1, because it is then relatively easier to induce him to be productive on task 2. The same effect exists for the variance of the error in the task 1 performance measure if it is sufficiently high. There is another effect, however. Higher variance implies that it is more difficult to measure the performance on task 1, and hence the use of information contained in the task 2 performance measure via relative performance evaluation is more valuable. This is the reason for ambiguity in the effect of the noisiness of the task 1 measure.

I next turn to the case in which the principal can restrict interaction between agents. If she wishes to utilize relative performance as incentive schemes, it is obvious that she wants to limit the agents' interaction in order to prevent unproductive sabotage. For example, in the traditional mass manufacturing factory, high in-process inventories make each operation done in isolation. Or two managers who are candidates for promotion could be assigned separately to an office in California and an office in Massachusetts. The possibility of the isolation may increase relative merits of the competitive incentive schemes.

The basic arguments do not change, however. The optimal share rates of agent A who specializes in task 1 under relative performance evaluation have already been given in ($*$). Furthermore, it can be shown that the net payoff to the principal under this scheme is increasing while the net payoff under induced cooperation is decreasing in the correlation coefficient. When there is no correlation, induced cooperation is better by Proposition 1. If errors are perfectly correlated, relative performance evaluation achieves the first best, and hence is better. These arguments result in the next proposition.

PROPOSITION 4. *There exists a cut-off level of the correlation coefficient, denoted by $\rho^{**} < 1$, such that induced cooperation is optimal for $\rho < \rho^{**}$ and relative performance evaluation (without sabotage) is optimal for $\rho > \rho^{**}$. Furthermore, $\rho^{**} < \rho^{*}$ holds.*

Note that in the case where the principal can limit interaction between agents, induced cooperation cannot be optimal for all values of the correlation coefficient, even if the two tasks are similar in the difficulty of inducing the agents to be productive. The comparative statics exercises are harder in this case, but it appears that the same results as those in Proposition 3 will hold.

## 3. Delegated Cooperation

One of the recent criticisms about the standard principal-multiagent analyses, which include the one in the previous section of the current paper, is that the relationship is characterized by a single grand contract designed by the principal.[14] Once the contract is accepted by agents, they behave independently of each other. In practice, they sometimes form a coalition and engage in side trades that are not directly controllable to the principal because she cannot observe them. Successful side trades cause the agents to behave as a group. Informal aspects of organizations, such as work norms and social relations, can be well represented by such group behavior. The importance of such informal features to organizational design has been emphasized by sociological theorists.[15]

Following Tirole (1986), in this section, I take an alternative approach: it is assumed that a group of members can costlessly write any side contract *based on information commonly*

[14] See, for example, Tirole (1986, 1988).
[15] See, for example, Baron (1988), Granovetter (1985), and Perrow (1986).

*observable among them.* These side contracts cannot be enforced explicitly when they are contingent on private information shared among coalitional members. This full-side-contract assumption thus can be rephrased as follows: it is assumed that a group of members, when forming a coalition, abide by their promises built on their common information, with probability one. This assumption is clearly extreme. However, it is also extreme to assume that promises that are not self-enforcing are respected with probability zero. As a first step, this paper adopts the extreme approach which is the other side of the traditional one.[16]

Given the full-side-contract assumption, Tirole (1986) analyzes a three- tier hierarchy of principal/supervisor/agent where the agent has private information on his productivity and his action, while the supervisor sometimes obtains evidence on the true productivity. He shows that although the principal can prevent the agent and the supervisor from forming a coalition and concealing the information on productivity by designing a coalition-free initial contract, the mere possibility of private trades reduces her net payoff from the one under no side trade: side trades lead to *collusion.* The principal thus wants to prevent them from side trading, if possible. For example, bureaucratic rules may be used instead of reliance on supervisory information.

Analyzing principal-multiagent relationships with moral hazard, Holmstrom and Milgrom (1990a), Itoh (1990), and Ramakrishnan and Thakor (1991) recently identify the case in which side trades increase the net payoff to the principal—the case where one may call side trading activities of the agents *cooperation.* I first summarize some of the major results from these papers as follows:

PROPOSITION 5. *When the agents cannot observe their actions $(a, b)$ each other (and hence they can observe public information $(x_1, x_2)$ only), the principal is never better off with side trades by the agents than the case of no side trade.*

PROPOSITION 6. *Suppose that production is technologically independent and the agents can monitor each other's efforts perfectly. Then there exists a positive threshold level of the correlation coefficient, denoted by $\bar{\rho} > 0$, such that the principal is better off with side trades by the*

[16] Alternatively, one could develop a repeated-game model of the principal-multiagent relationship to analyze side trades. A successful development of such a model would make it possible to analyze intermediate cases of promises followed with probability between zero and one. The costs of that approach are that they tend to be complex and messy, and we may not go very far with it. As a first step, the easier approach is taken in this paper.

