Moral hazard in teams

Bengt Holmstrom*

This article studies moral hazard with many agents. The focus is on two features that are novel in a multiagent setting: free riding and competition. The free-rider problem implies a new role for the principal: administering incentive schemes that do not balance the budget. This new role is essential for controlling incentives and suggests that firms in which ownership and labor are partly separated will have an advantage over partnerships in which output is distributed among agents. A new characterization of informative (hence valuable) monitoring is derived and applied to analyze the value of relative performance evaluation. It is shown that competition among agents (due to relative evaluations) has merit solely as a device to extract information optimally. Competition per se is worthless. The role of aggregate measures in relative performance evaluation is also explored, and the implications for investment rules are discussed.

1. Introduction

■ Orthodox economic theory has little to offer in terms of understanding how nonmarket 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 explicitly recognizes that agents have limited and different information, is a welcome invention, which promises to be helpful in understanding the intricacies of organizational design. Particularly important in this context are questions concerning the control of agents' incentives, which to a large degree dictate the structure of organizations and set the limits of its performance potential (Arrow, 1974).

The members of an organization may be seen as providing two kinds of services: they supply inputs for production and process information for decisionmaking. 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 for 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 he privately observes.

This article is concerned with moral hazard in teams.[1] By a team I mean rather

---

* Northwestern University.

This article 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 comments by Froystein Gjesdal, Milt Harris, Richard Kihlstrom, Ed Lazear, Paul Milgrom, Steve Ross, Karl Shell, an anonymous referee, and seminar participants at Yale, the Universities of Pennsylvania, California (San Diego), Chicago and Iowa. Financial support from the Center for Advanced Studies in Managerial Economics at Northwestern and from a J. L. Kellogg and NSF research grant is gratefully acknowledged.

[1] Earlier work on moral hazard in a single-agent setting includes Wilson (1969), Spence and Zeckhauser (1971), Ross (1973), Mirrlees (1976), Harris and Raviv (1979), Shavell (1979), Holmstrom (1979), and Grossman and Hart (1980). Models with many agents have been studied by Baiman and Demski (1980), Lazear and Rosen (1981), and Radner (1980). Papers contemporaneous and independent of this article are Atkinson and Feltham (1980), Green and Stokey (1982), and Nalebuff and Stiglitz (1982).

loosely a group of individuals who are organized so that their productive inputs are related.[2] The objective of the analysis is to derive some positive and normative implications regarding the organization of production with many agents whose inputs are imperfectly observed. The analysis will focus on two features that are specific to multiagent organizations: the problem of free riding when there is joint production and the role of competition in controlling incentives.

I start by considering the free-rider issue (Section 2). In contrast to the single-agent case, moral hazard problems may occur even when there is no uncertainty in output. The reason is that agents who cheat cannot be identified if joint output is the only observable indicator of inputs. Indeed, I show that noncooperative behavior will always yield an inefficient outcome if joint output is fully shared among the agents.

In a well-known paper, Alchian and Demsetz (1972) argue that efficiency can (and will) be restored by bringing in a principal who monitors the agents' inputs. My first point will be that the principal's role is not essentially one of monitoring. I show that under certainty group incentives alone can remove the free-rider problem. Such incentives require penalties that waste output or bonuses that exceed output. In both cases the principal is needed, either to enforce the penalties or to finance the bonuses. Thus, the principal's primary role is to break the budget-balancing constraint. The fact that capitalistic firms feature separation of ownership and labor implies that the free-rider problem is less pronounced in such firms than in closed organizations like partnerships.

Group incentives can also work quite well under uncertainty, but their effectiveness will be limited if there are many agents and if the agents are risk averse. This makes monitoring important. As a refinement of earlier results on monitoring in the single-agent context (Holmstrom, 1979; Shavell, 1979), I show in Section 3 that agents' sharing rules can, without loss in welfare, be written on a statistic of all observations if and only if this statistic is sufficient in the sense of statistical decision theory. This result parallels those in decision theory, but is not the same, since the context is a strategic game where agents rather than Nature choose the parameters of the distribution.

The most interesting implications of the sufficient statistic condition concern relative performance evaluation. These are explored in Section 4. One finds that relative performance evaluation will be valuable if one agent's output provides information about another agent's state uncertainty. Such will be the case if and only if agents face some common uncertainties. Thus, inducing competition among agents by tying their rewards to each other's performance has no intrinsic value. Rather, competition is the consequence of the efficient use of information.

An example of performance evaluation that can be rationalized in this fashion is the use of rank-order tournaments, which have recently been studied by Lazear and Rosen (1981). But since rank order is not a sufficient statistic for individual output except in special circumstances, rank-order tournaments will generally imply an inefficient use of available information. In contrast, I show that aggregate measures like peer averages may often provide sufficient information about common uncertainties and thus schemes that compare agents with such aggregate measures will be efficient. An example of relative performance evaluation of this kind is given by the new executive incentive packages, which base rewards on explicit comparisons with firms within the same industry.

For another application of the sufficient statistic condition, I show that the cost of common uncertainties can be essentially eliminated when the number of agents grows large. What remains to be coped with are the idiosyncratic risks of individual agents. This implies a particular concern for idiosyncratic risks in managerial decisions about invest-

[2] Although "team" is the most natural term to describe the setting I study, it should be noted that the term "team" has another precise technical meaning within team theory (Marschak and Radner, 1972), which is inconsistent with incentive problems. My model is also similar in structure to what Wilson (1968) called a syndicate.

ments, since only idiosyncratic risk will enter the process of evaluating managers. Thus, an agency-theoretic perspective implies changes in the normative implications of standard asset pricing models, which hold that systematic risk is the only risk that carries a premium.

The article concludes with some remarks on facets of the multiagent problem that are not addressed here but deserve consideration.

2. Group incentives and the role of the principal

■ **Certainty.** Consider the following simple model of team production. There are n agents. Each agent, indexed i, takes a nonobservable action $a_i \in A_i = [0, \infty)$, with a private (nonmonetary) cost $v_i: A_i \rightarrow \mathbb{R}$; v_i is strictly convex, differentiable, and increasing with $v_i(0) = 0$. Let $a = (a_1, \ldots, a_n) \in A \equiv \overset{n}{\underset{i=1}{\mathsf{X}}} A_i$ 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 $x: A \rightarrow \mathbb{R}$, which must be allocated among the agents. The function x is assumed to be strictly increasing, concave, and differentiable 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 additively separable in money and action, and linear in money. Hence, it is of the form $u_i(m_i, a_i) = m_i - v_i(a_i)$. Agents' initial endowments of money are finite and normalized to be zero.

