Discussion Paper # 471

MORAL HAZARD IN TEAMS

by

Bengt Holmstrom*

February 1981

J.L. Kellogg School of Management
Northwestern University
Evanston, Illinois 60201

* This paper grew out of my discussion of Stan Baiman and Joel Demski's paper: "Economically Optimal Performance Evaluation and Control Systems" presented at the Chicago Conference of Accounting 1980. Their paper provided ample sources of inspiration for which I am grateful. I have also benefited from discussions with Steve Ross and from comments received in seminars at Yale and the University of Pennsylvania.

I. Introduction

Orthodox economic theory has little to offer in terms of understanding how non-market organizations, like firms, form and function. This is so because traditional theory pays little or no attention to the role of information, which evidently lies at the heart of organizations. The recent development of information economics, which explictly recognizes that agents have limited and different information, is a welcomed invention, which promises to be helpful in understanding the intricacies of organizational structure. Particularly important in this context are questions concerning the control of agent's incentives, which to a large degree dictate the design of organizations and set the limits of its performance potential.

The members of an organization may be seen as providing two kinds of services: they supply inputs for production and process information for decision-making. Along with this dichotomy goes a taxonomy for incentive problems. Moral hazard refers to the problem of inducing agents to supply proper amounts of productive inputs when their actions cannot be observed and contracted upon directly. Adverse selection refers to a situation where actions can be observed, but it cannot be verified whether the action was the correct one given the agent's contingency, which is not jointly observed.

This paper is concerned with remedies to moral hazard in teams. By a team I mean rather loosely a group of individuals who are organized together for productive activities. There is no presumption that they share a common objective as in the more technical use of the term (see Marschak and Radner's team theory).

I will first consider team production under certainty. The model is as follows: There are many agents, whose privately taken actions jointly determine output. An agent's action is costly to himself and neither the cost nor the action can be observed by others. The problem is to induce agents to take efficient actions for optimal team production. Since it is not possible in general to infer the agents' actions from the outcome, agents can cover dysfunctional behavior by blaming each other. This is the well-known free-rider phenomenon. If agents form a partnership, and hence have to share the outcome fully between themselves, there exists no sharing rule based on the joint outcome alone, which induces proper incentives for action.

This observation is the starting point for Alchian and Demsetz's (1972) well-known article. Since a partnership is productively inefficient, they argue that competition will lead the partnership to fall apart and develop into an organization where there is a monitor who will control that agents take correct actions. In order to induce the monitor to perform his job properly, he should be given a residual of the outcome. This will guarantee efficiency according to Alchian and Demsetz.

This line of reasoning provides a theory of the firm. Firms develop since their organizational structure is superior to a partnership as argued above. Alchian and Demsetz's analysis does not explain, however, the existence of corporations, where part of the residual goes to stock owners who do a very limited amount of monitoring themselves. I will argue that monitoring can largely be dispensed with in the context of certainty simply by letting outsiders, who provide no inputs for production pick up the residual. Therefore, separating ownership from production will resolve moral hazard and restore efficiency. This provides for an alternative rudimentary rationale for the capitalistic firm, that is, a reason for why capital hires labor, rather than the reverse.

Separating ownership from production is one way of alleviating incentive problems. The other is the aforementioned monitoring. Of course, for monitoring to resolve the problem, a sufficiently rich set of observable measures of productionis required. I address the question how rich. Quite generally two measures will do. However, if we insist on shares being monotone in output, then as many measures are needed as there are agents and these measures must effectively discern the actions taken by the agents. This is the familiar theme of responsibility accounting, which states that agents should be responsible only for what they can control.

The second part of the paper considers the impact of uncertainty. First, I explore circumstances under which separation of ownership and production will resolve inefficiencies as in the case of certainty. Mirrlees (1974) observed that penalties can be quite effective in alleviating moral hazard problems. His model had a single agent. I extend the result to the context of team production by using group penalties. The case of certainty is but a special case of this more general setting.

An objection to penalties is that they may have to be infeasibly large. If there are constraints on agents' wealth, which restrict the extent to which agents may be penalized, this will generally hinder an efficient solution (Mirrlees' agrument normally requires unbounded penalties). The more agents there are, the less efficient it will be to police opportunism by group penalties. Consequently, bounded penalties lead to limits on the size of team organization when uncertainty is present. This contrasts with the certainty case for which one can find an efficient, individually feasible solution irrespective of team size and wealth constraints.

Given that group penalties will be ineffective in larger teams, I go on to study the second alternative for resolving incentive problems: monitoring. Two results are proved. The first one states a necessary and sufficient condition for a new output measure to improve on the second-best solution. This is an extension of a result of mine for the single agent case. It says that only the sufficent statistic of a set of measures is relevant for controlling agents' actions. This may sound obvious but it not. One should appreciate that we are not concerned with a statistical decision problem, but with a strategic game. It is appealing, but by no means evident, that the same sufficient statistic condition works in a strategic context.

The second result builds on the first. Departing somewhat from the notion of team production, I show that in a situation with independently producing agents, moral hazard can be better controlled by evaluating each agent against average performance, provided that the uncertainties are correlated. If agents share a common uncertain factor of production, then as the number of agents grows large, the use of an average will remove the commonly shared uncertainty from the problem and leave us only with solving that part of the problem that pertains to the agents' idiosyncratic risks.

It is interesting to note that a recent trend in executive incentive scheme design is to use performance of closely related firms as a base for executive compensation. The earlier practice of stock option compensation has the drawback of leaving the executive faced with risk he can in no way control. In accordance with theory (and common sense) such risk should be filtered away to the extent possible. This is the virtue of the new performance incentive packages.

Measuring performance against peer averages has been common in firms for a long time. Lazear and Rosen (1980) have studied a specific form of such practice, which they call rank order tournaments. In rank order tournaments bonuses are awarded based on performance rank. Agents are compared against each other, rather than absolute performance standards. The rationale, as will be shown, rests solely on the value that peer performance has in providing information about production uncertainties. If agents do not face any common sources of uncertainty, then it is inefficient to use rank order tournaments.

To summarize, the paper pursues two ways of combatting moral hzard in teams. The first one uses separation of ownership and labor to allow for non-budget-balancing penalities or bonuses. This way one may frequently induce proper incentives at low cost. The other one relies on increased monitoring of performance to uncover actual actions. The question addressed and answered is what type of monitoring will be valuable and improve efficiency.

The rest of the paper is organized as follows. Section 2 deals with the case of certainty when joint output is the only observable measure. Section 3 addresses the issue of how detailed a monitoring system should be in order to reach efficiency under certainty. Section 4 considers use of penalties and bonuses when there is uncertainty. Section 5 discusses the value of monitoring under uncertainty and Section 6 the use of relative performance measures. The last section offers a summary of the main implications of the model for understanding and improving organizational design.

2. Certainty - Single output measure

Consider the following simple model of team production. There are n agents. Each agent, indexed i, takes a non-observable action $a_i \varepsilon A_i = [0,\infty)$, with a private (possibly non-monetary) return $v_i: A_i \to \mathbb{R}$; v_i is strictly convex and decreasing with $v_i(0) = 0$. Let $a = (a_1,\ldots a_n) \ \varepsilon \ A \equiv \displaystyle\mathop{X}_{i=1}^{n} A_i$ and write $a_{-i} = (a_1,\ldots,a_{i-1}, a_{i+1},\ldots,a_n)$, $a = (a_i, a_{-i})$.

The agents' actions determine a joint monetary outcome $x : A \to \mathbb{R}$, which must be allocated among the agents. The function x is assumed strictly concave with $x(0) = 0$. Let $s_i(x)$ stand for agent i's share of the outcome x. The preference function of agent i is assumed (for simplicity) to be linearly separable and hence of the form $u_i(n_i, a_i) = n_i + v_i(a_i)$ over money and action.