*agents than the case of no side trade where relative performance evaluation is instead used, if and only if $\rho < \bar{\rho}$.*

If the agents can only observe the payoffs from their tasks, they will select their actions independently of each other, as in the previous case of no side trade. Then the only role of side contracting is that of mutual insurance. After accepting an initial contract offered by the principal, they will attain the optimal risk sharing between them via side contracting. They therefore select their efforts not based on their pay specified by the initial contract, but based on their final income after side trades. This distorts their effort choice. The principal can prevent them from engaging in reinsurance, without loss of profit, by incorporating this possibility into the initial contract. Because of the additional constraints, however, such reinsurance opportunities are never of value to the principal.[17]

The agents thus must share information not observable to the principal, in order for their side trades to be valuable to her. In the paper I consider the case where the agents can monitor each other's actions perfectly and write side contracts contingent on them. It is assumed that the agents select a Pareto optimal effort pair via side contracting. If such side contracting results in higher net payoff to the principal than the case of no side trade, I say that *delegated cooperation* is attained: The arrangement of cooperation is delegated to the agents who monitor each other's actions and enforce coordinated actions. Proposition 6 shows that in contrast to induced cooperation, the principal can enjoy delegated cooperation despite technological independence between the agents, if the correlation in the error terms is sufficiently small. Production is said to be technologically independent if $\mu_1$ depends on $a_1$ only and $\mu_2$ depends on $b_2$ only.

To understand the benefit of side trades contingent on actions, suppose that the error terms are stochastically independent. Then without side trade, the optimal contract is individual-based, that is, $\alpha_2 = 0$ and $\beta_1 = 0$ (e.g., Holmström, 1982). With side trades, the agents choose their efforts to maximize the sum of the certainty equivalents of their expected utilities.[18] Note that under individual-based schemes, the agents select the same efforts whether or not they side trade. Thus, *side trades are of no value under individual-based pay.*

[17] This observation is also found in Varian (1990). He points out that mutual insurance among agents may be beneficial to the principal if they share information about states of nature that are not available to the principal.

[18] Transferable utility exists in the model.

Now suppose that $(\alpha_1, \beta_2)$ are the optimal share rates when the agents do not side contract, and define new rates by

$$\hat{\alpha}_1 = \frac{r_B}{r}\alpha_1, \quad \hat{\alpha}_2 = \frac{r_B}{r}\beta_2, \quad \hat{\beta}_1 = \frac{r_A}{r}\alpha_1, \quad \text{and} \quad \hat{\beta}_2 = \frac{r_A}{r}\beta_2 \tag{**}$$

where $r = r_A + r_B$. Then under the new share rates, the agents select the same actions as those under $(\alpha_1, \beta_2)$, while the risk premiums are minimized. In summary, side trades between the agents who observe each other's actions are beneficial to the principal because she can implement the same efforts with less risk imposed on them. Note that this results from the perfect monitoring capabilities of the agents, not from the opportunities of mutual insurance.

The argument just presented clearly shows that each share parameter must be strictly positive under delegated cooperation. We thus face the same incompatibility of delegated cooperation with relative performance evaluation as that of induced cooperation discussed in the previous section: Under side trading behavior, relative performance evaluation is infeasible. The relative merits of delegated cooperation depends on the correlation coefficient in the way similar to Proposition 4.

*Induced cooperation and delegated cooperation*: We now return to the initial model in which the agents can perform two tasks and task-specific performance measures are available. Can the principal attain delegated cooperation when induced cooperation is optimal under no side trade? To examine this question, suppose that the correlation coefficient is sufficiently small that when no side trade occurs, under the optimal contracts, the principal induces both agents to perform each of the two tasks. That is, I assume $\rho < \rho^*$ if the principal cannot prevent sabotage under relative performance evaluation, or $\rho < \rho^{**}$ if she can do so. In addition, suppose for simplicity that for $i = 1, 2$, $\mu_i(a_i + b_i) = a_i + b_i$, and $c''_{Ai}$ and $c''_{Bi}$ are constant, as in Proposition 2.

PROPOSITION 7. *Under the assumptions given above, the principal attains delegated cooperation if two agents are sufficiently homogeneous, in the sense that for each $i$, $|r_A c''_{Ai} - r_B c''_{Bi}|$ is sufficiently small.*

The proof of the proposition goes as follows. First, suppose $r_A c''_{Ai} = r_B c''_{Bi}$ for $i = 1, 2$.

