AN ANALYSIS OF THE PRINCIPAL-AGENT PROBLEM

by

Sanford J. Grossman* and Oliver D. Hart**

Working Paper No. 15-80

RODNEY L. WHITE CENTER
FOR FINANCIAL RESEARCH

The Wharton School

University of Pennsylvania

Philadelphia, PA 19104

*University of Pennsylvania.

**Churchill College, Cambridge.

Support from the U.K. Social Science Research Council and NSF Grant No. SOC 79-13429 is gratefully acknowledged. We would like to thank Bengt Holmstrom and Jim Mirrlees for helpful comments on an earlier draft.

Abstract

Most analyses of the principal-agent problem assume that the principal chooses an incentive scheme to maximize expected utility subject to the agent's utility being at a stationary point. An important paper of Mirrlees has shown that this approach is generally invalid. We present an alternative procedure. If the agent's utility function is separable in action and reward, we show that the optimal way of implementing an action by the agent can be found by solving a convex programming problem. We use this to characterize the optimal incentive scheme and to analyze the determinants of the seriousness of an incentive problem.

An Analysis of the Principal-Agent Problem

by

Sanford J. Grossman* and Oliver D. Hart**

April 1980

[Revised July 1980]

1. Introduction

It has been recognized for some time that, in the presence of moral hazard, market allocations under uncertainty will not be unconstrained Pareto optimal (see Arrow [1971], Pauly [1968]). It is only relatively recently, however, that economists have begun to undertake a systematic analysis of the properties of the second-best allocations which will arise under these conditions. Much of this analysis has been concerned with what has become known as the principal-agent problem. Consider two individuals who operate in an uncertain environment and for whom risk sharing is desirable. Suppose that one of the individuals (known as the agent) is to take an action which the other individual (known as the principal) cannot observe. Assume that this action affects the total amount of consumption or money which is available to be divided between the two individuals. In general, the action which is optimal for the agent will depend on the extent of risk sharing between the principal and the agent. The question is: What is the optimal degree of risk sharing, given this dependence?

---

* University of Pennsylvania.

** Churchill College, Cambridge.

Support from the U.K. Social Science Research Council and NSF Grant No. SOC 79-13429 is gratefully acknowledged. We would like to thank Bengt Holmstrom and Jim Mirrlees for helpful comments on an earlier draft.

Particular applications of the principal-agent problem have been made to the case of an insurer who cannot observe the level of care taken by the person being insured; to the case of an insurer who cannot observe the extent of the loss incurred by the person being insured; and to the case of an owner of a firm who cannot observe the effort level of a manager or worker.[1/]

Most of the mathematical analyses of the principal-agent problem proceed by assuming that the principal chooses the risk-sharing contract, or incentive sheme, to maximize his expected utility subject to the constraints that (a) the agent's expected utility is no lower than some pre-specified level; (b) the agent's utility is at a stationary point; i.e., the agent satisfies his first-order conditions with respect to the choice of action. However, in an important paper, Mirrlees [1975] has shown that this procedure is generally invalid unless, at the optimum, the solution to the agent's maximum problem is unique. In the absence of uniqueness (and it is difficult to guarantee uniqueness in advance), the first-order conditions derived by the above procedure are not even necessary conditions for the optimality of the risk-sharing contract.

The purpose of this paper is to develop a method for analyzing the principal agent problem which avoids the difficulties of the "first-order condition" approach. Our approach is to break the principal's problem up into a computation of the costs and benefits of the different actions taken by the agent. For each action, we consider the incentive scheme which minimizes the (expected) cost of getting the agent to choose that action. We show that, under the assumption that the agent's utility function is additively or multiplicatively separable in action and reward, this cost minimization problem is a fairly straightforward convex programming problem. An analysis of these convex problems yields a number of results about the form of the optimal incentive scheme. We will also be able to analyze what factors determine how serious a particular incentive problem is; i.e., how great the loss is to the principal from having to operate in a second-best situation where the agent's action cannot be observed relative to a first-best situation where it can be observed.

In addition to providing greater rigour, the costs vs. benefits approach also provides a clear separation of the two distinct roles the agent's output plays in the principal-agent problem. On the one hand, the agent's output contributes positively to the principal's consumption, so the principal desires a high output. On the other hand, the agent's output is a signal to the principal about the agent's level of effort. This informational role may be in conflict with the consumption role. For example, there may be a moderate output level which is achieved when the agent takes low effort levels and never occurs at other effort levels. If the agent is penalized whenever this moderate output occurs, then he is discouraged from taking these low effort actions. However, there may be lower output levels which have some chance of occurring regardless of the agent's action. To encourage the agent to take high effort levels, it is then optimal to pay the agent more in low output states than in moderate output states, even though the principal prefers moderate output levels to low output levels.

The dual role of output makes it difficult to obtain conditions which ensure even elementary properties of the incentive scheme, such as monotonicity. In Section 3, sufficient conditions for monotonicity are given. It is also shown in this section that a monotone likelihood ratio condition, which the "first-order condition" approach suggests is a guarantee of monotonicity, must be strengthened once we take into account the possibility that the agent's action is not unique at the optimal incentive scheme.

The paper is organized as follows. In Section 2, we show how the principal's optimization problem can be decomposed into a costs vs. benefits problem. In Section 3, we use our approach to analyze the monotonicity and progressivity of the optimal incentive scheme. In Section 4, we give a simple algorithm for computing an optimal incentive scheme when there are only two outcomes associated with the agent's actions. Finally, Section 5 analyzes the effects of risk aversion and information quality on the incentive problem.