The question is whether there is a way of fully allocating the joint outcome x so that the resulting non-cooperative game between the agents has a Pareto optimal Nash equilibrium. That is, do there exist sharing rules $\{s_i(x)\}$ such that we have budget-balancing

$$(1) \qquad \sum_{i=1}^{n} s_i(x) = x, \text{ for all } x,$$

and the non-cooperative game with payoffs

$$(2) \qquad s_i(x(a)) + v_i(a_i), \; i=1,\ldots,n,$$

has a Nash equilibrium a*, which satisfies the condition for Pareto optimality,

$$(3) \qquad a^* = \underset{a \in A}{\operatorname{argmax}}\; x(a) + \sum_{i=1}^{n} v_i(a_i).$$

If the sharing rules are differentiable we find, since a* is a Nash equilibrium, that

$$(4) \qquad s_i' x_i + v_i' = 0, \; i=1,\ldots,n,$$

where $x_i \equiv \partial x/\partial a_i$. Pareto optimality again implies

(5) $$x_i + v_i' = 0, \; i=1,\ldots,n.$$

Consistency of (4) and (5) requires $s_i' = 1$, $i=1,\ldots,n$. But this is in conflict with (1) since differentiating (1) implies

(6) $$\sum_{i=1}^{n} s_i' = 1.$$

Therefore, with differentiable sharing rules we cannot reach efficient Nash equilibria. The same is true more generally as stated in the following:

Theorem 1. Assume there exists a Pareto optimal solution a* in the inferior of A for which $x_i(a^*) \neq 0$ for all i. Then there do not exist sharing rules $\{s_i(x)\}$ which satisfy (1) and for which a* is a Nash equilibrium in the non-cooperative game with payoffs (2).

Proof: Let $s_i(x)$, i=1,..,n, be arbitrary sharing rules satisfying (1). I will show that the assumption that a* is a Nash equilibrium will lead to a contradiction.

From the definition of a Nash equilibrium

(7) $$s_i\big(x(a_i, a^*_{-i})\big) + v_i(a_i) \leq s_i\big(x(a^*)\big) + v_i(a_i^*), \; \forall a_i \varepsilon A_i.$$

Let $\{\alpha^\ell\}$ be a strictly increasing sequence of real numbers converging to $x(a^*)$. Let $\{a_i^\ell\}$ be the corresponding n sequences satisfiying

(8) $$\alpha^{\ell} = x(a_i^{\ell},\ a_{-1}^*).$$

The sequences $\{a_i^{\ell}\}$ are well-defined (starting from a large enough ℓ if necessary) since $a^* \varepsilon$ int A, $x_i(a^*) \neq 0$ and $x(a)$ is strictly concave. Pareto optimality implies

$$v_i'(a_i^*) = -x_i(a^*) \ , \ \forall_i .$$

This in turn implies, using (8), that $v_i(a_i^{\ell}) - v_i(a_i^*) = x(a^*) - x(a_i^{\ell}, a_{-i}^*)$ $+ o(a_i^{\ell} - a_i^*)$, $\forall_i$, $\forall \ell$, where $o(h)/h \to 0$ as $h \to 0$. Substituting into (7), using (8) gives

(9) $$x(a^*) - \alpha^{\ell} + o(a_i^{\ell} - a_i^*) \leqslant s_i\big(x(a^*)\big) - s_i(\alpha^{\ell}),\ \forall_i,\ \forall \ell.$$

Sum (9) over i and use (1) to get

$$\sum_{i=1}^{n} \{ - x(a^*) - \alpha^{\ell} + o(a_i^{\ell} - a_i^*)\} \leqslant 0, \quad \forall \ell.$$

This can be written

(10) $$\sum_{i=1}^{n} \{- x_i(a^*)(a_i^{\ell} - a_i^*) + o(a_i^{\ell} - a_i^*) \leqslant 0,\ \forall \ell.$$

Since $\alpha^{\ell} < x(a^*)$ by the choice of α^{ℓ}, and $x_i(a^*) \neq 0$, the first term in the bracket is strictly positive. For large enough ℓ, this term dominates, which contradicts (10). Hence, the assumption that a* is a Nash equilibrium has led to a contradiction and must be false. Q.E.D.

Theorem 1 extends the intuition of the inconsistency of (4)-(6) to the case of arbitrary sharing rules. As long as we insist on budget-balancing (e.g. (1)) and there are externalities present, ($x_i \neq 0$) we cannot achieve efficiency. Agents can cover improper actions behind the uncertainty concerning who was at fault, since all agents cannot be penalized sufficiently for a deviation in the outcome. Therefore, some agent always has an incentive to capitalize on this control deficiency.

The result is indicative of the free-rider problems closed (budget-balancing) organizations face. Examples include labor-managed firms, firm cooperatives, management teams and professional services firms like CPA partnerships. In all cases labor and ownership are integrated and is likely to result in insufficient supply of productive inputs like effort.

This is the starting point for Alchian and Demsetz's (1972) reasoning. As I described in the introduction Alchian and Demsetz conclude that partnerships will break down and be replaced by firms in which an owner takes on the task to monitor agents. The owner is given the title to net earnings in order to have the proper incentive to monitor. This argument provides a simple theory of the capitalistic firm, which emerges as the most efficient organizational form in combatting adverse incentives.

Note, however, that it is not evident that full efficiency will be reached with an owner-monitor. Insofar that he expends effort in monitoring we will have an augmented team, albeit with more measures than the single joint outcome as a basis for sharing the output. In the next section I will take up the question of how rich a measurement system needs to be to reach efficiency. Here I will discuss a simpler, more fundamental solution to the problem: elimination of the budget-balancing requirement.

If budget-balancing (e.g.(1)) is not a constraint, one can make the efficient outcome a Nash equilibrium by giving each agent the total share of the outcome. From equations (4)-(6) we see that, indeed, that is the only differentiable (at a^*) scheme that works. It has the drawback that there will be insufficient funds to compensate agents if they for some reason choose a joint action a such that $x(a) > x(a^*)$. A more appropriate scheme has a kink at $x(a^*)$. For instance consider the step function:

$$(11) \qquad s_i(x) = \begin{cases} b_i & , x \geq x(a^*) \\ k_i & , x < x(a^*), \end{cases}$$

where $\sum_i b_i = x(a^*)$. This will work as long as $b_i + v_i(a_i^*) \geq k_i - v_i(0) = k_i$. Since Pareto optimality of a^* implies

$$x(a) + \sum_i v_i(a_i^*) \geq x(0) + \sum_i v_i(0) = 0,$$

we can clearly choose b_i:s such that $\sum_i b_i = x(a^*)$ and $b_i + v_i(a_i^*) \geq 0$. Therefore, by taking $k_i = 0$, $i=1,\ldots,n$, we can enforce efficiency without penalties exceeding individual resources. This shows,

Theorem 2. There exist feasible group incentives, which induce an efficient Nash Equilibrium without violating individual endowment constraints.

The fact that a feasible group incentive scheme exists regardless of the size of the team implies that under certainty, incentive problems do not impose constraints on team size. That will not be the case under uncertainty.

Group incentives, where all agents are penalized since the ones at fault cannot be discerned, are found in some types of contracting with labor teams. Usually it takes the form of a flat wage for team members with a group bonus which is paid only if the target is attained (whether we view the discontinuity in (11) as a bonus or penalty appears immaterial). An extreme example of group incentives is the dismissal of the board of directors of a firm.

We find then that no monitoring is needed in order to achieve efficiency, at least under certainty. Though the point is technically trivial, I think it is conceptually important: breaking the budget-balancing constraint is an effective way of resolving externalities in the team. A natural way of doing it is to separate ownership and labor, that is, introduce an owner to the organization who does not provide any productive inputs, but merely picks up the residual of the non-budget-balancing sharing rule. This provides for a rudimentary theory of the capitalistic firm, which differs somewhat from that of Alchian and Demsetz's in that monitoring plays no role in the argument. Note that stock owners do not generally exercise very close monitoring of the behavior of managers, only of managers' performance as measured by the total outcome. Perhaps this can be taken as an indication that not monitoring but ownership separation, is the essence in the argument. This is, of course, not saying that monitoring can generally be abandoned. Its role may be important when there is uncertainty, as will be seen shortly.

The theme that budget-balancing and efficiency are inconsistent when externalities are present is certainly not novel. A celebrated solution to the resulting free-rider problem in the public goods context is Groves scheme (Groves, 1973). I note that Groves solution is possible only by breaking the budget constraint; Groves scheme does not balance the budget (barring exceptional cases), because balancing the budget would necessarily result in inefficiencies in analogy with theorem 1.

3. Certainty-Monitoring

The conclusion from the previous section is that it is insufficient to observe the total outcome if one wants to achieve an efficient noncooperative equilibrium when budget-balancing is imposed. If budget-balancing cannot be relaxed the alternative is to observe additional signals about the agents' actions or, as I will view it, get a more detailed account of the outcome measure x.