Then under no side trade, the optimal second-best efforts for the principal satisfy

$$c'_{Ai}(a_i) = c'_{Bi}(b_i) \qquad \text{for} \qquad i = 1, 2. \tag{\#}$$

This implies that inputs into task $i$, $(a_i, b_i)$, minimize the total costs of actions given the total effort $a_i + b_i$ fixed. The principal can implement these efforts with less costs when the agents side trade, and hence delegated cooperation follows in this case.

However, it can be shown that the set of feasible efforts the principal can implement under side contracting by the agents contains only those satisfying (#): the principal has less control over the agents' allocation of efforts between them at each task than under no side trade. This problem costs the principal nothing when the agents are homogeneous because it is in fact the efforts satisfying (#) that the principal would like to implement in the case of no side trade. However, it turns out that the case of homogeneous agents is special: the optimal second-best efforts under no side trade never satisfy (#) unless $r_A c''_{Ai} = r_B c''_{Bi}$ holds for $i = 1, 2$. Of course, this does not imply that delegated cooperation is an exceptional case. What is shown is that the principal usually implements different efforts under these two regimes. And she still attains delegated cooperation if the agents are sufficiently homogeneous.[19]

In summary, if the systematic risk is not important so that the principal prefers induced cooperation to relative performance evaluation, she would also like to encourage the agents to monitor each other's efforts and to coordinate them, attaining delegated cooperation.

*Team production and hidden gaming*: The result similar to Proposition 7 also holds under team production, in which the payoffs from individual tasks $(x_1, x_2)$ are not observable to the principal while the total payoff $x = x_1 + x_2$ is publicly observable. This implies that the share parameters must satisfy $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$, and hence relative performance evaluation is infeasible. Delegated cooperation is therefore attained when the agents are sufficiently homogeneous, for all levels of the correlation coefficient.

When team production is inevitable because of the nature of tasks or production systems, the principal generally suffers from the agents' incentives to shirk because of weak tie between performance measures and individual efforts (Alchian and Demsetz, 1972; Holmström, 1982).

[19] This result also depends on the assumption that each agent is perfectly multi-skilled in the sense that his productivity at task 1 is equal to that at task 2. Otherwise, the optimal second-best efforts for homogeneous agents who do not side trade may not satisfy (#).

As is cited in Section 1, Levine and Tyson (1990), identifying team production as one of the most important features in employee participation arrangements, suggest that mutual monitoring and sanctioning among agents resolve the motivational problem. My analysis verifies their assertion if workers are sufficiently homogeneous.

Now suppose that the principal can hire a supervisor who can observe individual performances $x_1$ and $x_2$ and report to the principal. If the supervisor were benevolent, in the sense that he does not need to be motivated to provide truthfully his information, his information is valuable, and the optimal schemes (under no side trade) depend on individual performance. Laffont (1990) shows that if with sufficiently high probability, the supervisor can abuse his position so as to benefit himself at the expense of the agents,[20] the optimal schemes do not use individual performance.

This "hidden gaming" problem is mitigated if the agents side trade. Suppose that the production is technologically independent, the agents are homogeneous ($r_A c''_{Ai} = r_B c''_{Bi}$ for $i = 1, 2$), and the error terms are stochastically independent.[21] Then one can show the following:

PROPOSITION 8. *Under the assumptions given above, individual performance measures are of no value to the principal when the agents side trade.*

To see the reason, suppose that the supervisor is benevolent so that the principal can design pay schemes contingent on individual performance. However, when the agents side trade, the optimal schemes satisfy $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. That is, they are only contingent on the team performance $x_1 + x_2$. The result thus follows.

Proposition 8 implies that if the agents monitor and coordinate each other's effort, the principal does not need the supervisor, whether he is benevolent or not, and hence the hidden gaming problem does not arise, at least under the restricted assumptions made here. Delegated cooperation is sometimes quite powerful and dominates the collusion problems based on manipulation of information by supervisors.

[20] The supervisor can form a coalition with one of the agents at the expense of the other, or he can extract private benefits from each agent by threatening to favor the other.

[21] These are the assumptions made in Laffont (1990).

## 4. Concluding Remarks

Using the linear principal-multiagent model with moral hazard developed by Holmstrom and Milgrom (1987, 1990b), I have argued that "cooperation" of a subset of organizational members can be beneficial to the organization as a whole. In two kinds of cooperation defined in the paper, induced cooperation and delegated cooperation, the main conclusion is very similar: they are preferred if and only if the correlation of error terms is sufficiently low. I have also obtained other testable predictions: (i) Induced cooperation is more likely to be preferred to relative performance evaluation as agents are less risk averse or the difference in the difficulty of providing incentives (in terms of performance measures and costs of actions) is smaller among tasks; (ii) When agents are jointly responsible for tasks (induced cooperation), whether or not team production prevails, delegated cooperation is more likely to be attained as the agents are more homogeneous in terms of their risk attitudes and cost structure.