2. Statement of the Problem

The application of the principal-agent problem that we will consider is to the case of the owner of a firm who delegates the running of the firm to a manager. The owner is the principal and the manager the agent. The owner is assumed not to be able to monitor the manager's actions. The owner does, however, observe the outcome of these actions, which we will take to be the firm's profit. It is assumed that the firm's profit depends on the manager's actions, but also on other factors which are outside the manager's control -- we model these as a random component. Thus, in particular, if the firm does well, it will not generally be clear to the owner whether this is because the manager has worked well or whether it is because he has been lucky.[2/]

We will simplify matters by assuming that there are only finitely many possible gross profit levels for the firm, denoted $q_1,\ldots,q_n$, where $q_1 < q_2 < \ldots < q_n$. We will assume that the principal is interested only in the firm's net profit, i.e. gross profit minus the payment to the manager. We will also assume that the principal is risk neutral -- our methods of analysis can, however, be applied to the case where the principal is risk averse (see Remark 2).

Let A be the set of actions available to the manager. We will assume that A is a compact subset of a finite dimensional Euclidean space. Let $S = \{x \in R^n |\ x \geq 0,\ \sum_{i=1}^{n} x_i = 1\}$. We assume that there is a continuous function $\pi\colon\ A \to S$, where $\pi(a) = (\pi_1(a),\ldots,\pi_n(a))$ gives the probabilities of the n outcomes $q_1,\ldots,q_n$ if action a is selected. It is assumed that, when the agent chooses $a \in A$, he knows the probability function π but not the outcome which will result from his action. We assume that the agent has a von Neumann-Morgenstern utility function $U(a,I)$ which depends both on his action a and his remuneration I from the principal. We include a as an argument in order to capture the idea that the agent dislikes working hard, taking care, etc.

The crucial simplifying assumption that we will make is that $U(a,I)$ is additively or multiplicatively separable in a and I.

(A1) (Additive Separability) $U(a,I)$ can be written as $V(I) - G(a)$, where (1) V is a real-valued, continuous, strictly increasing, concave function defined on some interval $\mathcal{I} = (\underline{I},\infty)$ of the real line; (2) $\lim_{I \to \underline{I}} V(I) = -\infty$; (3) G is a real-valued, continuous function defined on A.

In the above, we allow for the case that $\underline{I} = -\infty$.

(A1') (Multiplicative Separability) $U(a,I)$ can be written as $V(I)G(a)$, where (A) V satisfies (1) and (2) above; (B) G is a strictly positive, real-valued, continuous function defined on A.

An interesting special case of multiplicative separability is when $V(I) = -e^{-kI}$, $G(a) = e^{ka}$ and A is a subset of the real line. Then $U(a,I) = -e^{-k(I-a)}$; i.e., effort appears just as negative income.

In the "first-best" situation where the principal can observe a, it is optimal for him to pay the agent according to the action he chooses. Let $\bar{U}$ be the agent's reservation price, i.e. the expected level of utility he can achieve by working elsewhere. We assume

(A2) If (A1) (resp. (A1')) holds, $\bar{U} + G(a)$ (resp. $\bar{U}/G(a)) \in V(\mathcal{I})$ for all $a \in A$.

Definition. Let $C_{FB}: A \to R$ be defined by $C_{FB}(a) = h(\bar{U} + G(a))$ in the case of additive separability and by $C_{FB}(a) = h(\bar{U}/G(a))$ in the case of multiplicative separability, where $h \equiv V^{-1}$. (Here C_{FB} stands for first-best cost.)

Then to get the agent to pick $a \in A$ in the first-best situation, the principal will offer him the following contract: I will pay you $C_{FB}(a)$ if you choose a and $\tilde{I}$ otherwise, where $\tilde{I}$ is close to $\underline{I}$.

Definition: Let $B: A \to R$ be defined by $B(a) = \sum_{i=1}^{n} \pi_i(a)q_i$. $B(a)$ is the expected benefit to the principal from getting the agent to pick a.

Definition: A first-best optimal action is one which maximizes $B(a) - C_{FB}(a)$ on A.

The function C_{FB} induces a complete ordering on A. For obvious reasons we will refer to actions with higher $C_{FB}(a)$'s as costlier actions.

In the second-best situation where a is not observed by the principal, it is not possible to make the agent's remuneration depend on a. Instead, the principal will pay the agent according to the outcome of his action, i.e. according to the firm's profit. An incentive scheme is therefore an n-dimensional vector $\underset{\sim}{I} = (I_1, I_2, \ldots, I_n) \in \mathcal{I}^n$, where I_i is the agent's remuneration in the event that the firm's profit is q_i. Given the incentive scheme $\underset{\sim}{I}$, the agent will choose $a \in A$ to maximize $\sum_{i=1}^{n} \pi_i(a)U(a,I_i)$.

We will assume that the principal knows the agent's utility function $U(a,I)$, the set A and the function $\pi: A \to S$. In other words, the principal is fully informed about the agent and about the firm's production possibilities. The incentive problem which we will study therefore arises entirely because the principal cannot monitor the agent's actions.[3/]

The principal's problem is to choose an incentive scheme $\underset{\sim}{I}^*$ and an action

a^* so that (1) under $\underset{\sim}{I}^*$, the agent will be willing to work for the principal and will find it optimal to choose a^* and (2) $\sum_i \pi_i(a)(q_i - I_i)$ is maximized at $a^*, \underset{\sim}{I}^*$. It simplifies matter considerably if we break this problem up into two parts. We consider first, given that the principal wishes to implement a^*, the least cost way of achieving this. We then consider which a^* should be implemented. Thus, to begin, suppose that the principal wishes the agent to pick a particular action $a^* \varepsilon A$. To find the least (expected) cost way of achieving this, the principal must solve the following problem:

(2.1) Choose $I_1, \ldots, I_n$ to minimize $\sum_{i=1}^{n} \pi_i(a^*)I_i$

$$\text{S.T.} \quad \sum_{i=1}^{n} \pi_i(a^*)U(a^*, I_i) \geq \bar{U},$$

$$\sum_{i=1}^{n} \pi_i(a^*)U(a^*, I_i) \geq \sum_{i=1}^{n} \pi_i(a)U(a, I_i) \quad \text{for all} \quad a \varepsilon A,$$

$$I_i \varepsilon \mathcal{I} \quad \text{for all} \quad i.$$

This problem can be simplified considerably in view of (A1) and (A1'). It will be convenient to regard $v_1 = V(I_1), \ldots, v_n = V(I_n)$ as the principal's control variables. Let $\mathcal{V} = V(\mathcal{I}) = \{v | v = V(I)$ for some $I \varepsilon \mathcal{I}\}$. By (A1) and (A1'), $\mathcal{V}$ is an interval of the real line $(-\infty, \bar{v})$. Furthermore (A1) and (A1') imply that the first constraint in (2.1) is binding -- for if not, costs can be reduced and all constraints will still be satisfied if we replace v_i by $v_i - \epsilon$ under (A1) or by $v_i(1-\epsilon)$ under (A1'), where $\epsilon > 0$ is small. Hence (2.1) can be rewritten as

(2.2) Choose $v_1, \ldots, v_n$ to minimize $\sum \pi_i(a^*)h(v_i)$

$$\text{S.T.} \quad \sum_{i=1}^{n} \pi_i(a^*)v_i = V(C_{FB}(a^*))$$

$$\sum_{i=1}^{n} \pi_i(a)v_i \leq V(C_{FB}(a)) \quad \text{for all} \quad a \in A,$$
$$v_i \in \mathcal{V} \quad \text{for all} \quad i \quad ,$$

where $h = V^{-1}$.

The important point to realize is that the constraints in (2.2) are linear in the v_i's. Furthermore, V concave $\Rightarrow$ h convex, and so the objective function is convex in the v_i's. Thus (2.2) is a rather simple optimization problem: minimize a convex function subject to (a possibly infinite number of) linear constraints. In particular, when A is a finite set, the Kuhn-Tucker theorem yields necessary and sufficient conditions for optimality. These will be analyzed later.

It is important to realize that, in the absence of (A1) or (A1'), it is not generally possible to convert (2.1) into a convex problem.

Definition: If $\underset{\sim}{I} = (I_1,\ldots,I_n)$ satisfies the constraints in (2.1) or $\underset{\sim}{v} = (v_1,\ldots,v_n)$ satisfies the constraints in (2.2), we will say that $\underset{\sim}{I}$ or $\underset{\sim}{v}$ implements action a^*. (We are assuming here that if the agent is indifferent between two actions, he will choose the one preferred by the principal.)

In order to establish the existence of a solution to (2.2), we need a further assumption.

(A3) For all $a \in A$ and $i = 1,\ldots,I$, $\pi_i(a) > 0$.

This assumption rules out cases studied by Mirrlees [1979] in which an optimum can be approached but not achieved.

Lemma 1. Assume (A1) or (A1'), (A2) and (A3). Then, if the constraint set of (2.2) is not empty for $a^* \in A$, (2.2) has a solution. If V is strictly concave, the solution is unique.

Proof. If V is linear, then all elements in the constraint set which satisfy $\sum\pi_i(a^*)U(a^*,I_i) = \bar{U}$ have the same expected cost and so a minimum certainly exists. Assume therefore that V is not linear. Now $\sum\pi_i(a^*)v_i$ is bounded over the constraint set. It therefore follows from a result of Bertsekas [1974] that unbounded sequences in the constraint set make $\sum\pi_i(a^*)h(v_i)$ unbounded (roughly because the variance of the $v_i \to \infty$). Hence we can bound the constraint set. The existence of a minimum follows from Weierstrass' theorem.

Uniqueness follows from the fact that V strictly concave $\Rightarrow$ h strictly convex. Q.E.D.

Definition: Let $C(a^*)$ be the minimized value of $\sum_{i=1}^{n} \pi_i(a^*)h(v_i)$ in (2.2) if the constraint set is non-empty. In the case where the constraint set of (2.2) is empty, write $C(a^*) = \infty$. This defines the second-best cost function $C: A \to R\cup\{\infty\}$.

Remark 1. The above analysis is based on the assumption that the function $V(I)$ is unbounded below (see (A1)). In the absence of this assumption, it is no longer the case that the constraint $\sum\pi_i(a^*)U(a^*,I_i) \geq \bar{U}$ necessarily holds with equality at an optimum. This complicates matters slightly, but does not alter the fact that (2.1) can be converted into a convex problem with linear constraints along the lines of (2.2).

The above constitutes the first step(s) of the principal's optimization problem: for each $a \in A$, compute $C(a)$. The second step is to choose which action to implement, i.e. to choose $a \in A$ to maximize $B(a) - C(a)$. This second problem will not generally be a convex problem. This is because even if $B(a)$ is concave in a, $C(a)$ will not generally be convex. As long as $\pi(a), C(a)$ are differentiable, however, the calculus will yield necessary conditions for optimality in the second problem. These conditions can then be combined with the conditions for optimality in (2.2) to yield overall necessary conditions for optimality.

Definition: A second-best optimal action is one which maximizes $B(a) - C(a)$ on A. A second-best optimal incentive scheme is one that implements a second-best optimal action at least expected cost.

Proposition 1 establishes the existence of a second-best optimum.

Proposition 1. Assume (A1) or (A1'), (A2) and (A3). Then $C(a)$ is a lower semicontinuous function of a, satisfying $C(a^*) = C_{FB}(a^*)$ if a^* minimizes $C_{FB}(a)$ on A. In particular, the problem $\max_{a \in A} (B(a)-C(a))$ has a solution.

Proof: If A is finite, then any function defined on A is continuous. Assume therefore that A is not finite. Let (a_r) be a sequence of points in A converging to a. Assume without loss of generality (w.l.o.g.) that $C(a_r) \to k$. Then, if $k = \infty$, we certainly have $C(a) \leq \lim_{r\to\infty} C(a_r)$. Suppose therefore that $k < \infty$. Let $(I_1^r, \ldots, I_n^r)$ be the cost minimizing way of implementing a_r. Then, if (A3) holds, the argument used in the proof of Lemma 1 shows that the sequence $((I_1^r, \ldots, I_n^r))$ is bounded. Let $(I_1, \ldots, I_n)$ be a limit point. Then $(I_1, \ldots, I_n)$ implements a and so $C(a) \leq \sum \pi_i(a) I_i = \lim_{r\to\infty} C(a_r)$. This proves lower semicontinuity.