Let x consist of the sum of m measures $x_k: A \to \mathbb{R}^1$, $k = 1,\dots,m$ i.e.,

$$\sum_{k=1}^{m} x_k(a) = x(a), \qquad \forall a \varepsilon K. \tag{12}$$

Call the set of functions x_k, an <u>accounting system</u>. Based on this accounting system, we can design an <u>allocation mechanism</u>, which is a set of sharing rules $s_i: \mathbb{R}^m \to \mathbb{R}$, $i=1,\dots, n$, satisfying:

$$\sum_{i=1}^{n} s_i(x_i,\dots,x_m) = x \qquad \forall x. \tag{13}$$

The pair consisting of an allocation mechanism and an accounting system will be called a <u>control system</u>. If a control system leads to a Nash equilibrium at the Pareto optimal action a*, the control system is <u>acceptable</u>. If an accounting system is rich enough so that an acceptable control system can be built upon it, the accounting system is <u>sufficient</u>.

With this terminology our problem can be posed as follows: find the conditions under which an accounting system is sufficient. The result from the previous section (Theorem 1) was that the total outcome alone is an insufficient accounting system.

The reason why a richer set of measures may help to control the agents better, is, of course, that several measures generally make it possible to infer more about individual actions. In the limit, a sufficiently rich accounting system may reveal exactly the actions of the agents. In that case an acceptable control system can easily be constructed. I state this formally in the following:

Theorem 3. If the accounting system is a one-to-one mapping from decisions to outcomes, then it is sufficient.

Proof: The measures will reveal each agent's action and so we can make the sharing rules directly dependent on these actions. Let a^* be a Pareto optimal action. Let the sharing rules at a^* be $s_i(a^*)$, $i=1,\ldots,n$. We will show that for any $a \varepsilon A$, the outcome can be shared so that:

$$(14) \qquad v_i(a_i) + s_i(a) \leq v_i(a_i^*) + s_i(a^*) \quad \forall i.$$

This clearly implies our claim.

Let $a \varepsilon A$ be arbitrary, and suppose (14) cannot be achieved. That implies there exist sharing rules $\hat{s}_i(a)$ such that

$$(15) \qquad v_i(a_i) + \hat{s}_i(a) \geq v_i(a_i^*) + s_i(a^*), \quad \forall i,$$

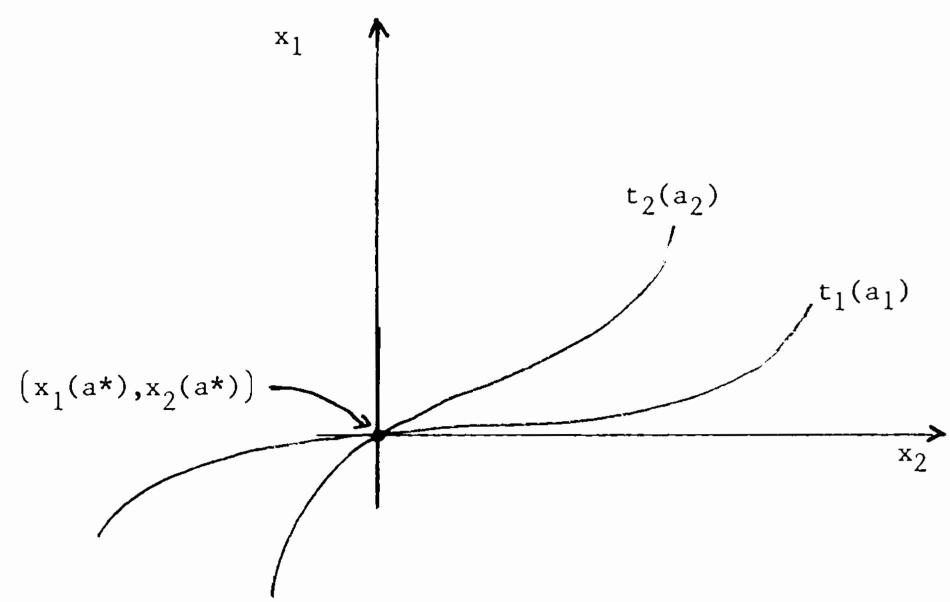
with strict inequality for at least one. Add (15) over all i's to get

$$\sum_{i=1}^{n} v_i(a_i) + x(a) > \sum_{i=1}^{n} v_i(a_i^*) + x(a^*),$$

using (1). This contradicts the Pareto optimality of a*. Q.E.D.

All Theorem 3 says is that if actions are observable or possible to infer with certainty, one can achieve efficiency. The payoffs of the noncooperative game can be redistributed in such a manner that the most desirable outcome is the only Nash equilibrium. Moreover, it can be done so as to satisfy individual feasibility constraints.

The assumption of observability is quite strong in Theorem 3 and can be weakened. It suffices that we can detect when an agent is the only one who deviates from the optimum. This will be possible if and only if the curves $t_i(a_i) = (x_1(a_i, a_{-i}^*), \ldots, x_m(a_i, a_{-i}^*)) \quad \varepsilon \mathbb{R}^m$, $i = 1, \ldots, n$ differ as illustrated in the figure below.



Call an accounting system <u>independent at a*</u> if there does not exist an $a \varepsilon A$, $a \neq a^*$ such that $t_1(a_1) = \ldots = t_n(a_n)$.

Theorem 4: An independent accounting system is sufficient.

Proof: Let $s_i(x_1(a^*),\ldots,x_n(a^*))$, $i=1,\ldots,n$ be an arbitrary split of $x(a^*)$, which satisfys (1). Define sharing rules s_i along the t_i-curves as follows:

$$s_i(t_i(a_i)) = s_i(x_1(a^*),\ldots,x_m(a^*)) + x(a_i,a^*_{-i}) - x(a^*) \tag{16}$$

for $i=1,\ldots,n$, and the others arbitrary but so that (1) holds. This is possible by the assumption of independence. With such a choice the agent's objective coincides with the social objective when others stick to their efficient action a^*_{-i}. Hence the agent's best response against a^*_{-i} will be a^*_i by definition of a^*. Q.E.D.

Notice that one measure does not constitute an independent account system. Independence is also a necessary condition in the sense that a sufficient accounting system has to be independent at least in the neighborhood of a^*. Notice further that if actions of the agents are perfect substitutes of each other, then no accounting system can be independent, since by definition of substitutability, for any a_i, there exists an a_j for each $j \neq i$, such that $t_i(a_i) = t_j(a_j)$.

From Theorem 5 we see that two measures may well be sufficiently rich to reveal individual deviation. However, if we make the assumption that x_k's are monotone in actions, e.g.,

$$\frac{\partial x_k(a)}{\partial a_i} \geqslant 0 \qquad \text{for every } a\varepsilon A,\ i=1,\ldots,n, \tag{17}$$

and constrain ourselves to differentiable and monotone sharing rules,

(18) $$\frac{\partial s_i}{\partial x_k}(x_1,\ldots,x_n) \geq 0 \qquad \forall x, \forall i,$$

then at least n measures are needed.

Theorem 5: Make the assumptions of theorem 1 and in addition that (17) holds. Then the accounting system has to include at least n measures if one wants to construct a monotone acceptable control system.

Proof: By (12) and (13) we have

(19) $$\sum_{i=1}^{n} s_{ik} = 1$$

where $s_{ik} = \frac{\partial s_i}{\partial x_k}\left(x_1(a^*),\ldots,x_n(a^*)\right)$.

By (3) and (12)

(20) $$v_i' + \sum_{k=1}^{m} x_{ki} = 0 \quad \forall i$$

where v_i' is evaluated at a^*, and $x_{ki} = \frac{\partial x_k}{\partial a_i}(a^*)$. From the Nash equilibrium property of a^*

(21) $$v_i' + \sum_{k=1}^{m} s_{ik}\, x_{ki} = 0$$

Combining (20) and (21) yields

(22) $$\sum_{k=1}^{m} x_{ki}(1 - s_{ik}) = 0$$

By (18) and (19) and (17),

$$x_{ki}(1 - s_{ik}) = 0 \quad \forall i, \forall k.$$

Since $x_i(a^*) \neq 0, x_{ki} > 0$ for at least one k for a given i, say k_i. Then $s_{ik_i} = 1$, which implies $s_{jk_i} = 0, \forall_j \neq i$, by (18) and (19). Hence, there must be at least n measures, since each agent is given the full share in at least one. Q.E.D.

The assumption of monotonicity is rather natural to make if we think of x_k's as monetary outcomes which improve with, say, increased effort. In practice most sharing rules are monotone. All agents get a positive share in the outcome. Under such circumstances efficiency can be achieved only if each agent is in charge of his own account. Moreover, the proof shows that it must be that his action does not affect the other agents' accounts. In other words, only when the whole system can be decoupled and externalities removed can we achieve efficiency.