The results of the paper offer an incentive-based explanation of the recent trends in organization structures, from hierarchy based on extensive division of labor to team-based organization.[22] Piore (1989) finds that one of the recent organizational reforms in American manufacturing is associated with the movement to reduce, and ultimately eliminate in-process inventory. The movement was partly motivated by the success of the Japanese just-in-time system, which began as an effort to adapt quickly and flexibly to changes in market demands for diverse products. The detailed discussion on the comparison of the Japanese production system with the American mass manufacturing system, from the technological viewpoint, is found in Aoki (1988, chapter 2).

As Piore (1989) points out, this reduction of in-process inventory has an important effect on the relationship among work stations: it forces fundamental changes on the way workers relate to each other, from isolated, independent operations to intense interaction between adjacent operations. That is, it drastically increases opportunities for productive interactions among workers. This change increases the threshold level of the correlation coefficient in my model, from $\rho^{**}$ to $\rho^{*}$, because of the possibility of the negative production externalities under relative performance evaluation: induced cooperation is now more likely to be preferred. Delegated cooperation is then also more likely to be optimal. This is then consistent with ca-

[22] Itoh (1991b) also applies the results to the analysis of human resource management practices of the stylized Japanese firm.

sual observation that the elimination of in-process inventory is accompanied by the changes in work organizational practices such as greater lateral communication and intensive job rotations within work teams. These practices create joint responsibility of workers to multiple tasks, as in induced cooperation, and facilitate mutual monitoring and sanctioning, as in delegated cooperation.

The discussion given above is still speculative, and more systematic tests of the results of the paper are high in my agenda.

## References


ALCHIAN, A. A. AND H. DEMSETZ (1972): "Production, Information Costs, and Economic Organization," *American Economic Review*, vol. 62, pp. 777–795.

AOKI, M. (1988): *Information, Incentives, and Bargaining in the Japanese Economy*. Cambridge: Cambridge University Press.

BARON, J. N. (1988): "The Employment Relation as a Social Relation," *Journal of the Japanese and International Economies*, vol. 2, pp. 492–525.

BLINDER, A. S. (1990) *Paying for Productivity: A Look at the Evidence*. Washington, D.C.: The Brookings Institution.

DERTOUZOS, M. L., R. K. LESTER, R. M. SOLOW AND THE MIT COMMISSION ON INDUSTRIAL PRODUCTIVITY (1989): *Made in America: Regaining the Productive Edge*. Cambridge: The MIT Press.

GRANOVETTER, M. (1985):"Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, vol. 91, pp. 481–510.

GROSSMAN, S. J. AND O. D. HART (1983): "An Analysis of the Principal-Agent Problem," *Econometrica,* vol. 51, pp. 7–45.

———— (1986): "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, vol. 94, pp. 691–719.

HOLMSTRÖM, B. (1982): "Moral Hazard in Teams," *Bell Journal of Economics,* vol. 13, pp. 324–340.

HOLMSTROM, B. AND P. MILGROM (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, vol. 55, pp. 303–328.

———— (1990a): "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics*, vol. 146, pp. 85–105.

———— (1990b): "Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design," mimeo. Yale School of Organization and Management.

HOERR, J. (1989): "The Payoff from Teamwork," *Business Week*, July 10.

ITOH, H. (1990): "Coalitions, Incentives, and Risk Sharing," mimeo. Faculty of Economics, Kyoto University, December.

———— (1991a): "Incentives to Help in Multi-Agent Situations," *Econometrica*, vol. 59.

———— (1991b): "Japanese Human Resource Management from the Viewpoint of Incentive Theory," mimeo. Faculty of Economics, Kyoto University.

LAZEAR, E. P. (1989): "Pay Equality and Industrial Politics," *Journal of Political Economy*, vol. 97, pp. 561–580.

LEVINE, D. I. AND L. D. TYSON (1990): "Participation, Productivity, and the Firm's Environment," in A. S. Blinder (ed.), *Paying for Productivity: A Look at the Evidence*, Washington, D.C.: The Brookings Institution, pp. 183–237.

MA, C. (1988): "Unique Implementation of Incentive Contracts with Many Agents," *Review of Economic Studies*, vol. 55, pp. 555–571.

MILGROM, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, vol. 12, pp. 380–391.

MOOKHERJEE, D. (1984): "Optimal Incentive Schemes with Many Agents," *Review of Economic Studies*, vol. 51, pp. 433–446.