To prove that $C(a^*) = C_{FB}(a^*)$ if a^* minimizes $C_{FB}(a)$ on A, note that a^* can be implemented by setting $I_i = C_{FB}(a^*)$ for all i in this case; i.e. there is is no trade-off between risk sharing and incentives when the action to be implemented is a cost-minimizing one. Q.E.D.

Definition: Let $L = \max_{a \in A} (B(a) - C_{FB}(a)) - \max_{a \in A} (B(a) - C(a))$ be the difference between the principal's expected profit in the first-best and second-best situations.

L represents the loss which the principal incurs as a result of being unable to observe the agent's action. Proposition 2 shows that, while there are some special cases in which $L = 0$, in general $L > 0$.

Proposition 2. Assume (A1) or (A1'), and (A2). Then: (1) $C(a) \geq C_{FB}(a)$ for all $a \in A$, which implies that $L \geq 0$. (2) If V is linear, $L = 0$. (3) If there exists a first-best optimal action $a^* \in A$ satisfying: for each i, $\pi_i(a^*) > 0 \Rightarrow \pi_i(a) = 0$ for all $a \in A$, $a \neq a^*$, then $L = 0$. (4) If A is a finite set and there is a first-best optimal action a^* which satisfies: for some i, $\pi_i(a^*) = 0$ and $\pi_i(a) > 0$ for all $a \in A$, $a \neq a^*$, then $L = 0$. (5) If there is a first-best optimal action $a^* \in A$ which minimizes $C_{FB}(a)$ on A, $L = 0$. (6) If (A3) holds, every maximizer $\tilde{a}$ of $B(a) - C_{FB}(a)$ on A satisfies $C_{FB}(\tilde{a}) > \min_{a \in A} C_{FB}(a)$, and V is strictly concave, then $L > 0$.

Proof: (1) is obvious since anything which is second-best feasible is also first-best feasible. To prove (2), let a^* maximize $B(a) - C_{FB}(a)$. Let the principal offer the agent the incentive scheme $I_i = q_i - t$, where $t = B(a^*) - C_{FB}(a^*)$. Then the principal's profit will be $B(a^*) - C_{FB}(a^*)$. On the other hand, by picking $a = a^*$, the agent can obtain expected utility $\bar{U}$.

(5) follows from Proposition 1. (3) and (4) follow from the fact that a^* can be implemented by offering the agent $I_i = C_{FB}(a^*)$ for those i such that $\pi_i(a^*) > 0$ and I close to $\underline{I}$ otherwise.

To prove (6), note that, if V is strictly concave,

$$\sum \pi_i(a^*)V(I_i) \geq V(C_{FB}(a^*))$$

implies

$$C(a^*) = \sum_{i=1}^{n} \pi_i(a^*)h(V(I_i)) > h(V(C_{FB}(a^*))) = C_{FB}(a^*)$$

unless I_i = constant with probability 1. But, since $\pi_i(a^*) > 0$ for all i, I_i = constant with probability $1 \Rightarrow I_i$ is independent of i. However, in this case, the constraints of problem (2.2) imply that $C_{FB}(a)$ is minimized at a^*.

Q.E.D.

Most of Proposition 2 is well known. Proposition 2(2) and (6) can be understood as follows. In the first-best situation, if the agent is strictly risk averse, the principal bears all the risk and the agent bears none. In the second best situation, this is generally undesirable. For if the agent is completely protected from risk, then he has no incentive to work hard; i.e., he will choose $a \in A$ to minimize $C_{FB}(a)$. Hence the second-best situation is strictly worse from a welfare point of view than the first-best situation. The exception is when the agent is risk neutral, in which case it is optimal both from a risk sharing and an incentive point of view for him to bear all the risk.

In the case of Proposition 2(4), a scheme in which the agent is penalized very heavily if certain outcomes occur can be used to achieve the first best. This relates to results obtained in Mirrlees [1979].

Remark 2. We have assumed that the principal is risk neutral. Our analysis generalizes to the case where the principal is risk averse, however. In this case, instead of choosing $\underset{\sim}{v}$ to minimize $\sum \pi_i(a^*)h(v_i)$ in problem (2.2), we choose $\underset{\sim}{v}$ to maximize $\sum \pi_i(a^*)U_p(q_i - h(v_i))$, where U_p is the principal's utility function. Note that (2.2) is still a convex problem. Although we can no longer analyze costs and benefits separately, we can, for each $a \in A$, define a a net benefit function $\underset{\underset{\sim}{v}}{\text{Max}} \sum \pi_i(a^*)U_p(q_i - h(v_i))$. An optimal action for the principal is now one that maximizes net benefits.

Remark 3. We have taken the outcomes observed by the principal to be profit levels.

Our analysis generalizes, however, to the case where the outcomes are more complicated objects, such as vectors of profits, sales, etc. The important point to realize is that output does not appear in the cost minimization problem (2.1) or (2.2). Thus, if the principal observes the realizations of a signal $\tilde{\theta}$, then I_i refers to the payment to the agent when $\tilde{\theta} = \theta_i$. Let $\hat{C}(a,\tilde{\theta})$ be the cost of implementing a when the information structure is $\tilde{\theta}$ (e.g. if $\tilde{\theta}$ reveals a exactly then $\hat{C}(a,\tilde{\theta}) = C_{FB}(a)$). Note that if the distribution of output is generated by a production function $f(a,\tilde{w})$, such that the marginal distribution of $\tilde{w}$ is independent of the information structure, then $B(a) = Ef(a,\tilde{w}) = E[E[f(a,\tilde{w}) | \theta]]$ is independent of the information structure, given a. It follows that the effect of changes in the information structure is summarized by the way that $C(a,\tilde{\theta})$ changes when the information structure changes. As will be seen in Section 5, this is easy to analyze.