The question is whether there is a way of fully allocating the joint outcome x so that the resulting noncooperative game among the agents has a Pareto optimal Nash equilibrium. That is, we ask whether there exist sharing rules $s_i(x) \geqslant 0$, $i = 1, \ldots, n$, such that we have budget-balancing

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

and the noncooperative game with payoffs,

$$s_i(x(a)) - v_i(a_i), \qquad i = 1, \ldots, n, \tag{2}$$

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

$$a^* = \underset{a \in A}{\operatorname{argmax}}\ [x(a) - \sum_{i=1}^{n} v_i(a_i)]. \tag{3}$$

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

$$s'_i x'_i - v'_i = 0, \qquad i = 1, \ldots, n, \tag{4}$$

where $x'_i \equiv \partial x/\partial a_i$. Pareto optimality implies that

$$x'_i - v'_i = 0, \qquad i = 1, \ldots, n. \tag{5}$$

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

$$\sum_{i=1}^{n} s'_i = 1. \tag{6}$$

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

Theorem 1. There do not exist sharing rules $\{s_i(x)\}$ which satisfy (1) and which yield a^* as a Nash equilibrium in the noncooperative game with payoffs (2).

Proof: See Appendix.

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—that is, (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, some agent always has an incentive to capitalize on this control deficiency.

The result indicates that in closed (budget-balanced) organizations like a labor-managed firm or a partnership, free-rider problems are likely to lead to an insufficient supply of productive inputs like effort. This observation is the starting point for Alchian and Demsetz' (1972) well-known theory of the firm. They argue that the inefficiency of a partnership will cause an organizational change. To secure a sufficient supply of effort, firms should hire a principal to monitor the behavior of agents. The monitor should be given title to the net earnings of the firm so that he has the proper incentives to work. Such an arrangement will restore efficiency. At the same time, it will change the partnership into a capitalistic firm with the monitor acting effectively as the owner.

There is a simpler solution, however, at least under certainty. The free-rider problem is not solely the consequence of the unobservability of actions, but equally the consequence of imposing budget-balancing. If we relax (1) to read:

$$\sum_i s_i(x) \leqslant x, \tag{7}$$

then there will exist efficient Nash equilibria.

Theorem 2. There exists a set of feasible sharing rules $s_i(x) \geqslant 0$, $i = 1, \ldots, n$, satisfying (7), such that a^* is a Nash equilibrium.

Proof: Let

$$s_i(x) = \begin{cases} b_i & \text{if} \quad x \geqslant x(a^*) \\ 0 & \text{if} \quad x < x(a^*). \end{cases} \tag{8}$$

Choose b_i's so that $\sum_i b_i = x(a^*)$ and $b_i > v_i(a_i^*) > 0$. This is possible because $x(a^*) - \sum_i v_i(a_i^*) > 0$ by Pareto optimality. It is clear that a^* is a Nash equilibrium when sharing rules are chosen in this fashion. *Q.E.D.*

The purpose of relaxing the budget-balancing constraint in (1) is to permit group penalties that are sufficient to police all agents' behavior. It is worth noting that the scheme in (8) does not violate individual endowment constraints and works independently of the team size. This feature is specific to the certainty case.

Schemes like (8) are observed in some types of contracting with labor teams. Usually it takes the form of a flat wage and a group bonus to be paid if a target is attained (whether one views the discontinuity in (8) as a bonus or a penalty appears immaterial). An extreme example of group punishment is the dismissal of the board of directors of a firm.

In a dynamic context the punishment in (8) can be interpreted as a threat to discontinue cooperation. There is a problem, however, with enforcing such group penalties if they are self-imposed by the worker team. Suppose something less than $x(a^*)$ is produced. *Ex post* it is not in the interest of any of the team members to waste some of the outcome. But if it is expected that penalties will not be enforced, we are back in the situation with budget-balancing, and the free-rider problem reappears. In the language of game theory, self-imposed penalties lead to an imperfect equilibrium in the sense of Selten (1975).

The enforcement problem can be overcome only by bringing in a principal (or a party) who will assume the residual of the nonbudget balancing sharing rules. The principal will not renegotiate the contract if for some reason the proper level of output is not attained. Note that it is important that the principal not provide any (unobservable) productive inputs or else a free-rider problem remains.

Of course, the scheme in (8) is only one of many possible solutions to the free-rider problem when (1) is not imposed. Bonding, where each agent pays up front $x(a^*)$ and receives a share $s_i(x) = x$, is another alternative, though it may be infeasible when there are endowment constraints. My point here is therefore not that group punishments are the only effective scheme, but rather that budget breaking is the essential instrument in neutralizing externalities from joint production. The primary role of the principal is to administer incentive schemes that police agents in a credible way rather than to monitor agents as in Alchian and Demsetz' story. The reason capitalistic firms enjoy an advantage over partnerships in controlling incentives is that they can (and will) independently of the level of internal monitoring employ schemes that are infeasible in closed (budget-balancing) organizations. There is little to suggest that either of the two forms of organization would stand at a comparative advantage when it comes to monitoring alone.

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; it does not balance the budget (barring exceptional cases), because by analogy with Theorem 1 balancing the budget would necessarily result in inefficiencies.[3]

□ **Uncertainty.** The reader may have noticed that when agents choose a^*, as they should in equilibrium, there will be no residual left with a scheme like (8). This fact may make the certainty solution appear extreme. That is not true, however. Penalties may work quite effectively under uncertainty as well. This was first observed by Mirrlees (1974) in a model with one agent. The argument is here extended to the multiagent case.

For the moment agents are assumed to be risk neutral.[4] Input costs are as above, output $x(a, \theta)$ is random through the state of nature θ, and agents have homogeneous beliefs concerning θ.

It is more convenient and illuminating to suppress θ and to consider the distribution function of x parameterized by a. (See Holmstrom (1979) for a more detailed argument.) Denote the conditional distribution of x, given the action vector a, by $F(x, a)$ and the conditional density function by $f(x, a)$. Assume that the partial derivatives $F_i(x, a) \equiv \partial F(x, a)/\partial a_i$ and $f_i(x, a) \equiv \partial f(x, a)/\partial a_i$ exist for all i and (x, a). The following assumptions regarding distributions will be used in the theorems below.

Assumption 1. $F(x, a)$ is convex in a.

Assumption 2. $F_i(x, a)/F(x, a) \to -\infty$ as $x \to -\infty$ (or its lower bound).

Assumption 3. $F_i(x, a)/(1 - F(x, a)) \to -\infty$ as $x \to +\infty$ (or its upper bound).