———— (1988): "Competition and Motivation: An Organizational Perspective," mimeo.

NALBANTIAN, H. R. (1988): "Incentive Compensation in Perspective," in H. R. Nalbantian (ed.), *Incentives, Cooperation, and Risk Sharing*. Savage: Rowman and Littlefield, pp. 3–43.

PERROW, C. (1986): *Complex Organizations: A Critical Essay* (3rd ed.). New York: Random House.

PIORE, M. J. (1989): "Corporate Reform in American Manufacturing and the Challenge to Economic Theory," MIT Working Papers no. 533.

RAMAKRISHNAN, R. T. S. AND A. V. THAKOR (1991): "Cooperation versus Competition in Agency," forthcoming in *Journal of Law, Economics, and Organization*.

ROOS, D., J. P. WOMACK, AND D. JONES (1990): *The Machine That Changed the World*. Rawson Associates.

TIROLE, J. (1986): "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economics, and Organization*, vol. 2, pp. 181–214.

———— (1988): "The Multicontract Organization," *Canadian Journal of Economics*, vol. 21, pp. 459–466.

VARIAN, H. R. (1990): "Monitoring Agents with Other Agents," *Journal of Institutional and Theoretical Economics*, vol. 146, pp. 153–174.

## Appendix

**Proof of Proposition 1.**

I first formally state the general optimization problem the principal solves as follows:

$$\max_{a,b,\alpha,\beta} \sum_{i=1}^{2} \mu_i(a_i, b_i) - C^A(a) - C^B(b) - \tfrac{1}{2} r_A \alpha\Sigma\alpha - \tfrac{1}{2} r_B \beta\Sigma\beta \tag{1a}$$

subject to

$$\sum_{i=1}^{2} \alpha_i \mu_i(a_i, b_i) - C^A(a) \geq \sum_{i=1}^{2} \alpha_i \mu_i(a_i', b_i) - C^A(a') \quad \text{for all } a', \tag{1b}$$

and

$$\sum_{i=1}^{2} \beta_i \mu_i(a_i, b_i) - C^B(b) \geq \sum_{i=1}^{2} \beta_i \mu_i(a_i, b_i') - C^B(b') \quad \text{for all } b', \tag{1c}$$

where the risk premium terms are given by

$$\alpha\Sigma\alpha = \alpha_1^2\sigma_1^2 + \alpha_2^2\sigma_2^2 + 2\alpha_1\alpha_2\rho\sigma_1\sigma_2 \tag{2a}$$

and

$$\beta\Sigma\beta = \beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2 + 2\beta_1\beta_2\rho\sigma_1\sigma_2. \tag{2b}$$

(1b) and (1c) are the incentive compatibility constraints stating that $(a, b)$ is a Nash equilibrium.

Because of the assumption of no correlation ($\rho = 0$), it is sufficient to consider share rates that are nonnegative. Then the first-order conditions for (1b) and (1c) are given as follows:

$$a_i\{\alpha_i\mu_i'(a_i + b_i) - c_{Ai}'(a_i)\} = 0 \quad \text{for } i = 1, 2; \tag{3a}$$

$$b_i\{\beta_i\mu_i'(a_i + b_i) - c_{Bi}'(b_i)\} = 0 \quad \text{for } i = 1, 2. \tag{3b}$$

These constraints incorporate the constraints $a_i \geq 0$ and $b_i \geq 0$ into the local incentive compatibility constraints. (See Itoh (1991) for the detail.) The principal solves (1a) with $\mu_i(a_i, b_i) = \mu_i(a_i + b_i)$ and $\rho = 0$ subject to (3a) and (3b). Let $\gamma_i \geq 0$ and $\xi_i \geq 0$ be the Lagrange multipliers for (3a) and (3b), respectively.

Suppose that one agent, say agent B, is solely responsible to one of the tasks, say task 2 ($a_2 = 0$ and $b_2 > 0$) at optimum. Then by the Kuhn-Tucker conditions, there exist $\gamma_i$ and $\xi_i$ ($i = 1, 2$) such that

$$\mu_2'(b_2) - c_{A2}'(a_2) - \gamma_2 c_{A2}'(a_2) + \xi_2 b_2 \beta_2 \mu_2''(b_2) \leq 0 \tag{4a}$$

and

$$\mu_2'(b_2) - c_{B2}'(b_2) + \xi_2 b_2[\beta_2 \mu_2''(b_2) - c_{B2}''(b_2)] = 0. \tag{4b}$$

Solving (4b) for $\mu_2'$ and substituting into (4a) yield

$$c_{B2}'(b_2) + \xi_2 b_2 c_{B2}''(b_2) \leq (1 + \gamma_2) c_{A2}'(a_2). \tag{5}$$

This inequality never holds for any nonnegative $\gamma_2$ and $\xi_2$ under the assumptions of the proposition. (Note in particular $c_{A2}'(a_2) = 0$.) Therefore, sole responsibility cannot be true for either task at optimum.