3. Some Characteristics of Optimal Incentive Schemes

It is of interest to know whether the optimal incentive scheme is monotone increasing (i.e., whether the agent is paid more when a higher output is observed) and whether the scheme is progressive (i.e., whether the marginal benefit to the agent of increased output is decreasing in output). These questions are quite difficult to answer because of the informational role of output. As we noted in the introduction, the agent may be given a low income at intermediate levels of output in order to discourage particular effort levels. Nevertheless, some general results about the shape of optimal schemes can be established. We begin with the following lemma.

Lemma 2. Assume (A1) or (A1'), (A2) and (A3). Let $(I_i)_{i=1}^n, (I_i')_{i=1}^n$ be incentive schemes which cause a and a' to be optimal choices for the agent, respectively, and minimize the respective costs (i.e. (2.1) or (2.2) is solved). Let $v_i = V(I_i)$ and $v_i' = V(I_i')$. Then

$$(3.1) \qquad \sum_i [\pi_i(a') - \pi_i(a)](v_i' - v_i) \geq 0 \quad .$$

Proof: From (2.2),

$$\sum_i \pi_i(a')v_i \leq V(C_{FB}(a')) = \sum \pi_i(a')v_i' \qquad \text{and}$$

$$\sum_i \pi_i(a)v_i' \leq V(C_{FB}(a)) = \sum \pi_i(a)v_i \quad .$$

Adding these two inequalities yields (3.1). Q.E.D.

We now use Lemma 2 to show that an optimal incentive scheme will have the property that the principal's and agent's returns are positively related over some range of output levels; i.e., it is not optimal to have, for all output levels q_i, q_j, $I_i > I_j \Longrightarrow q_i - I_i < q_j - I_j$. The proof proceeds by showing that, if the principal's and agent's payments are negatively related, then there is a twist in the incentive schedule which raises the agent's payment in high return states for the principal and lowers it in low return states for the principal, and which is good for incentives since it gets the agent to put more probability weight on states yielding the principal a high return. Such a twist is also desirable for risk-sharing. Since the incentive and risk-sharing effects reinforce each other, the principal is made better off.

In order to bring about both the incentive and risk-sharing effects, the twist in the incentive scheme must be chosen carefully. It is for this reason that the proof of the next proposition may seem rather complicated at first sight.

Proposition 3. Assume (A1) or (A1'), (A2), (A3) and V strictly concave. Let $(I_1,\dots,I_n)$ be a second-best optimal incentive scheme. Then the following cannot be true: $I_i > I_j \Longrightarrow q_i - I_i \leq q_j - I_j$ for all $1 \leq i,j \leq n$ and for some i,j ,

$I_i > I_j$ and $q_i - I_i < q_j - I_j$.

Proof: Suppose that

(3.2) $$I_i > I_j \Rightarrow q_i - I_i \leq q_j - I_j \text{ for all } 1 \leq i,j \leq n \text{ and for some } i,j\ ,\ I_i > I_j \text{ and } q_i - I_i < q_j - I_j \ .$$

Let $(I_1',\ldots,I_n')$ be a new incentive scheme satisfying

(3.3) $$v_i' + \lambda h(v_i') = v_i + \lambda q_i - \mu \quad \text{for all } i\ ,$$

where $v_i = V(I_i)$, $v_i' = V(I_i')$, $\lambda > 0$ and μ is such that

(3.4) $$\lambda \operatorname*{Max}_i (q_i - h(v_i)) \geq \mu \geq \lambda \operatorname*{Min}_i (q_i - h(v_i)) \ .$$

If $\lambda = \mu = 0$, then $v_i' = v_i$ solves (3.3). The implicit function theorem therefore implies that (3.3) has a solution as long as λ,μ are small. (Even if h is not differentiable it has left and right hand derivatives.)

It follows from (3.2) and (3.4) that the change to the new incentive scheme has the effect of increasing the lowest I_i's and decreasing the highest ones. Pick μ so that $\sum\pi_i(a')v_i' - V(C_{FB}(a')) = \operatorname*{Max}_{a\epsilon A} [\sum\pi_i(a)v_i' - V(C_{FB}(a))] = 0$. This ensures that the agent's expected utility remains at $\bar{U}$. We now show that the principal's expected profit is higher under the new incentive scheme than under the old, which contradicts the optimality of $(I_1,\ldots,I_n)$.

Substituting (3.1) of Lemma 2 into (3.3) yields:

$$\sum_i \pi_i(a')(q_i - h(v_i')) \geq \sum_i \pi_i(a)(q_i - h(v_i')) \ .$$

If we can show that $\sum\pi_i(a)h(v_i') < \sum\pi_i(a)h(v_i)$, it will follow that

$$\sum \pi_i(a')(q_i - h(v_i')) > \sum \pi_i(a)(q_i - h(v_i))\ ,$$

i.e., the principal is better off.

To see that $\sum\pi_i(a)h(v_i') < \sum\pi_i(a)h(v_i)$, note that

$$\sum \pi_i(a)(h(v_i) - h(v_i')) \geq \sum \pi_i(a)h'(v_i')(v_i - v_i')$$

by the convexity of h (here h' is the right-hand derivative if h is not differentiable). It suffices therefore to show that the latter expression is positive. By (3.3),

$$\sum \pi_i(a)h'(v_i')(v_i - v_i') = \sum\pi_i(a)h'(v_i')(\lambda h(v_i') - \lambda q_i + \mu) \quad .$$

Suppose that this is nonpositive for small λ. Dividing by λ, letting $\lambda \to 0$, assuming that μ/λ converges to $\hat{\mu}$ w.l.o.g., and using the fact that $v_i' \to v_i$, we get

(3.5) $$\sum \pi_i(a)h'(v_i)(h(v_i) - q_i + \hat{\mu}) \leq 0 \quad .$$

However, by (3.2), $h'(v_i)$ and $(h(v_i) - q_i))$ are positively correlated; i.e., as one moves up so does the other. Therefore, by Hardy, Littlewood and Polya [1952, p.43],

(3.6) $$\sum \pi_i(a)h'(v_i)(h(v_i) - q_i + \hat{\mu}) > (\sum\pi_i(a)h'(v_i))(\sum\pi_i(a)(h(v_i) - q_i + \hat{\mu})) \geq 0 \quad ,$$

where the last inequality follows from the fact that (1) $h' \geq 0$; (2) $\sum\pi_i(a)v_i' - V(C_{FB}(a)) \leq 0 = \sum\pi_i(a)v_i - V(C_{FB}(a))$, which implies that $\lim_{\lambda\to 0} (1/\lambda)\sum\pi_i(a)(v_i-v_i') \geq 0$. (3.6) contradicts (3.5).