Technically, the role of Assumption 1 will be to assure global optimality of the agents' actions. Unfortunately, it is not an assumption that is satisfied by natural specifications of $x(a, \theta)$, for example, $x(a, \theta) = (\Sigma\, a_i)\theta$ with θ normally distributed. It has an economic interpretation, though. Since $F(x, \lambda a_1 + (1 - \lambda)a_2) \leq \lambda F(x, a_1) + (1 - \lambda)F(x, a_2)$, the distribution on the left dominates the distribution on the right in the sense of first-order

---

[3] The idea of breaking the budget-balancing constraint to resolve externalities has figured more or less explicitly in many articles on property rights. This is particularly true of the bonding solution. Part of the analysis in Green (1976) is closely related to the one presented here, although its context and emphasis are different.

[4] Note that risk neutrality does not here mitigate the moral hazard problem as is the case with a single agent (Harris and Raviv, 1979).

stochastic dominance. Therefore, Assumption 1 corresponds to a particular form of stochastically diminishing returns to scale.

Assumptions 2 and 3 are implied by $f_i/f \rightarrow -\infty$ as $x \rightarrow -\infty$ and $f_i/f \rightarrow +\infty$ as $x \rightarrow +\infty$, respectively, which hold for many natural specifications of $x(a, \theta)$. These are likelihood ratio conditions and can be interpreted as stating that for very small or very large values of x, one can discern very precisely whether the right actions were taken (Milgrom, 1981).

Let output x be shared according to sharing rules $s_i(x)$, $i = 1, \ldots, n$, satisfying (7). For the time being I omit consideration of endowment constraints. The following theorem extends the insights about group incentives from the certainty case to the uncertainty case. The idea of the proof is taken from Mirrlees (1974).

Theorem 3: Under assumptions 1 and 2, a first-best solution can be approximated arbitrarily closely by using group penalties.

Proof: Consider the following sharing rules:

$$s_i(x) = \begin{cases} s_i x, & x \geqslant \bar{x}, \\ s_i x - k_i, & x < \bar{x}, \end{cases} \tag{9}$$

where $k_i > 0$, and $\Sigma\, s_i = 1$. Evidently (9) satisfies (7). The rules in (9) prescribe a penalty k_i to each agent i if a critical output level $\bar{x}$ is not achieved. Otherwise, the entire output is shared. For a^* to be a Nash equilibrium with (9), it is necessary and sufficient (by Assumption 1) that

$$s_i E_i(a^*) - k_i F_i(\bar{x}, a^*) - v_i'(a_i^*) = 0, \qquad i = 1, \ldots, n, \tag{10}$$

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 (10) holds. The expected residual is given by $W = \sum_i k_i F(\bar{x}, a^*)$. I need to show that W can be made arbitrarily small. From (10),

$$k_i = A_i/F_i(\bar{x}, a^*), \qquad \text{with} \qquad A_i = s_i E_i(a^*) - v_i'(a_i^*). \tag{11}$$

Let $\bar{x}$ decrease and adjust k_i so that (11) holds. Then the residual is given by

$$W = \sum_i A_i F(\bar{x}, a^*)/F_i(\bar{x}, a^*), \tag{12}$$

which, by Assumption 2, goes to zero with $\bar{x}$. *Q.E.D.*

From (10) and Assumption 2 it follows that the k_i's generally go to infinity as $\bar{x}$ is decreased. Finite endowments will make this arrangement infeasible. If the x-distribution is tight in the sense that there is an $\bar{x}$-value for which $|F_i|$ is large while $|F/F_i|$ is small, then we can get good approximations to first best even with endowment constraints. But if the distribution has much spread, so that $|F_i|$ is small for all $\bar{x}$ (as is the case for a lognormal distribution when the variance is big), then efficiency losses become substantial with wealth constraints. Assuming that the distribution of output becomes more diffuse when the number of agents grows large (more specifically that $|F_i| \rightarrow 0$ for all x), while $E_i(a)$ stays bounded, implies that a^* will converge to 0 because of (10) and that $s_i \rightarrow 0$ (because $\Sigma\, s_i = 1$). Thus, in contrast to the certainty case, endowment constraints will generally limit the size of a team that can effectively be policed by penalties.

A resolution of this dilemma can occasionally be found by paying bonuses as the following theorem indicates:

Theorem 4: Under Assumptions 1 and 3, the first best can be enforced at a negligible cost to a principal with unbounded wealth, even when agents' endowments are limited.

The proof is omitted because it is similar to that of Theorem 3. The scheme that the principal will use pays a bonus b_i if $x > \bar{x}$ and pays $s_i x(a^*)$, $(\Sigma\, s_i = 1)$, if $x \leq \bar{x}$. The bonuses and $\bar{x}$ can be adjusted simultaneously so that a^* remains an equilibrium, while the principal's expected cost $(\Sigma\ b_i)(1 - F(\bar{x}, a^*))$ goes to zero because of Assumption 3.

The theorems above were developed under risk neutrality assumptions. If agents are risk averse with utility tending to $-\infty$ as wealth decreases, Theorem 3 remains valid (Mirrlees, 1974). Theorem 4, however, appears quite dependent on the risk neutrality assumption. An unverified conjecture is that asymptotic risk neutrality suffices.

3. Sufficient statistics

■ The preceding analysis suggests that under some circumstances efficient team production can be approached via simple penalty or bonus schemes. The role of the principal is important even if there are no risk-sharing advantages, which contrasts with the single-agent case. In the case of penalties the principal is needed to enforce them, and in the case of bonuses he is needed to pay them. The central feature is that the principal allows the budget-balancing constraint to be broken.

If there is uncertainty in production and if agents are risk averse or have limited endowments, monitoring becomes an important instrument in remedying moral hazard, since the first best is not attainable. I shall investigate below what type of monitoring provides valuable information in the sense that it helps improve welfare.

The model includes a risk-neutral principal and n risk-averse agents. The ith agent's utility function is additively separable in money and action, with $u_i(m_i)$ denoting the agent's utility of money function and $v_i(a_i)$ the agent's disutility of action. 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.[5] 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 $g(y, a)$. I assume that the derivative of g with respect to a_i, denoted g_{a_i}, exists for all i. The welfare problem can be stated as

$$\max_{a, s_i(y)} \int \{E(x|y, a) - \sum_i s_i(y)\} dG(y, a), \tag{13}$$

subject to:

$$\text{(i)} \quad \int u_i(s_i(y))dG(y, a) - v_i(a_i) \geqslant \bar{u}_i, \qquad i = 1, \ldots, n.$$

$$\text{(ii)} \quad a_i \in \underset{a_i'}{\operatorname{argmax}} \int u_i(s_i(y))dG(y, a_i', a_{-i}) - v_i(a_i'), \qquad i = 1, \ldots, n.$$

Here, $E(x|y, a)$ is the expected output of x, given y and a. Of course, it equals x if x is part of y. Condition (ii) implies that a is a Nash equilibrium. This behavioral assumption may be unreasonable at times, but the results to be presented do not depend on it in a critical way. (See p. 333.)