The conclusion is that if budget-balancing is required, the only way to reduce inefficiencies is to create a richer accounting system which better discerns individual deviations. Two measures may be sufficient, but if they are monontone and we want a monotone allocation mechanism, then n independent measures are needed which in effect decouple the organization. The desire to decouple the organization is familiar from responsibility accounting. The analysis supports the widely accepted accounting principle that managers should be able to control the measures that are used for evaluation of their performance (see Horngren 1972, Chapter 6 on responsibility accounting and motivation).

We have not discussed the possibility that some decision, say an allocation of the firm's resources, may make agents' actions dependent. If one tries to promote goal congruence, in order to guarantee an efficient allocation of resources by giving each agent a share in the firm's outcome, this will again lead to insufficient supply of effort. It is interesting to note that Groves' scheme is able to get around this problem. By effectively decoupling the organization, it can assure both optimal allocation of resources and efficient supply of effort. Once the allocation of resources is determined, each agent is in charge of his own account as required for efficiency.

4. Uncertainty-Single Output Measure.

In section 2 we found that efficiency can be attained by using penalties (or bonuses). A team can be made efficient by committing itself to wasting some of the output if the desired target level is not reached. In equilibrium no output is wasted. Under uncertainty output would generally be wasted and this fact may make the certainty case appear extreme. That is not true, however. Penalties may work quite effectively also under uncertainty. This was first observed by Mirrlees (1974) in the context of a principal-agent relationship. The argument is here extended to the team case.

For the moment agents are assumed risk neutral. Output $x(a,\theta)$ is random through the state of nature θ. For all θ, x is assumed concave. Agents have homogenous beliefs concerning θ.

It is more convenient and illuminating to supress θ and consider the distribution function of x parametrized by a. Denote it $F(x,a)$. Output is shared according to sharing rules $s_i(x)$, $i=1,\ldots,n$, for which

$$(23) \qquad \sum_{i=1}^{n} s_i(x) \leq x, \text{ for all } x.$$

Equation (23) permits waste. Generally, we would have to require in addition that $s_i(x) \geq w_i$, where w_i is agent i's endowment, but for the time being I omit consideration of such a constraint. It is assumed that the partial derivatives

$$F_i(x,a) \equiv \partial F(x,a)/\partial a_i, \; i=1,\ldots,n$$

exist for all (x,a).

Theorem 6: Assume

(i) $F(x,a)$ is convex in a.

(ii) $F_i(x,a)/F(x,a) \to -\infty$.

Then a first-best solution can be approximated arbitrarily closely using group penalties.

Assumption (ii) is equivalent to assuming that $f_i(x,a)/f(x,a) \to \infty$. This ratio is known as the monotone liklihood ratio (Milgrom, 1980) and is in effect a measure of how sharply one can distinguish actions based on the output. Assumption (i) may not be very reasonable. Its purpose is to guarantee that first-order conditions correspond to global optima. In many situations it can be relaxed.

Proof: Consider the following sharing rules:

$$(24) \qquad s_i(x) = \begin{cases} s_i x, & x \geq \bar{x}, \\ s_i x - k_i, & x < \bar{x}, \end{cases}$$

where $k_i > 0$, and $\Sigma s_i = 1$. Evidently (24) satisfies (23). The rules in (24) prescribe a penalty k_i to each agent i if a critical output level $\bar{x}$ is not achieved. Otherwise output is shared wholly. In order for a* to be a Nash equilibrium with (24) it is necessary and sufficient (by assumption (i)) that

$$(25) \qquad s_i E_i(a^*) - k_i F_i(\bar{x}, a^*) + v_i'(a_i^*) = 0, \ i=1,\ldots,n,$$

where $E(a) = Ex(a)$ the expected value of x given a and $E_i(a) = \partial Ex(a)/\partial a_i$. For fixed $\bar{x}$, choose k_i so that (25) holds. The expected waste is given by

$$W = \Sigma_i k_i F(\bar{x}, a^*).$$

We need to show that W can be made arbitrarily small. From (25),

$$(26) \qquad k_i = A_i / F_i(\bar{x}, a^*), \text{ with } A_i = s_i E_i(a^*) + v_i'(a_i^*).$$

Let $\bar{x}$ decrease and adjust k_i so that (26) holds. Then waste is given by

$$(27) \qquad W = \sum_i A_i F(\bar{x}, a^*)/F_i(\bar{x}, a^*),$$

and goes by assumption (ii) to zero with $\bar{x}$. Q.E.D.

Theorem 6 says that if lower values of x admit increasingly sharp inferences to be made about a, then first-best can be approximated arbitrarily closely by using group incentives where all agents are penalized for low output. This is the force of assumption (ii). It should be noted that (ii) holds for a number of standard distributions, including the case where $x(a,\theta) = x(a) + \theta$ and θ is normal. Assumption (i) is false for the normal distribution over the whole range x, but it is valid for x-values below the mean. Therefore, since $\bar{x} \to -\infty$, the argument goes through for the normal case as well.

Note that penalities k_i have to be designed individually unless the problem is symmetric and agents are identical.

Through (26) and (ii), k_i:s go to infinity with $\bar{x}$. This may violate endowment constraints. If the x-distribution is tight (low variance) we can get a good approximation already with small k_i:s, but in general not. Thus, efficiency losses may be substantial with constraints on wealth. The fact that the degree to which we can approximate first-best depends on the amount of penalties that can be imposed on agents, contrasts with the certainty case in which first-best can be achieved without exceeding endowments.

Under uncertainty limited endowments put a constraint on the size of a team as well. More team members imply increased production potential (if the production possibility set is superadditive) but requires that s_i':s decrease. A smaller s_i' requires a higher k_i to force the agent to choose a particular level of a_i. Since k_i is bounded we will in the limit have agents choose $a_i = 0$ when team size grows large. Thus, there is a tradeoff which will determine an optimal team size in contrast to the case of certainty.

A resolution to this dilemma may be found by bringing in a principal who has resources to pay bonuses.

Theorem 7: Make the assumptions:

(i) $F(x,a)$ in convex in a.

(ii) $-F_i(x_1 a)/\big(1-F(x_1,a)\big) \to \infty$ as $x \to \infty$.

Then first-best can be enforced at negligible expected cost to a principal with unbounded wealth, even under the restriction that agents' endowments are limited.

Proof: Consider the following scheme:

$$(28) \qquad s_i(x) = \begin{cases} b_i, & x \geqslant \bar{x}, \\ k_i, & x < \bar{x}, \end{cases}$$

where $\Sigma k_i = E(a^*)$ and b_i':s are bonuses paid if a critical level $\bar{x}$ is exceeded. Define b_i:s by:

$$(29) \qquad -(b_i - k_i)\, F_i(x,a^*) + v_i'(a_i^*) = 0, \ i=1,\ldots,n.$$

By (i), (29) guarantees that a* is an equilibrium. The principal's expected revenue is

$$(30) \qquad E(a^*) - \sum_i b_i\big(1-F(\bar{x},a^*)\big) - \sum_i k_i F(\bar{x},a^*) =$$

$$E(a^*) - \sum_i k_i - \sum_i (b_i - k_i)\left(1 - F(\bar{x}, a^*)\right)$$

$$E(a^*) - \sum_i k_i + \sum_i B_i\left(1 - F(\bar{x}, a^*)\right)/F_i(\bar{x}, a^*),$$

where $B_i = -v_i^*(a_i^*)$, are constants. By (ii), this expression goes towards $E(a^*) - \sum_i k_i = 0$ as $n \to \infty$. Q.E.D.

Theorem 6 tells us that a principal in order to restore efficiency, can promise to pay bonuses to agents for exceptionally high outcomes but otherwise pay them their expected product.

The theorems above were developed under risk neutrality assumptions. If agents are risk averse with marginal utility tending to $-\infty$ as wealth decreases, theorem 6 remains valid (cf. Mirrlees, 1974). However theorem 7 appears quite dependent on the risk neutrality assumption. An unverified conjecture is that asymptotic risk neutrality suffices.

5. Uncertainty-Monitoring.

The results above suggest that under certain circumstances efficient team production can be approached via simple penalty or bonus schemes. A critical feature again is that the budget-balancing constraint is broken. Obviously, there remain plenty of situations where such schemes do not remedy moral hazard due to endowment constraints and a second-best solution has to be implemented. Monitoring then becomes a crucial element. I will investigate below what type of monitoring will provide valuable information in the sense that it helps improve welfare.