One only needs to show that both agents select positive efforts at *both* tasks: $a_1 = b_1 = 0$, for example, cannot hold at optimum. This can be easily shown by the procedure similar to that presented above.

**Proof of Proposition 2.**

By the simplifying assumptions of the proposition, one can decompose the optimization problem to one for agent A and one for agent B. Here I focus on the optimization problem for agent A, which is given as follows:

$$\max_{a,\alpha} a_1 + a_2 - c_{A1}(a_1) - c_{A2}(a_2) - \tfrac{1}{2} r_A \alpha \Sigma \alpha \tag{6a}$$

subject to

$$\alpha_1 - c_{A1}'(a_1) = 0 \tag{6b}$$

and

$$\alpha_2 - c_{A2}'(a_2) = 0. \tag{6c}$$

By solving the Kuhn-Tucker conditions for the program, one can easily obtain the following solution:

$$\alpha_1 = D^{-1}(1 + r_A \sigma_2 (m_{A2} - \rho m_{A1})); \tag{7a}$$

$$\alpha_2 = D^{-1}(1 + r_A \sigma_1 (m_{A1} - \rho m_{A2})) \tag{7b}$$

where $D = 1 + r_A(\sigma_1^2 c_{A1}'' + \sigma_2^2 c_{A2}'') + r_A^2(1 - \rho^2)\sigma_1^2 \sigma_2^2 c_{A1}'' c_{A2}'' > 0$. Under the assumption $m_{A2} > m_{A1}$, $\alpha_1$ is always strictly positive, regardless of the correlation coefficient. Also $\alpha_2$ is always strictly positive if

$$1 + r_A \sigma_1 (m_{A1} - \rho m_{A2}) \geq 1 + r_A \sigma_1 (m_{A1} - m_{A2}) > 0$$

or $m_{A2} - m_{A1} < (r_A\sigma_1)^{-1}$. Otherwise, define $\rho^*$ by

$$\rho^* = \frac{1 + r_A\sigma_1^2 c''_{A1}}{r_A\sigma_1\sigma_2 c''_{A2}} = \frac{1}{r_A\sigma_1\sigma_2 c''_{A2}} + \frac{m_{A1}}{m_{A2}}. \tag{8}$$

Then $\alpha_2 > 0$ if and only if $\rho < \rho^*$.

**Proof of Proposition 3.**

This follows directly from (8).

**Proof of Proposition 4.**

I first state the principal's problem for agent A when $\mu_1 = a_1$ and $\mu_2 = b_2$ as follows:

$$\max_{a_1,\alpha} a_1 - c_{A1}(a_1) - \tfrac{1}{2} r_A \alpha\Sigma\alpha \tag{9a}$$

subject to

$$\alpha_1 - c'_{A1}(a_1) = 0. \tag{9b}$$

The solution to this program yields the optimal shares given in $(*)$ in the main text.

We know that by Proposition 1, induced cooperation is better than relative performance evaluation when $\rho = 0$. On the other hand, when $\rho = 1$, the optimal share rates under relative performance evaluation are given by $\alpha_1 = 1$ and $\alpha_2 = -\rho\sigma_1/\sigma_2$. Substituting these into the risk premium term (2a) yields $\alpha\Sigma\alpha = 0$. That is, the first-best is attained, and hence relative performance evaluation is better. What remains to be shown is that the net payoff is decreasing in $\rho$ under induced cooperation and increasing in $\rho$ under relative performance evaluation.

Let $I_A(\rho) = (a(\rho), \alpha(\rho))$ be the optimal incentive scheme for agent A under the correlation coefficient $\rho$ and $P(I_A(\rho); \rho)$ be the total payoff from agent A's actions. The higher $P$ is, the higher the net payoff to the principal is.

First consider induced cooperation. $I_A(\rho)$ then is the solution to program (6a–6c) with $\alpha_1 > 0$ and $\alpha_2 > 0$, and $P(I_A(\rho); \rho)$ is the total payoff (6a). We then want to show $P(I_A(\rho''); \rho'') < P(I_A(\rho'); \rho')$ for $\rho'' > \rho'$. To show this, note that $\rho$ affects only the risk premium term in (6a). $I(\rho'')$ hence satisfies (6b–6c) regardless of correlation, and is feasible under the program with $\rho = \rho'$. However, if $P(I_A(\rho''); \rho'') \geq P(I_A(\rho'); \rho')$, then $P(I_A(\rho''); \rho') > P(I_A(\rho''); \rho'')$ holds. By transitivity, the net payoff with correlation $\rho'$ is higher under $I_A(\rho'')$ than under $I_A(\rho')$. Contradiction. The net payoff to the principal is therefore decreasing in the correlation under induced cooperation.