The following definition is an extension of that in Holmstrom (1979).

Definition:[6] A function $T_i(y)$ is said to be *sufficient for y with respect to a_i*, if there exist functions $h_i(\cdot) \geqslant 0$, $p_i(\cdot) \geqslant 0$ such that:

---

[5] It is not necessarily enough, as the literature frequently suggests, that y is observable to the principal and the agent for $s(y)$ to be enforceable. Legal enforceability requires that the enforcing authorities are also able to observe y when needed. On the other hand, implicit contracts may include signals that are not observed perfectly by the agents.

[6] Gjesdal (1982) provides a related extension of my earlier sufficient statistic condition.

$$g(y, a) = h_i(y, a_{-i})p_i(T_i(y), a), \qquad \text{for all } y \text{ and } a \text{ in the support of } g. \tag{14}$$

The vector $T(y) = (T_1(y), \ldots, T_n(y))$ is said to be sufficient for y with respect to a, if each $T_i(y)$ is sufficient for a_i.

Equation (14) is the well-known condition for a sufficient statistic in ordinary statistical decision theory (deGroot, 1970). Note, however, that the action a_i is a parameter chosen not by Nature but by a strategic agent. This notwithstanding, it will be shown below (Theorems 5 and 6) that agent i's sharing rule should be based solely on $T_i(y)$ if and only if T_i is sufficient, which parallels results in statistical decision theory.

Theorem 5. Assume $T(y) = (T_1(y), \ldots, T_n(y))$ is sufficient for y with respect to a. Then, given any collection of incentive schemes $\{s_i(y)\}$, there exists a set of schemes $\{\tilde{s}_i(T_i)\}$ that weakly Pareto dominates $\{s_i(y)\}$.

Proof: I consider first the case of a single agent, dropping subscripts throughout. Define $\tilde{s}(T)$ as follows:

$$u(\tilde{s}(T)) = \int_{T(y)=T} u(s(y))g(y, a)dy/p(T, a) = \int_{T(y)=T} u(s(y))h(y)dy. \tag{15}$$

By Jensen's inequality (15) implies:

$$\int \tilde{s}(T(y))g(y, a)dy \leqslant \int s(y)g(y, a)dy. \tag{16}$$

From (14) and (15) it follows that the agent will enjoy the same expected utility for all a, whether faced with $s(y)$ or $\tilde{s}(T)$. Thus, he will take the same action under $\tilde{s}(T)$ as he takes under $s(y)$. From this and (16), it follows that the principal is at least as well off with $\tilde{s}(T)$ as with $s(y)$.

With n agents, the proof is identical once it is observed that the other agents' actions can be viewed as known constants because they can be inferred from the equilibrium. *Q.E.D.*

The converse to Theorem 5 requires brief preparation. I wish to state that if $T(y)$ is not sufficient, then we can strictly improve welfare by observing y. There is a problem, however, with the meaning of an insufficient statistic $T(y)$. Equation (14) 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 vector of optimal sharing rules.[7] Moreover, for a fixed a, equation (14) can always be satisfied by an appropriate choice of $h_i(\cdot)$ and $p_i(\cdot, \cdot)$.

To handle such cases, I shall define $T(y)$ as *sufficient at a* if it is the case that for all i and all T_i,

$$\frac{g_{a_i}(y_1, a)}{g(y_1, a)} = \frac{g_{a_i}(y_2, a)}{g(y_2, a)} \qquad \text{for almost all } y_1, y_2 \in \{y | T_i(y) = T_i\}. \tag{17}$$

Note that (17) follows from (14). Conversely, (17) implies (14) (by integrating) if it holds for all a. I shall say that $T(y)$ is *globally sufficient* if (17) is true *for all a and i* and *globally insufficient* if *for some i* (17) is false *for all a.*

Theorem 6. Assume $T(y)$ is globally insufficient for y. Let $\{s_i(y) = \tilde{s}_i(T(y))\}$ be a collection of nonconstant sharing rules such that the agents' action choices are unique in equilibrium. Then there exist sharing rules $\{\hat{s}_i(y)\}$ that yield a strict Pareto improvement. Moreover,

[7] I am indebted to Steve Ross for this observation.

$\{\hat{s}_i(y)\}$ can be chosen so as to induce the same equilibrium actions as the rules $\{s_i(y)\}$ do.[8]

Proof: Again, for notational simplicity, I consider only the single-agent case. The fact that improvements can be achieved by keeping the agent's action unaltered implies that changes in one agent's sharing rule will not affect the other agents' behavior. Therefore, the steps shown here for a single agent can be repeated for all other agents in turn. The strategy of the proof is to show that if $T(y)$ is globally insufficient, there exists a different set of sharing rules that (i) leave the agent's action unchanged and (ii) leave the principal no worse off and the agent better off in terms of risk sharing.

Since $T(y)$ is globally insufficient, there exist a T_1 and sets of positive measure Y_{11}, Y_{12} which are disjoint and subsets of $Y_1 = \{y|T(y) = T_1\}$ such that

$$\frac{g_a(Y_{11}, a)}{g(Y_{11}, a)} \neq \frac{g_a(Y_{12}, a)}{g(Y_{12}, a)}, \tag{18}$$

where a is the agent's response to $s(y)$. Here $g(Y_{kl}, a) = \Pr\{y \in Y_{kl}|a\}$. 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 of positive measure and $\tilde{s}(T_1) \neq \tilde{s}(T_2)$. Define the following variation:

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

where $I_{11}(y)$ is the indicator function for the event $\{y \in Y_{11}\}$ and similarly for $I_{12}(y)$, $I_2(y)$; and ds_{11}, ds_{12}, and ds_2 are numbers that we shall choose.