The set-up assumes a risk-neutral principal and n risk-averse agents.

Since the principal is risk-neutral there are no gains to risk-sharing per se. To the extent output is used in determining agents' payoffs, its value is solely in providing incentives. Put differently, output will merely be used as a signal about the actions taken by the agents.

Let y be the vector of signals observed, so that y can be used as the basis for sharing. This vector may or may not contain x. The distribution of y as a function of a is given by G(y,a), with density g(y,a). The welfare problem can be stated as

$$(29) \qquad \max_{a, s_i(y)} \int \left\{ E(x|y,a) - \sum_i s_i(y) \right\} dG(y,a)$$

$$\text{s.t. (i)} \int u_i\big(s_i(y)\big) dG(y,a) + v_i(a_i) \geq \bar{u}_i, \quad i=1,\ldots,n.$$

$$\text{(ii)} \quad a_i \varepsilon \operatorname*{argmax}_{a_i'} \int u_i\big(s_i(y)\big) dG(y, a_i', a_{-i}) + v_i(a_i'), \quad i=1,\ldots,n.$$

Here, E(x|y,a) is the expected output of x given y and a. It equals, of course x if x is part of y.

I will first consider the single-agent case n=1 and extend results in Harris and Raviv (1979), Holmstrom (1979) and Shavell (1979). From this extension the multi-agent results follow easily. Since n=1, the index i will be temporarily dropped.

<u>Definition</u>: A function T(y) is said to be <u>sufficient for</u> y <u>with respect to a</u>, if there exist functions $h(\cdot) \geq 0$, $p(\cdot) \geq 0$ such that:

$$(30) \qquad g(y,a) = h(y)p\big(T(y),a\big), \text{ for all } y \text{ and } a.$$

Equation (30) is the well-known condition for a sufficient statistic in ordinary statistical decision theory (deGroot, 1970). Note, however, that the action a is not a parameter chosen by nature but by a strategic agent. This notwithstanding, it will be shown below that sharing rules should be based solely on T(y) if and only if T is sufficient for y, which parallels results in statistical decision theory.

Theorem 8. Assume T(y) is sufficient for y with respect to a. Given any scheme s(y), there exists a scheme $\tilde{s}(T(y))$, which weakly Pareto dominates s(y).

Proof. Define $\tilde{s}(T(y))$ as follows:

$$(31) \qquad u(\tilde{s}(T)) = \int_{T(y)=T} u(s(y))g(y,a)dy = p(T,a) \int_{T(y)=T} u(s(y))h(y)dy.$$

By Jensen's inequality,

$$(32) \qquad \tilde{s}(T) \leq \int_{T(y)=T} s(y)g(y,a)dy.$$

From (30) and (31) follows that the agent will enjoy the same expected utility for all a, whether faced with s(y) or $\tilde{s}(T(y))$. Thus, he will not change his action from that which he took under s(y). From this, and (32) follows that the principal is at least as well off with $\tilde{s}(T(y))$ as with s(y).

Q.E.D.

The main import of Theorem 8 is that randomization does not pay if the agent's utility function is separable.[1] Indeed, any pure noise should be filtered away from the agent's sharing rule. To the extent the sharing rule is random it should be through signals that are informative about the agent's action.

The converse to theorem 8 requires brief preparation. We wish to state that if T(y) is not sufficient, then we can improve welfare strictly by observing y. There is a problem, however, with the meaning of an insufficient statistic T(y). Equation (30) may not hold for all a and yet a particular T(y) will be sufficient in the sense that welfare improvements cannot be made. Such is obviously the case if we take T(y) equal to the optimal sharing rule.[2] Moreover, for a fixed action a,equation (30) can always be satisfied by an appropriate choice of $h(\cdot)$ and $p(\cdot,\cdot)$.

To rule out such cases, I will define T(y) as <u>insufficient</u> at a if it is <u>not</u> the case that:

$$(33) \qquad \frac{g_a(y_1,a)}{g(y_1,a)} = \frac{g_a(y_2,a)}{g(y,a)}, \text{ for almost all } y_1,y_2 \varepsilon \{y|T(y) = \text{constant}\}.$$

Note that (33) follows from (30). Conversely, (33) implies (30) (by integrating), if it holds for all a. I will say that T(y) is <u>globally insufficient</u> if (33) is false for all a.

<u>Theorem 9</u>. Assume T(y) is globally insufficient for y with respect to a. Let $s(y) = \tilde{s}(T(y))$ be a sharing rule such that the agent's response is unique. Then there exists a sharing rule $\hat{s}(y)$ which yields a strict Pareto

---

1. Separability of the utility function is crucial for the result that pure randomization does not pay; see Gjesdal (1980). This is further evidence that our sufficient statistic condition and the related theorems are not direct consequences of the decision theoretic results (i.e. Blackwell's theorem).
2. I am indebted to Steve Ross for this observation.

improvement. Moreover $\hat{s}(y)$ can be chosen so as to induce the same action response as $s(y)$ unless $s(y)$ is a constant.

Proof. Since $T(y)$ is globally sufficient, there exists a T_1 and sets Y_{11}, Y_{12} which are disjoint and subsets of $Y_1 = \{y | T(y) = T_1\}$ such that

$$(34) \qquad \frac{g_a(Y_{11},a)}{g(Y_{11},a)} \neq \frac{g_a(Y_{12},a)}{g(Y_{12},a)},$$

where a is the agent's response to $s(y)$. Since $s(y)$ is not constant, there exists a $T_2 \neq T_1$ such that the set $Y_2 = \{y | T(y) = T_2\}$ is non-empty and $\tilde{s}(T_1) \neq \tilde{s}(T_2)$. Define the following variation:

$$\hat{s}(y) = \tilde{s}\big(T(y)\big) + I_{11}(y)\ ds_{11} + I_{12}(y)\ ds_{12} + I_2(y)\ ds_2,$$

where $I_{11}(y)$ is the indicator function for the event $\{y \varepsilon Y_{11}\}$ and similarily for $I_{12}(y)$, $I_2(y)$. Let $g(Y_{11},a) = \Pr\{y \varepsilon Y_{11} | a\}$ and similarily for $g(Y_{12},a)$, $g(Y_2,a)$.

The effect on the principal's and the agent's welfare (exluding any change in action) from a change to $\hat{s}(y)$ is given by

$$(35) \qquad \Delta P = -\big[ds_{11}g(Y_{11},a) + ds_{12}g(Y_{12},a) + ds_2 g(Y_2,a)\big],$$

$$(36) \qquad \Delta A = u_1'\big(ds_{11}g(Y_{11},a) + ds_{12}g(Y_{12},a)\big) + u_2' ds_2 g(Y_2,a).$$

Here $u_1' = u_1'\big(\tilde{s}(T_1)\big)$ and $u_2' = u_2'\big(\tilde{s}(T_2)\big)$. Since $\tilde{s}(T_1) \neq \tilde{s}(T_2)$, $u_1' \neq u_2'$. Assume for concreteness that $u_2' < u_1'$.

The effect on the agent's action from this variation is given by

$$(37) \qquad \Delta a = u_1'\left(ds_{11}g_a(Y_{11},a) + ds_{12}g_a(Y_{12},a)\right) + u_2' ds_2 g_a(Y_2,a).$$

Choose ds_{11}, ds_{12}, ds_2 as follows. Let $\Delta P = 0$, and substitute for ds_2 in (36) requiring:

$$(38) \qquad \Delta A = (u_1' - u_2')\left(ds_{11}g(Y_{11},a) - ds_{12}g(Y_{12},a)\right) = k > 0$$

Keeping ds_2 fixed, require furthermore:

$$(39) \qquad ds_{11}g_a(Y_{11},a) + ds_{12}g_a(Y_{12},a) = -u_2'/u_1' \; ds_2 g_a(Y_2,a)$$

This makes $\Delta a = 0$. The system (38), (39) has a unique solution because (34) implies that the determinant is non-zero. This shows that ds_{11}, ds_{12}, ds_2 can be chosen so that $\Delta P = 0$, $\Delta A > 0$ and $\Delta a = 0$, in other words so that the principal is no worse off and the agent is better off in terms of risk-sharing while the action remains unchanged. Q.E.D.