This proves that $\sum\pi_i(a)h(v_i') < \sum\pi_i(a)h(v_i)$, which establishes that the principal's expected profit is higher under $(I_1',\ldots,I_n')$. Contradiction.

Remark 4. There is an interesting contrast between Proposition 3 and results found in the literature on optimal risk sharing in the absence of moral hazard.

In this literature, it is shown that (if the individuals are risk averse) it is optimal for the individuals' returns to be positively related over the whole range of outputs, rather than over just some range of outputs. See Borch [1968].

Proposition 3 may be used to establish the following result about the monotonicity of the optimal incentive scheme.

Proposition 4. Assume (A1) or (A1'), (A2), (A3) and V strictly concave. Let $(I_1,\ldots,I_n)$ be a second-best optimal incentive scheme. Then (1) there exists $1 \leq i \leq n-1$ such that $I_i \leq I_{i+1}$, with strict inequality unless $I_1 = I_2 = \ldots = I_n$; (2) there exists $1 \leq j \leq n-1$ such that $q_j - I_j < q_{j+1} - I_{j+1}$.

Proof: (1) follows directly from Proposition 3. So does (2) once we rule out the case $q_1 - I_1 = q_2 - I_2 = \ldots = q_n - I_n$. We do this by a similar argument to that used in Proposition 3. Suppose that $\underset{\sim}{I}$ is an optimal incentive scheme satisfying

$$(3.7) \qquad q_1 - I_1 = q_2 - I_2 = \ldots = q_n - I_n = k \quad .$$

Then $I_1 < I_2 < \ldots < I_n$. Consider a new incentive scheme $\underset{\sim}{I}' = (I_1+\epsilon, I_2+\epsilon, \ldots, I_{n-1}+\epsilon, I_n-\mu\epsilon)$ where $\epsilon > 0$ and μ is chosen so that $\underset{a\epsilon A}{\text{Max}}[\sum\pi_i(a)V(I_i')-V(C_{FB}(a))] = 0$. We show that the principal's expected profit is higher under $\underset{\sim}{I}'$ than under $\underset{\sim}{I}$ for small ϵ. Suppose not. Then

$$\sum \pi_i(a')(q_i - I_i') \leq \sum \pi_i(a)(q_i - I_i) = k \quad ,$$

where a' (resp. a) is optimal for the agent under $\underset{\sim}{I}'$ (resp. $\underset{\sim}{I}$). Substituting for $\underset{\sim}{I}'$ yields

$$-(1 - \pi_n(a'))\epsilon + \pi_n(a')\mu\epsilon \leq 0 \quad .$$

Take limits as $\epsilon \to 0$. W.l.o.g. $a' \to \hat{a}$. Hence we have

$$(3.8) \qquad -(1 - \pi_n(\hat{a})) + \pi_n(\hat{a})\mu \leq 0 \quad .$$

Now since a' is an optimal action for the agent under $\underset{\sim}{I}'$, it follows by uppersemicontinuity that $\hat{a}$ is optimal under $\underset{\sim}{I}$. Hence we have

$$\sum \pi_i(\hat{a})V(I_i') - V(C_{FB}(\hat{a})) \leq 0 = \sum \pi_i(\hat{a})V(I_i) - V(C_{FB}(\hat{a})) \quad .$$

Using the concavity of V and taking limits as $\epsilon \to 0$, we get

$$\sum_{i=1}^{n-1} \pi_i(\hat{a})V'(I_i) - \pi_n(\hat{a})V'(I_n)\mu \leq 0 \quad .$$

But since $V'(I_i)$ is decreasing in i, this contradicts (3.8). (If V is not differentiable, V' denotes the right-hand derivative.)

This proves that the principal does better under $\underset{\sim}{I}'$ than under $\underset{\sim}{I}$. Hence we have ruled out the case $q_1 - I_1 = \cdots = q_n - I_n$. This establishes Proposition 4. Q.E.D.

Proposition 4 says that it is not optimal for the agent's marginal reward as a function of income to be negative everywhere or to be greater than or equal to one everywhere.[4/] However, the proposition does allow for the possibility that either of these conditions can hold over some interval. To see when this may occur, it is useful to consider in more detail the case where A is a finite set. When A is finite, we can use the Kuhn-Tucker conditions for problem (2.2) to characterize the optimum. If (A3) holds and h is differentiable, these yield:

$$\text{(3.9)} \qquad h'(v_i) = \lambda - \sum_{\substack{k \in A \\ a_k \neq a^*}} \mu_k \frac{\pi_i(a_k)}{\pi_i(a^*)} \quad ,$$

where λ, (μ_k) are nonnegative Lagrange multipliers and $\mu_k > 0$ only if the agent is indifferent between a* and a_k at the optimum. The following proposition states that $\mu_k > 0$ for at least one action which is less costly than a*. This implies that at an optimum the agent must be indifferent between at least two actions (unless a* is the least costly action, i.e. where there is no incentive problem).

Proposition 5. Assume (A1) or (A1'), (A2), (A3) and A finite. Suppose that (2.2) has a solution for $a^* \varepsilon A$. Then if $C_{FB}(a^*) > \min_{a \varepsilon A} C_{FB}(a)$, this solution will have the property that $\sum \pi_i(a^*)v_i - V(C_{FB}(a^*)) = \sum \pi_i(a_k)v_i - V(C_{FB}(a_k))$ for some $a_k \varepsilon A$ with $C_{FB}(a_k) < C_{FB}(a^*)$. Furthermore, if V is strictly concave and differentiable, the Lagrange multiplier μ_k will be strictly positive for some a_k with $C_{FB}(a_k) < C_{FB}(a^*)$.