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

$$\Delta P = -[ds_{11}g(Y_{11}, a) + ds_{12}g(Y_{12}, a) + ds_2g(Y_2, a)], \tag{19}$$

$$\Delta A = u'_1[ds_{11}g(Y_{11}, a) + ds_{12}g(Y_{12}, a)] + u'_2ds_2g(Y_2, a). \tag{20}$$

Here $u'_1 = u'(\tilde{s}(T_1))$ and $u'_2 = u'(\tilde{s}(T_2))$. Since $\tilde{s}(T_1) \neq \tilde{s}(T_2)$, $u'_1 \neq u'_2$. Assume for concreteness that $u'_2 > u'_1$.

The sign of the effect on the agent's action from the variation $\hat{s}(y)$ is given by

$$\text{sgn}(\Delta a) = \text{sgn}[u'_1(ds_{11}g_a(Y_{11}, a) + ds_{12}g_a(Y_{12}, a)) + u'_2ds_2g_a(Y_2, a)]. \tag{21}$$

Choose ds_{11}, ds_{12}, ds_2 as follows. Fix $ds_2 > 0$ and require that $\Delta P = 0$ and $\Delta a = 0$ so that

$$ds_{11}g(Y_{11}, a) + ds_{12}g(Y_{12}, a) = -ds_2g(Y_2, a), \text{ and} \tag{22}$$

$$ds_{11}g_a(Y_{11}, a) + ds_{12}g_a(Y_{12}, a) = -u'_2/u'_1ds_2g_a(Y_2, a). \tag{23}$$

From (20) and (22), it follows that $\Delta A > 0$, because $u'_2 > u'_1$. The system (22) and (23) has a solution because (18) implies that the determinant is nonzero. 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.*

The main import of Theorem 5 is that randomization does not pay if the agent's utility function is separable.[9] Indeed, any pure noise should be filtered away from the

---

[8] The assumption that sharing rules are nonconstant is not an essential restriction; a constant sharing rule does not provide any incentives for effort and would obviously not be employed.

The assumption that actions are unique is restrictive. For conditions that guarantee uniqueness at the optimum, see Grossman and Hart (1980).

[9] Separability of the utility function is crucial for the result that pure randomization does not pay; see Gjesdal (1982). On the other hand, Karl Shell alerted me to the fact that even if the utility function is not separable, $T(y)$ is all we need to know about y, since randomization, when desired, can be generated in other ways than through the noninformative part of y.

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.

Conversely, Theorem 6 states that if $T(y)$ is not a sufficient statistic for y (at the optimal a), 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.

Theorems 5 and 6 differ somewhat from my earlier result on informativeness, which dealt with a comparison of the two information systems $\eta_1 = x$ and $\eta_2 = (x, z)$. In η_1, the outcome x is observed, while in η_2 both the outcome x and an additional signal z are observed. In Holmstrom (1979), I called z an informative signal if $T(x, z) = x$ is not sufficient for (x, z) as defined in (14), and I proved that z is valuable if and only if z is informative.[10] Obviously, Theorems 5 and 6 generalize this result.

More importantly, though, for the application to the multiagent case, Theorem 6 shows that if $T(y)$ is not sufficient, welfare improvements can be made without changing the agents' actions, but merely by improving risk sharing. The proofs in Holmstrom (1979) and Shavell (1979) both do the reverse: maintain the same level of risk sharing, while providing incentives for the agent to change the action in a direction desirable for the principal. Such changes in action would complicate considerably the situation with many agents, since the whole equilibrium would then change.[11]

Working with risk sharing improvements, while keeping actions intact, has another advantage: the results do not hinge critically on the Nash equilibrium assumption made in the earlier formulation of the problem (see (13)). Theorem 5 is valid no matter what behavior is assumed; in particular, cooperation among agents would be acceptable. Theorem 6 is not so insensitive, because in the proof the agents' payoffs (as a function of actions) are changed globally, and this may induce collusive behavior. But as long as actions are chosen noncooperatively, the specific equilibrium concept is inessential, even in Theorem 6.

The sufficient-statistic results above are closely related to Blackwell's well-known theorem in decision theory. Blackwell's theorem states that all decisionmakers in all decision situations will prefer one information system (that is, experiment) to another if and only if the former is sufficient for the latter. (I shall explain shortly what this means.) But, Theorems 5 and 6 are not corollaries of Blackwell's results. Most importantly, the agency problem does not fit directly the framework of statistical decision theory. No inferences are being drawn from y concerning a, because actions are determined by the choice of the incentive scheme. In fact, the outcome y does not tell the principal anything at all about the actions of the agents! It is quite appealing, therefore, that the agency results conform with the intuition that information is being extracted from y, although logically that is not what is occurring.[12]

Secondly, Theorem 6 does not have a counterpart in Blackwell's analysis. Whether the necessary part of Blackwell's theorem, that information systems cannot be compared unless one is sufficient for the other, is true in the agency framework is still an open question.

Grossman and Hart (1980) and Gjesdal (1982) discuss the value of different infor-

---

[10] The sufficiency part was also proved by Shavell (1979) without explicitly using the notion of a sufficient statistic.

[11] Baiman and Demski (1980) prove a special case of Theorem 6 by 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).

[12] Mathematically, the agency problem can be set up as a statistical decision problem with the restriction that the "decision rule," $s(y)$, has to satisfy the agents' incentive constraints. These are nonconvex in general, which explains why randomization may be valuable (see Gjesdal (1982) for an example). Randomization is never worthwhile in Blackwell's set-up.

mation systems from the point of view of Blackwell's notion of sufficiency. The following discussion will reveal its relationship to the sufficiency property used here.[13]

Let $\eta_1 = y_1$ and $\eta_2 = y_2$ be two information systems. Blackwell calls η_1 sufficient for η_2 if we have for all y_2, a:

$$g_2(y_2, a) = \int h(y_2, y_1) g_1(y_1, a) dy_1 , \tag{24}$$

where g_2 and g_1 are the marginal densities of y_2 and y_1. Condition (24) looks much like (14), but is conceptually quite different, since (24) involves only marginal distributions, whereas (14) presumes a joint distribution for $y = (y_1, y_2)$. But the solutions to the Pareto problem (13) corresponding to η_1 or η_2 do not depend on the joint distribution of y_1 and y_2. Hence we may define a joint density by letting

$$g(y_1, y_2, a) = h(y_2, y_1) g_1(y_1, a). \tag{25}$$

This will have as its marginals g_1 and g_2 (as in (24)). But (25) 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 we do by basing it on (y_1, y_2). Therefore, $\eta_1 = y_1$ is as good an information system as $\eta = (y_1, y_2)$. But, of course, η can be no worse than $\eta_2 = y_2$. The conclusion is that η_1 is as good as η_2 if η_1 is sufficient for η_2 in the Blackwell sense. The Blackwell ordering can therefore be viewed as a corollary to the results of Theorems 5 and 6. The converse is true as well, since (24) is weaker than (14).