Theorem 9 is the converse of 8 and says that if T(y) is not a globally sufficient statistic for y, we can do strictly better by using all of y instead of T(y) as a basis for the sharing rule. The intuition, of course, is that y reveals more information about a than T(y) does if and only if T(y) is not sufficient for y.

Theorem 9 differs in two respects from a corresponding theorem in Holmstrom (1979). Rather importantly for the application to the multi-agent case, the improvement in welfare is here achieved, not by forcing a change in the agent's action, but by keeping the action the same and improving risk-sharing instead. For this, s(y) had to be assumed non-constant, since obviously a constant function cannot be improved upon for risk-sharing.

Secondly, theorem 9 is more general, as my earlier result concerned the following special case:[3]

Application 1: Informative signals.

Consider two information systems $I_1 = x$ and $I_2 = (x,z)$. In I_1 we observe the outcome x, in I_2 we observe the outcome x and some additional signal z. Suppose $T(x,z) = x$ is not sufficient for (x,z). Then I called z an informative signal in Holmstrom (1979), and showed that z can be used to design a sharing rule s(x,z) which is strictly better than any s(x). This, of course, is a special case of theorem 9.

Application 2: Finer signals

Grossman and Hart (1980) and Gjesdal (1980) discuss the value of different information systems from the point of view of Blackwell's notion of fineness. The following discussion will reveal its relationship to the sufficiency property used here.[4]

Let $I_1 = y_1$ and $I_2 = y_2$ be two information systems. I_1 is said to be finer than I_2 if we have for all y_2,a:

$$(40) \qquad g_2(y_2,a) = \int h(y_2,y_1)\, g_1(y_1,a)dy_1,$$

where g_2 and g_1 are the marginal densities of y_2 and y_1. Condition (40) looks much like (30) but is conceptually quite different, since (40) deals with

3. Gjesdal (1980) provides a closely related extension of my earlier informativeness condition.
4. I am grteful to Paul Milgrom for discussing this relationship.

marginal distributions only, whereas (30) presumes a joint distribution for $y=(y_1, y_2)$.

However, the solution to the Pareto problem with I_1 or I_2 does not depend on the joint distribution of y_1 and y_2. So we may define one by letting

$$(41) \qquad g(y_1,y_2,a) = h(y_2,y_1)\, g_1(y_1,a).$$

This will have as its marginals g_1 and g_2 (as in (40)). But (41) says that $T(y_1,y_2) = y_1$ is sufficient for (y_1,y_2). Hence, we get as good results by basing the sharing rule on y_1 as on (y_1,y_2). Therefore, $I_1 = y_1$, is as good an information system as $I = (y_1,y_2)$. But, of course, I can be no wore than $I_2 = y_2$. The conclusion is that I_1 is as good as I_2, if I_1 is finer than I_2 in the Blackwell sense. The Blackwell finessess result can therefore be viewed as a corollary to the sufficiency results of theorems 8 and 9.

Let me now turn to the multi-agent case. Condition (30) will still define the notion of a sufficient statistic.

Theorem 8': Assume $T(y)$ is sufficient for y with respect to a. Given a set of sharing rules $\{s_i(y)\}$, there exist another set $\{\tilde{s}_i(T(y))\}$ which weakly Pareto dominates $\{s_i(y)\}$.

Proof: As theorem 8. Q.E.D.

Theorem 9': Assume $T(y)$ is globally insufficient for y with respect to a. Let $s_i(y) = \tilde{s}_i(T(y))$, $i=1,\ldots,n$, be a set of sharing rules for which the Nash equilibrium is unique. Then there exists sharing rules $\{\hat{s}_i(y)\}$ which yield a strict Pareto improvement.

Proof: As theorem 9. Note that since the variations $\hat{s}_i(y)$ do not change the other agent's actions, we need not worry about the Nash equilibrium changing.

Q.E.D.

If one had employed the line of proof in Holmstrom (1979), in which the agent's action is changed, one would have had to be concerned with the effects of one agent's change in action on the other agents' choice of action.[5]

Application 3: Dividing output in teams

Consider the team production case. Let total output be $x(a,\theta)$. A finer information system is obtained by measuring $x(a,\theta)$ as the sum:

$$x(a,\theta) = x_1(a,\theta) + x_2(a,\theta). \tag{42}$$

From the general results obtained above, we know that $T(x_1,x_2) = x_1 + x_2 = x$ is sufficient for (x_1,x_2) and hence as good as observing x_1, x_2 separately if and only if:

$$g(x_1,x_2,a) = h(x_1,x_2)\ p(x_1 + x_2,a), \tag{43}$$

where g is the joint distribution of (x_1,x_2) and p is the distribution of x induced by θ.

When will (43) hold? Suppose,

5. Baiman and Demski (1980) prove a special case of theorem 9' using an extra assumption that appears unnecessary. This assumption is called for, since they do not exploit the possibility of improving risk-sharing benefits, but rather use the line of proof in Holmstrom (1979).

$$x_1(a,\theta) = m\big(x(a,\theta),\ \ell(\theta)\big)$$

(44)

$$x_2(a,\theta) = x(a,\theta) - m\big(x(a,\theta),\ \ell(\theta)\big)$$

for some functions $m(\cdot)$ and $\ell(\cdot)$ such that $x(a,\theta)$ is <u>independent</u> of $\ell(\theta)$. (44) maps x into two random variables x_1, x_2 which are independent of x and it follows easily that (43) holds. The converse is also true. Condition (43) implies (44) with $\ell(\theta)$ indepedent of $x(a,\theta)$. From our general results we know then that (44) implies that it is valueless to observe x_1, x_2 rather than $x = x_1 + x_2$, which is quite intuitive as we obtain x_1 and x_2 by going from x to a given (possibly) random split.

However, whenever x_1 and x_2 divide x so that we can infer something more about a from seeing x_1, x_2 rather than x, such a finer accounting system is valuable.

6. Relative Performance Measures

In this final section I will consider the case where total output is given by

(45) $$x(a,\theta) = \sum_i x_i(a_i,\theta).$$

All x_i: s are assumed to be observed. We do not have team production since each x_i depends only on a_i and some exogenous random shock θ. If θ were not random, efficiency would be easily achieved by holding each agent i responsible for his output x_i in accordance with the general principles of responsibility accounting. We saw this in section 3.

What will interest us here is when it will be desirable to depart from this general principle and actually have s_i, the sharing rule of i, depend on the vector of outcomes $x=(x_1,\ldots,x_n)$ rather than x_i alone. We frequently observe agents being evaluated based on peer performance. Almost in all organizations agents compete with each other in one form or another. Sometimes there is an explicit prize for the best ones, as for instance, among sales personnel ("salesman of the month" - awards etc.). The special case of rank order tournaments, in which relative performance as measured by rank alone, has been analyzed by Lazear and Rosen (1980). Further examples are provided by the recent performance incentives for executives in which executive performance is compared with that in competing firms.

The rationale for relative performance evaluation is easily understood in light of the results on the value of information that are given in this paper and in Holmstrom (1979). We have found that essentially any information that is useful for inferring the agent's action is valuable and should enter the contract (ignoring transactions and information costs). Therefore, if the x_i:s are correlated with each other (through θ), s_i should depend not just on x_i but on the whole vector x. This will help reduce the randomness in the agent's share that stem from circumstances outside his control. It makes sense (both intuitively and theoretically) to evaluate executive performance against performance of other firms, because economy-wide shocks are thereby absorbed. Recessions have made managerial stock option-plans pay off poorly independently of managerial performance and it is possible that the above cited move towards performance incentives was triggered by this fact.[6]

Before considering the form sharing rules may take when x_i:s are correlated, let me first dispense with the independent case.

---

6. A similar result is proved in Biaman and Demski (1980).

Theorem 10: Suppose x_i:s are independent. Then, optimal sharing rules $s_i(x)$ will only depend on x_i.

Proof: If x_i:s are independent, then

$$(46) \qquad f(x,a) = \prod_{i=1}^{n} f_i(x_i,a_i).$$

Considering agent i, we see at once that (46) implies that $T_i(x) = x_i$ is sufficient for x with respect to a_i. By theorem 8, it will be enough to let s_i depend on x_i alone. Q.E.D.

The import of theorem 10 is that it does not pay to force agents to compete with each other unless there is some common underlying uncertainty. The benefits from competition itself are nil. What is of value, is the information that may be gained from peer performance. Competition among agents is a consequence of exploiting this information.