Under relative performance evaluation without sabotage, $I_A(\rho)$ is the optimal scheme to program (9a–9b) which satisfies $\alpha_1\alpha_2 < 0$ for $\rho > 0$. Then if $P(I_A(\rho''); \rho'') \leq P(I_A(\rho'); \rho')$

hold for $\rho''.\rho'$, then $P(I_A(\rho');\rho') < P(I_A(\rho');\rho'')$ holds. Since $I_A(\rho')$ is feasible under program with $\rho = \rho''$, this contradicts the optimality of $I_A(\rho'')$. We thus have the net payoff increasing in the correlation under relative performance evaluation.

Finally, $\rho^{**} < \rho^*$ is immediate since the net payoff under relative performance without sabotage is strictly higher that with sabotage.

**Proof of Proposition 5.**

It is obvious from the discussion after the proposition.

**Proof of Proposition 6.**

In general, when agents side trade based on their actions, the principal solves the following program: (1a) subject to

$$(a,b) \in \arg\max_{a',b'}\{\sum_{i=1}^{2}(\alpha_i+\beta_i)\mu_i(a_i',b_i') - C^A(a') - C^B(b')\} \tag{10a}$$

and

$$(\alpha,\beta) \in \arg\min_{\alpha',\beta'}\{r_A\alpha'\Sigma\alpha' + r_B\beta'\Sigma\beta'; \alpha'+\beta' = \alpha+\beta\}. \tag{10b}$$

(10a) states that the agents select a Pareto optimal effort pair. (10b) states that the contract prevents the agents from reinsuring between them.

Suppose $\mu_1(a_1,b_1) = \mu_1(a_1)$ and $\mu_2(a_2,b_2) = \mu_2(b_2)$. I first show that the principal enjoys delegated cooperation when $\rho = 0$. Suppose $(a_1,b_2,\alpha_1,\beta_2)$ is the optimal contract under no side trade. (Note $a_2 = b_1 = \alpha_2 = \beta_1 = 0$ by independence.) They then satisfy

$$\alpha_1 = \frac{c'_{A1}(a_1)}{\mu_1'(a_1)} \qquad \text{and} \qquad \beta_2 = \frac{c'_{B2}(b_2)}{\mu_2'(b_2)}. \tag{11}$$

Consider the following program:

$$\min_{\alpha',\beta'}\sum_{i=1}^{2}\{(r_A(\alpha_i')^2 + r_B(\beta_i')^2)\sigma_i^2\} \tag{12a}$$

subject to

$$\alpha_1' + \beta_1' = \alpha_1 \qquad \text{and} \qquad \alpha_2' + \beta_2' = \beta_2. \tag{12b}$$

Let $(\hat{\alpha},\hat{\beta})$ be the solution to this program, which is given by $(**)$ in the main text after the proposition. By (11) and (12b), the solution satisfies (10a), and hence the side trading agents select $(a_1,b_2)$ under $(\hat{\alpha},\hat{\beta})$. In addition, it is obvious that $(\hat{\alpha},\hat{\beta})$ satisfies (10b) when $\rho = 0$. Therefore $(\hat{\alpha},\hat{\beta})$ is the optimal scheme implementing $(a_1,b_2)$ under side trading, and is preferred to $(\alpha_1,\beta_2)$ by the principal because it attains lower risk premiums than $(\alpha_1,\beta_2)$.

When $\rho = 1$, relative performance evaluation under no side trade is preferred because it attains the first-best outcome as shown in the proof of Proposition 4. The latter proof also shows that the net payoff to the principal is increasing in $\rho$ under relative performance evaluation. The argument similar to that in the proof of Proposition 4 also applies to show that the net payoff is decreasing in $\rho$ under side trading. The existence of the threshold value thus follows.

**Proof of Proposition 7.**

*Claim (a). When the agents side trade based on their actions, the set of feasible efforts the principal can implement contains only those efforts which satisfy (#):*

Under the assumptions, the side trading agents select their efforts by

$$\max_{a,b} \sum_{i=1}^{2} \{(\alpha_i + \beta_i)(a_i + b_i) - c_{Ai}(a_i) - c_{Bi}(b_i)\}. \tag{13}$$

The first-order conditions thus yield $\alpha_i + \beta_i = c'_{Ai}(a_i) = c'_{B_i}(b_i)$ for $i = 1, 2$.