4. Relative performance evaluation and competition among agents

■ The result on sufficient statistics has many useful applications in agency theory. I shall illustrate its applicability by analyzing relative performance evaluation, which is of considerable relevance in the multiagent setting.

Theorem 6 lends support to the effort firms make in creating information systems that separate out individual contributions to total output. It is easy to show (by using (18)) that two independent measures of joint output are together more informative than a single measure of output. Thus, refined budgetary measures have positive value when costs are excluded.

Now consider the case where the information system is so rich that total output can be itemized according to the contribution of each individual; that is,

$$x(a, \theta) = \sum_i x_i(a_i, \theta_i), \qquad \theta = (\theta_1, \ldots, \theta_n),$$

where all x_i's are separately observed. If the θ_i's are not random, then efficiency can be achieved by holding each agent responsible for his own output.[14] This is in accordance with the general principles of responsibility accounting. What will be examined here is when under uncertainty it will be valuable to deviate 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 on x_i alone.

We frequently observe agents being evaluated on the basis of peer performance. In almost 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 is measured by rank alone, has been analyzed by Lazear and

---

[13] I am grateful to Paul Milgrom for discussing this relationship.

[14] In an earlier version of this article, I showed a partial converse: to achieve efficiency under certainty by using monitoring alone, n measures are needed that locally separate the contributions of individuals as in (26); this was shown under the restriction that sharing rules and measures are monotonic functions.

Rosen (1981). Related examples are provided by recent executive incentive packages in which 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 given in the previous section. The following theorem shows how those results apply.

Theorem 7:[15] Assume the x_i's are monotone in θ_i. Then the optimal sharing rule of agent i depends on individual i's output alone if and only if outputs are independent.

Proof: If the θ_i's are independent, then

$$f(x, a) = \prod_{i=1}^{n} f_i(x_i, a_i), \tag{26}$$

which obviously implies that $T_i(x) = x_i$ is sufficient for x with respect to a_i. By Theorem 5, it will be optimal to let s_i depend on x_i alone.

Suppose instead that θ_1 and θ_2 are dependent. Since, in equilibrium, the value of a_2 can be inferred, we can, without loss of generality, assume $x_2 = \theta_2$. Keeping a_2 at its equilibrium value (and suppressing, it notationally), the joint distribution of x_1 and $x_2 = \theta_2$ conditional on a_1 is given by

$$f(x_1, \theta_2, a_1) = \tilde{f}(x_1^{-1}(a_1, x_1), \theta_2), \tag{27}$$

where x_1^{-1} is the inverse of $x_1(a_1, \theta_1)$ and $\tilde{f}(\theta_1, \theta_2)$ is the joint distribution of (θ_1, θ_2). It follows that

$$\frac{f_{a_1}(x_1, \theta_2, a_1)}{f(x_1, \theta_2, a_1)} = \frac{\tilde{f}_1(x_1^{-1}(a_1, x_1), \theta_2)}{\tilde{f}(x_1^{-1}(a_1, x_1), \theta_2)} \frac{\partial x_1^{-1}(a_1, x_1)}{\partial a_1}. \tag{28}$$

Since θ_1 and θ_2 are dependent, $\tilde{f}_1/\tilde{f}$ depends on θ_2. Thus, (17) does not hold, and Theorem 6 applies. Consequently, the sharing rule for agent 1 should depend on both x_1 and x_2. *Q.E.D.*

An important implication of Theorem 7 is that forcing agents to compete with each other is valueless if there is no common underlying uncertainty. In this setting, the benefits from competition itself are nil. What is of value is the information that may be gained from peer performance.[16] Competition among agents is a consequence of attempts to exploit this information.

At this point it is appropriate to comment on the use of rank-order tournaments (Lazear and Rosen, 1981). 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, $w_1 \geqslant \ldots \geqslant w_n$. The agent with the highest output gets w_1, second highest gets w_2, and so on.

From Theorem 7 it follows that if the agents' outcomes are unrelated, then rank-order tournaments will perform worse than rewarding agents on the basis of their individual outcomes alone. Pitting agents against each other will only result in more randomness in the reward scheme without any gains in the power of inference about actions.[17] On the other hand, as first noted by Lazear and Rosen, rank-order tournaments may be valuable if outcomes are related. The analysis above supports this contention. But it should be observed that rank-order tournaments may be informationally quite wasteful

[15] A related result is proved in Baiman and Demski (1980).

[16] In the extreme, if agents' outputs are completely dependent, in the sense that any x_j will reveal all θ_i, $i \neq j$, then the first-best outcome can be easily achieved by using relative performance evaluations.

[17] Lazear and Rosen (1981) find that rank-order tournaments may dominate piece rates even if agents' outputs are independent. Theorem 7 shows that this is the case not because of the value of competition, but because piece rates are sometimes far from optimal in the single-agent case.

if performance levels can be measured cardinally rather than ordinally. It is clear that the mapping from the agents' outcomes $x = (x_1, \ldots, x_n)$ into the statistic $T(x) = (k_1(x), \ldots, k_n(x))$, where $k_i(x)$ is the rank order of agent i, is not a sufficient statistic for a, except in trivial cases. Therefore, Theorem 6 tells us that there should be a better way of making use of x than what the rank-order tournament does. (Of course, if the output is very complex so that only ordinal measures are possible, then rank-order tournaments will be the best one can do.)

A general characterization of how information about the optimal use of peer performance with many agents can be developed formally as an extension of the characterization in Mirrlees (1976) or Holmstrom (1979). I shall not pursue that issue here. Instead I shall indicate how the sufficient statistic condition can be used to rationalize schemes that only use aggregate information about peer performance.

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

I: $$x_i(a_i, \theta_i) = a_i + \eta + \epsilon_i, \qquad i = 1, \ldots, n,$$

II: $$x_i(a_i, \theta_i) = a_i(\eta + \epsilon_i), \qquad i = 1, \ldots, n.$$

Here $\theta_i = (\eta, \epsilon_i)$, where η is a common uncertainty parameter, while the ϵ_i's are independent, idiosyncratic risks.

Theorem 8: Let the technology be given by either I or II. Assume that $\eta, \epsilon_1, \ldots, \epsilon_n$ are independent and normally distributed. Let $\bar{x} = \sum \alpha_i x_i$ be a weighted average of the agents' outcomes. In the case of technology I, let $\alpha_i = \tau_i/\bar{\tau}$, where τ_i is the precision (the inverse of the variance) of ϵ_i and $\bar{\tau} = \sum \tau_i$. In the case of technology II, let $\alpha_i = \tau_i/\bar{\tau}a_i$, where a_i is the equilibrium response of agent i. In both cases, an optimal set of sharing rules $\{s_i(x)\}$ will have s_i depend on $\bar{x}$ and x_i alone.