I have not investigated whether the converse of theorem 10 is true; does dependence through Θ always imply that it is suboptimal to base the agent's reward on his own outcome alone. The following general condition for when rewards should depend only on x_i for agent i follows from the previous analysis:

$$(47) \qquad f_i(x,a) = h_i(x,a_{-i})\ p_i(x_i,a).$$

When (47) holds nothing can be gained by including x_{-i} in s_i. Note that a_{-i} can enter h_i and p_i since we can in equilibrium infer the actions of the agents, given any particular schedule.

The opposite extreme to independence is complete dependence. This can be phrased as the condition that given x_i, x_j does not vary when actions are fixed. We have the obvious:

Theorem 11: Assume x_i:s are fully dependent. Then there exists a first-best Nash equilibrium.

Proof: The following scheme will yield a first-best Nash equilibrium:

$$s_i(x_i,x_j) = \begin{cases} w_i^*, \text{ if } x_i = \phi_i\left(x_j,(a_j^*, \Theta),a_i^*,a_j^*\right) \\ b_i, \text{ if not,} \end{cases}$$

where b_i is a sufficient penalty, w_i^* is the first-best (constant) payment, and a^* is the first-best action and $\phi_i(\cdot)$ is the assumed deterministic relationship between x_j and x_i, a_i, a_j.

Q.E.D.

This theorem corresponds to theorem 4 in the certainty case. It could also have been shown that as x_i:s become more dependent, first-best can be approached in the limit.

A general characterization of how information about peer performance should be optimally used with many agents can be formally developed as an obvious extension of the characterizations in Mirrlees (1976) or Holmstrom (1979). The validity of such a characteriztion depends at least for the time being on rather restrictive assumptions as explained in Grossman and Hart (1980). I will not pursue that issue here. My purpose is to point out how the sufficient statistic conditions developed above can be used to rationalize schemes which only use aggregate information about peer performance. I will also consider what happens when we have a large number of agents.

I will restrict attention to the following two particular output structures:

I: $x_i(a_i,\theta) = a_i + \eta + \varepsilon_i$, $i=1,\ldots,n$,

II: $x_i(a_i,\theta) = a_i(\eta + \varepsilon_i)$, $i=1,\ldots,n$,

where $\theta = (\eta, \varepsilon_1,\ldots,\varepsilon_n)$ is a random vector.

Theorem 12: Let the technology be given by either I or II. Assume $\eta,\varepsilon_1,\ldots,\varepsilon_n$ are independent and normally distributed. Let $\bar{x} = \Sigma \frac{\tau_i}{\bar{\tau}} x_i$ be the weighted average of the agents' outcomes, where τ_i is the precision (the inverse of the variance) of ε_i and $\bar{\tau} = \Sigma\tau_i$. An optimal set of sharing rules $\{s_i(x)\}$ will have s_i depend on $\bar{x}$ and x_i alone; in other words, each agent will be judged relative to a weighted average, which will be equal to the mean outcome if precisions are equal.

Proof: Consider technology I. The joint density function for x given a is:

$$(48) \qquad f(x,a) = K \int \exp \left\{-\frac{1}{2}\left[\sum_j \tau_j(x_j - a_j - \mu_j - \eta)^2 + \tau_o(\eta - \mu_o)^2\right]\right\} d\eta$$

where K is a constant, τ_o is the precision and μ_o is the mean of η, and τ_j,μ_j are the precision and mean of ε_j. In view of theorems 8 and 9 we need to show that we can write (48) in the form:

$$(49) \qquad f(x,a) = h_i(x,a_{-i})\; p_i\; (\bar{x},x_i,a)$$

for each $i=1,\dots,n$. The reason we may allow h_i and p_i to depend on a_{-i} is that the other agents' actions can be taken as given when designing agent i's scheme s_i (confer proofs of theorems 8' and 9'). Let

$$\bar{z}_{-i} = \sum_{k\neq i} \tau_k/\tau_{-i}(x_k - a_k - \mu_k), \qquad \bar{\tau}_{-i} = \sum_{k\neq i} \tau_k.$$

Then we can write:

$$\sum_j \tau_j(x_j - a_j - \mu_j - \eta)^2 =$$

$$\sum_{j\neq i} \tau_j(x_j - a_j - \mu_j - \bar{z}_{-i} + \bar{z}_{-i} - \eta)^2 + \tau_i(x_i - a_i - \mu_i)^2 =$$

$$\sum_{j\neq i} \tau_j(x_j - a_j - \mu_j - \bar{z}_{-i})^2 + (n-1)(\bar{z}_{-i} - \eta)^2 + \tau_i(x_i - a_i - \mu_i)^2.$$

Substituting this expression into (48), we find upon integrating over η that we can write

$$f(x,a) = h_i(x,a_{-i})\,\hat{p}_i(\bar{z}_{-i},x_i,a).$$

However, since

$$\bar{z}_{-i} = (\bar{\tau}\bar{x} - \tau_i x_i)/\tau_{-i} - \sum_{k\neq i} \tau_k/\bar{\tau}_{-i}(a_k + \mu_k),$$

$$\hat{p}_i(\bar{z}_{-i},x_i,a) = p_i(\bar{x},x_i a),$$

completing the proof of writing (48) in the form (49).

The proof for technology II is similar and is omitted.

Of course, if $\tau_i = \tau$ for all i, then $\bar{x} = \frac{1}{n} \Sigma x_i$, i.e. the average of observations. Q.E.D.

Notice that theorem 11 does not make the claim that s_i should depend on $x_i - \bar{x}$, only that it will have the form $s_i(x_i, \bar{x})$. The fact that the outputs of different agents are generally weighted differently in calculating $\bar{x}$, reflects possible differences in values of these information sources. If ε_j has high precision (low variance) then x_j tells rather sharply the value of η and should count more in the average. This is another way of saying that x_j:s which are correlated strongly with x_i should be more significant indicators in evaluating agent i:s performance. Conversely, as $\tau_j \to 0$, x_j will essentially tell nothing about η because of the noise in ε_j and hence should count very little.

Theorem 12 suggests that sometimes (perhaps quite often) an aggregate measure like the weighted average of peer performance will capture all the relevant information about the common uncertainty. This provides a rationale for the common practice of comparing performance against peer aggregates, though, of course, the sufficiency of a weighted average is specific to the normal distribution. Other distributions will have other sufficient statistics.

At this point it is appropriate to comment on the use of rank-order tournaments (Lazear and Rosen, 1980). A rank-order tournament awards agents merely on their performance rank, not on the value of the output itself. With n agents there are n prizes $w_1 \geq \ldots \geq w_n$. The agent with the highest output gets w_1, second highest gets w_2 and so on.

From theorem 10 follows that if the agents' outcomes are unrelated, then rank-order tournaments will perform worse than rewarding agents based on their individual outcomes alone. Pitching agents against each other will only result in more randomness in the reward scheme without any gains in the power of inference about actions. On the other hand, rank-order tournaments may be valuable if outcomes are related as was first noted by Lazear and Rosen. The analysis above supports this contention. But it should be observed that rank-order tournaments may be informationally quite wasteful if performance levels can be measured cardinally rather than ordinally. It is clear that the mapping from the agents' outcomes $x=(x_1,..,x_n)$ into the statistic $T(x) = (k_1(x),...,k_n(x))$, where $k_i(x)$ is the rank-order of agent i, is not a sufficient statistic for a, except in trivial cases. Therefore, theorem 9' tells us that there should be a better way of making use of x than what the rank-order tournament does.

Let me finally turn to an analysis of large teams under the simplifying (but unnecessary) assumption that the technology is given by I or II. First, note that if we knew η ex post, this common uncertainty could and should be "filtered" away to yield an improved solution to the agency problem (c.f. the discussion of finer signals in Application 2). And if we knew η ex post then there would be no need to compare individual agents' outputs, since conditional on η they are independent (theorem 10). Thus, the solution to the incentive problem with n agents coincides with the solution of the n individual agency problems when η is known ex post. For these individual problems we will have the optimal schemes depend on $a_i + \varepsilon_i$ (for I) and $a_i\ \varepsilon_i$ (for II), since the observation of η will allow us to observe these variables.

Now, η is not observed ex post. But it is intuitively clear that as the number of agents grows large, we can essentially observe η by infering it from the independent signals about η provided by the x_i:s. Therefore we would expect that with many agents we will be able to achieve approximately the same solution as if there were no common uncertainty at all. This is correct:

Theorem 13: Consider technology I or II where $\eta, \varepsilon_1, \ldots, \varepsilon_n$ are independent and ε_i:s identical. Then the solution to the single agent problem without common uncertainty (i.e. $\eta = 0$) can be approximated arbitrarily closely as the number of agents grows large.