*Claim (b). If the principal wishes to implement efforts satisfying (#), then she attains delegated cooperation:*

Let $(a, b, \alpha, \beta)$ be the optimal contract under no side trade and suppose for $i = 1, 2$, $c'_{Ai}(a_i) = c'_{Bi}(b_i)$. Then by the incentive compatibility constraints, $\alpha_i = \beta_i$ for $i = 1, 2$. Let $(\hat{\alpha}, \hat{\beta})$ be the solution to the following program:

$$\min_{\alpha', \beta'} r_A \alpha' \Sigma \alpha' + r_B \beta' \Sigma \beta' \tag{14a}$$

subject to

$$\alpha' + \beta' = \alpha = \beta. \tag{14b}$$

Then clearly side trading agents select $(a, b)$ under $(\hat{\alpha}, \hat{\beta})$. Furthermore, they have no incentive to reinsure themselves, and hence $(\hat{\alpha}, \hat{\beta})$ satisfies (10b). Finally, the risk premiums are lower under $(\hat{\alpha}, \hat{\beta})$ than under $(\alpha, \beta)$ because

$$\begin{aligned} r_A \hat{\alpha} \Sigma \hat{\alpha} + r_B \hat{\beta} \Sigma \hat{\beta} &< \min_{\alpha', \beta'} r_A \alpha' \Sigma \alpha' + r_B \beta' \Sigma \beta' \quad \text{subject to } \alpha' + \beta' = \alpha + \beta \\ &\leq r_A \alpha \Sigma \alpha + r_B \beta \Sigma \beta \end{aligned}$$

where the first inequality follows because by assumption, induced cooperation is optimal in the case of no side trade, and hence $\alpha + \beta > \alpha = \beta$.

*Claim (c). When no side trade occurs, the optimal efforts for the principal satisfy (#) if and only if $r_A c''_{Ai} = r_B c''_{Bi}$ for $i = 1, 2$:*

Given the total effort for task $i$, denoted by $t_i$, the principal solves the following program when agents do not side trade:

$$\min_{a,\alpha,\beta} \sum_{i=1}^{2} \{c_{Ai}(a_i) + c_{Bi}(t_i - a_i)\} + \tfrac{1}{2} r_A \alpha \Sigma \alpha + \tfrac{1}{2} r_B \beta \Sigma \beta \tag{15a}$$

subject to

$$\alpha_i - c'_{Ai}(a_i) = 0 \quad \text{for } i = 1,2 \tag{15b}$$

and

$$\beta_i - c'_{Bi}(t_i - a_i) = 0 \quad \text{for } i = 1,2. \tag{15c}$$

Let $\gamma_i$ and $\xi_i$ be the Lagrange multipliers for (15b) and (15c), respectively. Then the Kuhn-Tucker necessary and sufficient conditions yield the following:

$$c'_{Ai} - c'_{Bi} + \gamma_i c''_{Ai} - \xi_i c''_{Bi} = 0 \quad \text{for } i = 1,2; \tag{16a}$$

$$r_A \alpha_i \sigma_i^2 + r_A \alpha_j \rho \sigma_1 \sigma_2 = \gamma_i \quad \text{for } i,j = 1,2 \text{ and } j \neq i; \tag{16b}$$

$$r_B \beta_i \sigma_i^2 + r_B \beta_j \rho \sigma_1 \sigma_2 = \xi_i \quad \text{for } i,j = 1,2 \text{ and } j \neq i. \tag{16c}$$

By (15b–15c) and (16a–16c), we obtain

$$c'_{Ai} - c'_{Bi} + r_A c''_{Ai}(\sigma_i^2 c'_{Ai} + \rho \sigma_1 \sigma_2 c'_{Aj}) - r_B c''_{Bi}(\sigma_i^2 c'_{Bi} + \rho \sigma_1 \sigma_2 c'_{Bj}) = 0 \quad \text{for } i,j = 1,2 \text{ and } j \neq i. \tag{17}$$

The conclusion follows from (17).

From Claims (a)–(c), the conclusion of the proposition follows.

**Proof of Proposition 8**

By the symmetry of the agents and the technological independence, the optimal efforts satisfy $a_1 = b_2$. Then $\alpha_1 + \beta_1 = \alpha_2 + \beta_2$ holds at optimum. Since the agents have the same coefficient of absolute risk aversion, they share risk equally, i.e., $\alpha_i = \beta_i$ for $i = 1,2$, and hence $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. This implies that although individual performance measures are available, the optimal pay schemes depend only on the team performance, when the agents side trade.