Proof: Suppose that the agent strictly prefers a^* to all actions less costly than a^* at the solution. Then, since (2.2) is a convex problem, we can drop all the constraints in (2.2) which refer to less costly actions without affecting the solution. In other words, we can substitute $A' = \{a \varepsilon A | a$ is at least as costly as $a^*\}$ for A in (2.2) and the solution will not change. But since a^* is now the least costly action, we know from the proof of Proposition 1 that it is optimal to set $I_i = I_j$ for all i,j. However, $I_i = I_j$ is not optimal for the original problem since, under these conditions, the agent will pick an a which minimizes $C_{FB}(a)$, and by assumption $C_{FB}(a^*) > \min_{a \varepsilon A} C_{FB}(a)$. Contradiction.

That $\mu_k > 0$ follows from the fact that if all the $\mu_k = 0$, then $h'(v_i)$ is the same for all i, which implies that $I_1 = \dots = I_n$; however, this means that the agent will choose a cost-minimizing action, contradicting $C_{FB}(a^*) > \min_{a \varepsilon A} C_{FB}(a)$. Q.E.D.

The simplest case occurs when $\mu_k > 0$ for just one a_k with $C_{FB}(a_k) < C_{FB}(a^*)$ (this will be true in particular if A contains only two actions). In this case, we can rewrite (3.9) as

$$(3.10) \qquad h'(v_i) = \lambda - \mu_k \frac{\pi_i(a_k)}{\pi_i(a^*)} .$$

We see that what determines v_i, and hence I_i, in this case is the relative likelihood that the outcome $q = q_i$ results from a_k rather than a^*. In particular, since h convex $\Longrightarrow h'$ increasing in v_i, a sufficient condition for the optimal incentive scheme to be increasing everywhere, i.e. $I_1 \leq I_2 \leq \cdots \leq I_n$, is that $\pi_i(a_k)/\pi_i(a^*)$ is decreasing in i, i.e. the relative likelihood that $a = a_k$ rather than $a = a^*$ produces the outcome $q = q_i$ is lower the better is the outcome i.

This observation has led some to suggest that the following is a sufficient condition for the incentive scheme to be increasing.

<u>Monotone Likelihood Ratio Condition</u> (MLRC). Assume (A3). Then MLRC holds if, given $a, a' \in A$, $C_{FB}(a') \leq C_{FB}(a)$ implies that $\pi_i(a')/\pi_i(a)$ is decreasing in i.[5/]

It should be noted that the "first-order condition" approach described in the introduction, which is based on the assumption that the agent is indifferent between a and $a + da$ at an optimum, does yield MLRC as a sufficient condition for monotonicity.[6/] We now show, however, that, once we take into account the possibility that the agent may be indifferent between several actions at an optimum, i.e. $\mu_k > 0$ for more than one a_k, MLRC does not guarantee monotonicity.

<u>Example 1</u>. $A = \{a_1, a_2, a_3\}$, $n = 3$. $\pi(a_1) = (\frac{2}{3}, \frac{1}{4}, \frac{1}{12})$, $\pi(a_2) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, $\pi(a_3) = (\frac{1}{12}, \frac{1}{4}, \frac{2}{3})$. Assume additive separability with $G(a_1) = 0$, $G(a_2) = \frac{1}{12}\sqrt{2} + \frac{1}{4}\sqrt{7/4}$, $G(a_3) = \frac{7}{12}\sqrt{7/4}$, $V(I) = (3I)^{1/3}$ (i.e. $h(v) = \frac{1}{3}v^3$) and $\bar{U} = \frac{1}{4}\sqrt{2} + \frac{1}{12}\sqrt{7/4}$. Note that MLRC is satisfied here.[7/]

We compute $C(a_1)$, $C(a_2)$, $C(a_3)$. Obviously, $C(a_1) = C_{FB}(a_1) = \frac{1}{3}(G(a_1) + \bar{U})^3 = 0.033$. To compute $C(a_2)$, we use the first-order conditions (3.9). These are

$$v_1^2 = \lambda - 2\mu_1 - \tfrac{1}{4}\mu_2 ,$$

$$v_2^2 = \lambda - \frac{3}{4}\mu_1 - \frac{3}{4}\mu_2 ,$$

$$v_3^2 = \lambda - \tfrac{1}{4}\mu_1 - 2\mu_2 ,$$

plus the complementary slackness conditions. These equations are solved by setting $\lambda = \frac{17}{4}$, $\mu_1 = 2$, $\mu_2 = 1$. This yields $v_1 = 0$, $v_2 = \sqrt{2}$, $v_3 = \sqrt{7/4}$, and the agent is then indifferent between a_1, a_2 and a_3:

$$\frac{2}{3}v_1 + \frac{1}{4}v_2 + \frac{1}{12}v_3 - G(a_1) = \frac{1}{3}v_1 + \frac{1}{3}v_2 + \frac{1}{3}v_3 - G(a_2)$$

$$= \frac{1}{12}v_1 + \frac{1}{4}v_2 + \frac{2}{3}v_3 - G(a_3) = \bar{U} .$$

Since the first-order conditions are necessary and sufficient, we may conclude that $C(a_2) = \frac{1}{3}(\frac{1}{3}v_1^3 + \frac{1}{3}v_2^3 + \frac{1}{3}v_3^3) = 0.571$.

Note that the incentive scheme which implements a_2, $I_1 = 0$, $I_2 = \frac{1}{3}2^{3/2}$, $I_3 = \frac{1}{3}(\frac{7}{4})^{3/2}$, is not monotonically increasing.