Proof: Consider technology I. Let

$$q_j = \eta + \varepsilon_j, \qquad j=1,\ldots,n.$$

$$\bar{q}_{-i} = \frac{1}{n-1} \sum_{j \neq i} q_j, \qquad i-1,\ldots,n.$$

Let $s_i^*(x_i)$ be the optimal solution to the single-agent problem when there is no common uncertainty, and let a_i^* be the agent's optimal response.

By the strong law of large numbers, $\bar{q}_i$ goes a.s. to η. Therefore,

$$\int u_i\left(s_i^*(a_i^* + \eta + \varepsilon_i - \bar{q}_{-i})\right) dP(\eta, \varepsilon_1, \ldots, \varepsilon_n)$$

converges to

$$\int u_i\left(s_i^*(a_i^* + \varepsilon_i)\right) dP(\varepsilon_i).$$

Provided a_i^* is a unique solution to

$$\max_{a_i} \int u_i\left(s_i^*(a_i + \varepsilon_i)\right) dP(\varepsilon_i) + v_i(a_i),$$

we find that for large enough n, the agent will choose an action arbitrarily close to a_i^* when solving

$$\max_{a_i} \int u_i\big(s_i^*(a_i^* + \eta + \varepsilon_i - \bar{q}_i)\big)\, dP(\eta, \varepsilon_1,\ldots,\varepsilon_n).$$

Since $\bar{q}_{-i}$ can be inferred from the other agents' outcomes by calculating $x_j - a_j = \eta + \varepsilon_j$, where a_j is the response of agent j, this proves the claim. The proof of technology II is similar.

Q.E.D.

It is clear that constraining attention to technology I and II is quite inessential. A number of other specifications will yield the same result and a quite general statement of the theorem should be possible to prove. In particular, the assumption that ε_i:s are identically distributed can be avoided in a way similar to theorem 12. The assumption that ε_i:s are independent while η is common is a natural way of providing a limit to what can be gained from relative performance evaluations.

Results analogous to theorem 13 could be arrived at with different informational assumptions. For instance, suppose the agent observes $\eta + \varepsilon_i$, i.e. η with a noise term, before taking his action. I conjecture that with many agents one would be able to come close to the solution of the same problem with η ex post known. Alternatively, if agents observe η before taking their actions, we can in the limit get the same solution as when η is commonly observed ex ante (this is simple to prove for the two technologies I and II).

The result that one may use relative performance measures to filter away common uncertainties has an interesting connection to financial theory. A standard model of financial markets is the capital asset pricing model. One of its normative implications is that investments carried out by firms should be decided upon without reference to the investment's idiosyncratic risk as the market through diversification can neutralize such risk (see e.g. Mossin, 1969). Investment decisions should instead be made with reference to systematic risk alone.

This prescription is in apparent contradiction with the analysis above. What we found was that the manager need not worry about systematic risk as that risk will be filtered away. Instead, he will have to be concerned with the idiosyncratic risk of an investment as that risk will have a major influence on how the firm and the managerial labor market will evaluate his managerial performance. This conflict of interests has not been studied here, but deserves attention in future research.

7. Summary

This paper has been concerned with ways of alleviating moral hazard when there are many agents. In contrast to the situation with a single agent, a multi-agent organization may experience moral hazard problems even under certainty. An analysis of the certainty case was first presented, which served as an introduction to the subsequent discussion of the more realistic case of uncertainty.

The paper has not attempted a characterization of a solution to moral hazard in teams. Instead, its objective has been to explore two common ways of alleviating the problem. One is to separate ownership and labor so as to neutralize externalities in joint production. The other is to use monitoring, richer measures of output and other additional information to discern more accurately the actions of agents.

Despite its simplicity, the model provides us with a number of positive and normative implications. The main ones are summarized here:

1. When there is no state uncertainty, moral hazard can be resolved by a threat to punish agents as a group if the appropriate level of output is not attained. This can be done without violating endowment constraints independently of the number of agents. Thus, no monitoring is needed and no constraints on firm size are imposed. The key is to break the budget-balancing constraint, which is suggestive of the value of separating ownership and labor in organizing production.

2. If one attempts to resolve the certainty case through monitoring while maintaining budget-balancing, it may suffice to have only two measures of output. However, for an efficient monotone control system in which shares are positively related to output measures, it is necessary to have as many measures as there are agents and these measures should effectively discern actions. This is in accordance with common principles of control, which attempt to hold agents responsible only for what they alone can have influence over.

3. The use of group penalties may be an effective way to control joint production even under uncertainty, but in contrast to the certainty case limited endowments will constrain the effectiveness of such a solution. These constraints will be felt more the larger the number of agents, which generally implies that the size of a team which can be effectively controlled through penalties is limited. This points to inefficiencies in large partnerships, when there is uncertain production.

4. Since group penalties will frequently be an insufficient means to

control moral hazard under uncertainty, monitoring becomes essential. As in the single-agent case valuable monitoring is characterized by the condition that it improves the ability to infer the agents' actions - a condition that extends a well-known result in statistical decision theory to the realm of strategic games.

5. Relative performance evaluation, as commonly observed, can be understood as a way of using information about state uncertainty efficiently. In accordance with the above mentioned necessary and sufficient criterion for what type of additional information is valuable, one finds that relative performance evalatuion is valuable if and only if agents face some common uncertainties. Thus, inducing competition among agents does not have intrinsic value. Rather, competition is a consequence of efficient information usage. An example of relative performance evaluation across firms is the new trend towards performance incentives for executives. These are rationalized by the analysis as a way of reducing unnecessary risk imposed on executives from factors outside their control.

6. Another example of performance evaluation that can be rationalized by the theory is the use of rank-order tournaments that have recently been discussed in the literature. It was noted, though, that they generally fail to make efficient use of available information.

7. Relative performance evaluation may often involve an index of comparison which aggregates peer performance, for instance through an average, without any loss of efficiency.

8. With a large reference group of agents, the cost of uncertainty that comes from a common source, can essentially be eliminated through relative performance measures. What remains to be coped with are the idiosyncratic risks of individual agents. This implies, from an agency

theoretic point of view, a particular concern for idiosyncratic risks in managerial decision making as these (but not systematic risks) will enter into the evaluation process of managers. This stands in contrast to implications of the capital asset pricing model, which prescribes the opposite, namely to ignore idiosyncratic risks and only to be concerned with systematic risk when making investment and production decisions.

REFERENCES

Alchian A., and H. Demsetz (1972), "Production, Information Costs and Economic Organization", American Economic Review, Vol. 62

Baiman S., and J. Demski (1980), "Economically Optimal Performance Evaluation and Control Systems", forthcoming in Journal of Accounting Research

Gjesdal, F. (1980), "Information and Incentives; the Agency Incentive Problem", DP 12/80, Norwegian School of Economics and Business Administration, Bergen, Norway, June.

Grossman S., and O. Hart (1980), "An Analysis of the Principal-Agent Problem", CARESS WP No. 80-17, University of Pennsylvania, April.

Harris M., and A. Raviv (1979), "Optimal Incentive Contracts with Imperfect Information", Journal of Economic Theory.

Holmstrom B. (1979) "Moral Hazard and Observability" Bell Journal of Economics, Spring.

Holmstrom B. (1980), "A Discussion of Baiman and Demski:' Economically Optimal Performance Evaluation and Control Systems'", forthcoming in Journal of Accounting Research.

Horngren C. (1972), Cost Accounting: A Managerial Emphasis (3rd ed.), Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Lazear E., and S. Rosen (1980), "Rank-Order Tournaments as Optimum Labor Contracts", mimeographed, University of Chicago, June.

Marschak J. and R. Radner (1972). Economic Theory of Teams, Yale: Yale University.

Milgrom P (1980), "Good News and Bad News: Some Representation Theorems", DP, Northwestern University.

Mirrlees, J. (1974), ""Notes on Welfare Economics, Information and Uncertainty", in Balch, McFadden and Wu, Essays on Economic Behavior Under Undertainty: North Holland Amsterdam.

Mirrlees, J. (1976), "The Optimal Structure of Incentives and Authority Within an Orgamization", Bell Journal of Economics, Spring.

Mossin J. (1969), "Security Pricing and Investment Criteria in Competitive Markets", American Economic Review

Shavell S. (1979), "Risk Sharing and Incentives in the Principal and Agent Relationship", Bell Journal of Economics, Spring.