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

$$f(x, a) = K \int \exp\{-\tfrac{1}{2}[\sum_j \tau_j(x_j - a_j - \mu_j - \eta)^2 + \tau_0(\eta - \mu_0)^2]\}d\eta, \tag{29}$$

where K is a constant, τ_0 is the precision of η, μ_0 is the mean of η, and τ_j, μ_j are the precision and mean of ϵ_j. In view of Theorem 5, we need to show that we can write (29) in the form of (14) for each $i = 1, \ldots, n$. Let

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

Then we can write:

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

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

Substituting this expression into (29), we find upon integrating over η that we can write $f(x, a) = h_i(x, a_{-i})\hat{p}_i(\bar{z}_{-i}, x_i, a)$. But, since

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

we have $\hat{p}_i(\bar{z}_{-i}, x_i, a) = p_i(\bar{x}, x_i, a)$, which completes the proof of writing (29) in the form (14).

The proof for technology II is similar and is omitted. *Q.E.D.*

Theorem 8 suggests that sometimes 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, although, of course, the sufficiency of a weighted average is a specific feature of the normal distribution.

Notice that Theorem 8 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 scale and in the value of these information sources. Differences in scale are corrected for by dividing x_j by a_j, which can be interpreted as a rate of return measure. Information values differ if ϵ_j's have different precision. If ϵ_j has high precision (low variance), then x_j tells rather sharply the value of η and should receive more weight 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 $\tau_j \rightarrow 0$, x_j will essentially tell nothing about η because of the noise in ϵ_j, and hence should count very little.

The predictions of Theorem 8 conform well with the recent trend in executive incentive design. After stock options lost their tax advantage (and perhaps also because the market had been depressed in general), performance incentive packages became popular. These tie executive compensation to performance measured explicitly in relation to other firms in the industry (in particular, industry averages).

Let me turn finally 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 (by Theorem 5) be filtered away to yield an improved solution to the agency problem. Also, if we knew η *ex post*, there would be no need to compare individual agents' outputs, since conditional on η they are independent (cf. Theorem 7). 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, the optimal schemes will depend on $a_i + \epsilon_i$ (for I) and $a_i\epsilon_i$ (for II), since the observation of η will allow us to observe these variables.

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

Theorem 9: Consider technology I or II. Assume $\eta, \epsilon_1, \ldots, \epsilon_n$ are independent with uniformly bounded variance. Assume that in the solution to the single-agent problem without common uncertainty (i.e., $\eta = 0$) the agent's response is unique. Then this solution can be approximated arbitrarily closely as the number of agents grows large.

Proof: Consider technology I. Define $q_j = \eta + \epsilon_j$, $j = 1, \ldots, n$, and $\bar{q}_{-i} = 1/(n-1) \sum_{j \neq i} q_j$, $i = 1, \ldots, n$. 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 almost surely to η. Therefore,

$$\int u_i(s_i^*(a_i + \eta + \epsilon_i - \bar{q}_{-i}))dP(\eta, \epsilon_1, \ldots, \epsilon_n)$$

converges uniformly to

$$\int u_i(s_i^*(a_i + \epsilon_i))dP(\epsilon_i).$$

Since a_i^* is a unique solution to

$$\max_{a_i} \int u_i(s_i^*(a_i + \epsilon_i))dP(\epsilon_i) - v_i(a_i),$$

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

when solving

$$\max_{a_i} \int u_i(s_i^*(a_i^* + \eta + \epsilon_i - \bar{q}_{-i}))dP(\eta, \epsilon_1, \ldots, \epsilon_n) - v_i(a_i).$$

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

It is clear that constraining attention to technology I or II is not essential. The important thing is that all common uncertainty can be discerned in the limit and therefore removed from the agent's responsibility. The specifications above were just simple ways to illustrate the point.

The result that one may use relative performance measures to filter away common uncertainties has some interesting implications for financial theory. A standard model of financial markets is the capital asset pricing model (CAPM). 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 be made solely with reference to systematic risk.

The analysis above shows that when incentives are a relevant concern, the normative implications of CAPM are altered. A new cost component related to moral hazard must be added. The interesting fact is that the costs due to moral hazard depend only on the idiosyncratic risk, which is precisely the converse of the costs due to risk bearing in the market. Therefore, the total costs of investment risk can be separated into two components: a price for the systematic risk given by the market according to CAPM (or more generally by arbitrage pricing) and a price for idiosyncratic risk determined by the project-specific returns to effort and the risk preferences of the manager.[18]

The agency perspective implies that between two projects with the same level of systematic risk, the one with lower idiosyncratic risk is strictly preferred. Consequently, diversification within the firm can be useful, since it helps to measure more accurately the manager's input. Furthermore, the concern for more accurate performance evaluation tends to make the firm choose projects which are more correlated with the market portfolio than efficient risk taking would prescribe. This implies that in their efforts to reduce idiosyncratic risk, firms provide society with a market portfolio which is not so diversified as (and hence riskier than) it would be were there no incentive problems present.[19]

This last point has a counterpart in the internal organization of a firm. Specialization is good from a purely technological point of view. In this respect agents' tasks should be well diversified. But problems with performance evaluation place a limit on the value of specialization. Thus, it may be optimal to have agents' tasks overlap (or duplicate) each other. One manifestation of this principle comes in the form of job rotation, which provides independent readings of the circumstances in which tasks are being carried out and thereby reduces moral hazard costs.

5. Concluding remarks

■ This article has made two general points. One is that the free-rider problem, which may arise in a multiagent setting, can largely be resolved if ownership and labor are partly

[18] Note, however, that including agency costs does not require a change in the positive theories of asset pricing. Prices of assets are based on expected returns in which agents' actions have already been factored. Whether expected returns are high because of the state variable in the outcome function $x(a, \theta)$ or because of the agent's action is immaterial for pricing the asset.

[19] One should not interpret insufficient diversification due to moral hazard as an inefficiency, since the framework is second best. There is a source of true inefficiency, though, when investment decisions are made independently. A firm which moves its project portfolio closer to the market's for improved performance evaluation does not account for the benefits that accrue to other firms from such a move.

separated. This gives capitalistic firms an advantage over partnerships. The other point is that relative performance evaluation can be helpful in reducing moral hazard costs, because it provides for better risk sharing. It is worth noting that many of the insights in the literature on rank-order tournaments and the value of competition (Lazear and Rosen, 1981; Green and Stokey, 1982; Nalebuff and Stiglitz, 1982) are fundamentally derivatives of the sufficient-statistic condition and therefore of much broader scope as shown here. In this respect generality proved both tractable and rich in implications.

There are other factors of the multiagent problem that have not been addressed in this article, but are worth studying. One concerns the possibility of collusion among agents when relative performance evaluations are used. Collusion may imply restrictions on reward structures. In this regard rank-order tournaments, which induce a zero-sum game between the agents, seem to have an advantage over schemes which are not zero-sum.

Another important issue relates to monitoring hierarchies. In this article monitoring technologies were exogenously given. In reality, they are not. The question is what determines the choice of monitors; and how should output be shared so as to provide all members of the organization (including monitors) with the best incentives to perform? Satisfactory answers to these questions would take us a big step toward understanding nonmarket organization.

Appendix

Proof of Theorem 1

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

From the definition of a Nash equilibrium

$$s_i(x(a_i, a^*_{-i})) - v_i(a_i) \leqslant s_i(x(a^*)) - v_i(a^*_i), \ \forall \ a_i \in A_i \, . \tag{A1}$$

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

$$\alpha^l = x(a^l_i, a^*_{-i}). \tag{A2}$$

The sequences $\{a^l_i\}$ are well defined (starting from a large enough l if necessary), since $a^* \in \text{int } A$, $x'_i(a^*) \neq 0$, and $x(a)$ is strictly concave. Pareto optimality implies $v'_i(a^*_i) = x'_i(a^*)$, $\forall \ i$. This in turn implies, using (A2), that $v_i(a^*_i) - v_i(a^l_i) = x(a^*) - x(a^l_i, a^*_{-i}) + o(a^l_i - a^*_i)$, $\forall \ i$, $\forall \ l$, where $o(h)/h \to 0$ as $h \to 0$. Substituting into (A1), using (A2) gives

$$x(a^*) - \alpha^l + o(a^l_i - a^*_i) \leqslant s_i(x(a^*)) - s_i(\alpha^l), \ \forall \ i, \ \forall \ l. \tag{A3}$$

Sum (A3) over i, use (1), rearrange and multiply by $n/(n-1)$. This gives:

$$\sum_{i=1}^{n} \{x(a^*) - \alpha^l + o(a^l_i - a^*_i)\} \leqslant 0, \ \forall \ l,$$

which can be written

$$\sum_{i=1}^{n} \{-x'_i(a^*)(a^l_i - a^*_i) + o(a^l_i - a^*_i)\} \leqslant 0, \ \forall \ l. \tag{A4}$$

Since $\alpha^l < x(a^*)$ 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 (A4). Hence, the assumption that a^* is a Nash equilibrium has led to a contradiction and must be false. *Q.E.D.*

References

ALCHIAN, A. AND DEMSETZ, H. "Production, Information Costs, and Economic Organization." *American Economic Review*, Vol. 62 (1972).

ARROW, K. *Limits to Organization.* New York: Norton & Co., 1974.

ATKINSON, A.A. AND FELTHAM, G.A. "Information in Capital Markets: An Agency Theory Perspective." Mimeographed, University of British Columbia, July 1980.

BAIMAN, S. AND DEMSKI, L. "Economically Optimal Performance Evaluation and Control Systems." *Journal of Accounting Research*, Supplement (1980).

DEGROOT, M. *Optimal Statistical Decisions.* New York: McGraw-Hill, 1970.

GJESDAL, F. "Information and Incentives: The Agency Incentive Problem." *Review of Económic Studies*, Vol. 49(3), No. 157 (July 1982).

GREEN, J. "On Optimal Structure of Liability Laws." *Bell Journal of Economics*, Vol. 7, No. 2 (Autumn 1976).

——— AND STOKEY, N. "A Comparison of Tournaments and Contracts." Working Paper No. 840, NBER, January 1982.

GROSSMAN, S. AND HART, O. "An Analysis of the Principal-Agent Problem." CARESS WP No. 80-17, University of Pennsylvania, April 1980 (forthcoming in *Econometrica*).

HARRIS, M. AND RAVIV, A. "Optimal Incentive Contracts with Imperfect Information." *Journal of Economic Theory*, Vol. 24 (1979).

HOLMSTROM, B. "Moral Hazard and Observability." *Bell Journal of Economics*, Vol. 10, No. 1 (Spring 1979).

———. "A Discussion of Economically Optimal Performance Evaluation and Control Systems." *Journal of Accounting Research*, Supplement (1980).

LAZEAR, E. AND ROSEN, S. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy*, Vol. 89, No. 5 (1981).

MARSCHAK, J. AND RADNER, R. *Economic Theory of Teams.* New Haven: Yale University Press, 1972.

MILGROM, P. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics*, Vol. 12, No. 2 (Autumn 1981).

MIRRLEES, J. "Notes on Welfare Economics, Information and Uncertainty" in Balch, McFadden, and Wu, eds., *Essays on Economic Behavior under Uncertainty*, Amsterdam: North Holland, 1974.

———. "The Optimal Structure of Incentives and Authority within an Organization." *Bell Journal of Economics*, Vol. 7, No. 1 (Spring 1976).

MOSSIN, J. "Security Pricing and Investment Criteria in Competitive Markets." *American Economic Review*, Vol. 59 (1969).

NALEBUFF, B. AND STIGLITZ, J. "Prizes and Incentives: Toward a General Theory of Compensation and Competition." Research Memorandum No. 293, Princeton University, January 1982.

RADNER, R. "Optimal Equilibria in Some Repeated Games with Imperfect Monitoring." Bell Laboratories Discussion Paper, May 1980.

ROSS, S. "The Economic Theory of Agency: The Principal's Problem." *American Economic Review*, Vol. 63 (May 1973).

SELTEN, R. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games." *International Journal of Game Theory*, Vol. 4 (1975).

SHAVELL, S. "Risk Sharing and Incentives in the Principal and Agent Relationship." *Bell Journal of Economics*, Vol. 10, No. 1 (Spring 1979).

SPENCE, A.M. AND ZECKHAUSER, R. "Insurance, Information and Individual Action." *American Economic Review*, Vol. 61 (May 1971).

WILSON, R. "The Theory of Syndicates." *Econometrica*, Vol. 36 (January 1968).

———. "The Structure of Incentives for Decentralization under Uncertainty." *La Decision*, No. 171 (1969).