Observe that $C(a_3) \geq C_{FB}(a_3) = \frac{1}{3}(G(a_3) + \bar{U})^3 = 0.635 > C(a_2)$. Since $C(a_3) > C(a_2) > C(a_1)$, it is easy to show that we can find $q_1 < q_2 < q_3$ such that $B(a_2) - C(a_2) > \max [B(a_3) - C(a_3), B(a_1) - C(a_1)]$. But this means that it is optimal for the principal to get the agent to pick a_2. Hence the optimal incentive scheme is as described above. It is not increasing despite the satisfaction of MLRC.

The reason that monotonicity breaks down in Example 1 is because, at the optimum, the agent is indifferent between a_2, the action to be implemented, a_1 a less costly action, and a_3 a more costly action. By MLRC, $\frac{\pi_i(a_1)}{\pi_i(a_2)}$, $\frac{\pi_i(a_2)}{\pi_i(a_3)}$ are decreasing in i. However, $\mu_1 \frac{\pi_i(a_1)}{\pi_i(a_2)} + \mu_2 \frac{\pi_i(a_3)}{\pi_i(a_2)}$ need not be monotonic.

This observation suggests that one way to get monotonicity is to strengthen MLRC so that it holds for weighted combinations of actions as well as for the basic actions themselves. In particular, suppose that

(3.11) Given any finite subset $\{a_1,\ldots,a_n\}$ of A, $a \in A$, and nonnegative weights $w_1,\ldots,w_n$ summing to 1, it is the case that $\left(\sum_{j=1}^{n} w_j\pi_i(a_j)/\pi_i(a)\right)$ is either monotonically increasing or decreasing in i.

Then, by the first-order conditions (3.9),

$$h'(v_i) = \lambda - \left(\sum_{\substack{k\in A\\ a_k\neq a^*}} \mu_k\right)\left(\sum_{\substack{k\in A\\ a_k\neq a^*}} w_k\pi_i(a_k)/\pi_i(a)\right), \tag{3.12}$$

where $w_k = \mu_k / \sum_{\substack{h\in A\\ a_h\neq a^*}} \mu_h$. But, by (3.11), the right-hand side (RHS) of (3.12) is monotonic. Hence, the v_i's are either monotonically increasing or decreasing. By Proposition 4, however, they cannot be monotonically decreasing; hence they are monotonically increasing.

Unfortunately, (3.11) turns out to be a very strong condition. In fact, it is equivalent to the following spanning condition.

Spanning Condition (SC). There exist $\hat{\pi}, \hat{\pi}' \in S$ such that (1) for each $a \in A$, $\pi(a) = \lambda(a)\hat{\pi} + (1-\lambda(a))\hat{\pi}'$ for some $0 \leq \lambda(a) \leq 1$; (2) $\hat{\pi}_i/\hat{\pi}'_i$ is monotonic in i.

That SC implies (3.11) is easy to see. We are grateful to Jim Mirrlees for pointing out and proving the converse.[8/]

Proposition 6. Assume (A1) or (A1'), (A2), (A3), V strictly concave and differentiable. Suppose that SC holds. Then a second-best optimal incentive scheme satisfies $I_1 \leq I_2 \leq \cdots \leq I_n$.

Proof: If A is finite, the argument following (3.12) establishes the result. To establish the result for the case A infinite, we approximate A by a finite subset and take limits (recall that A is compact). Q.E.D.

An alternative sufficient condition for monotonicity may be found in the work of Mirrlees [1979], who establishes a similar result to Proposition 7 below. For each $a \in A$, let $F(a) = (\pi_1(a), \pi_1(a) + \pi_2(a), \ldots, \pi_1(a) + \ldots + \pi_n(a))$. In the following proposition, the notation $F(a) \geq F'(a)$ is used to mean $F_i(a) \geq F_i'(a)$ for all $i = 1,\ldots,n$.

Concavity of Distribution Function Condition (CDFC). CDFC holds if $a, a', a'' \in A$ and $V(C_{FB}(a)) = \lambda V(C_{FB}(a')) + (1-\lambda)V(C_{FB}(a''))$, $0 \leq \lambda \leq 1$, implies that $F(a) \geq \lambda F(a') + (1-\lambda)F(a'')$.

Proposition 7. Assume (A1) or (A1'), (A2), (A3), V strictly concave and differentiable. Suppose that MLRC and CDFC hold. Then a second-best optimal incentive scheme $(I_1,\ldots,I_n)$ satisfies $I_1 \leq I_2 \leq \ldots \leq I_n$.

Proof: Assume first that A is finite. Let a^* maximize $B(a) - C(a)$. Let $A' = \{a \in A | C_{FB}(a) \leq C_{FB}(a^*)\}$. Consider the cost minimizing way of getting the agent to pick a^* given that he can choose only from A'. It is clear from (3.9) that, since $\pi_i(a_k)/\pi_i(a^*)$ is decreasing in i by MLRC, the incentive scheme $(I_1,\ldots,I_n)$ is monotonically increasing. We will be home if we can show that $(I_1,\ldots,I_n)$ is optimal when A' is replaced by A. Since adding actions cannot reduce the cost of implementing a^*, all we have to do is to show that $(I_1,\ldots,I_n)$ continues to implement a^*, i.e. there does not exist a'', $C_{FB}(a'') > C_{FB}(a^*)$, such that

$$(3.13) \qquad \sum \pi_i(a'')v_i - V(C_{FB}(a'')) > 0 = \sum \pi_i(a^*)v_i - V(C_{FB}(a^*)) \quad .$$

However, we know from Proposition 5 that

$$(3.14) \qquad \sum \pi_i(a')v_i - V(C_{FB}(a')) = 0 = \sum \pi_i(a^*)v_i - V(C_{FB}(a^*)) \quad .$$

for some a' with $C_{FB}(a') < C_{FB}(a^*)$. Writing $V(C_{FB}(a)) = \lambda V(C_{FB}(a')) + (1-\lambda)V(C_{FB}(a''))$ and using CDFC and the fact that $v_1 \leq v_2 \leq \cdots \leq v_n$, we get

$$\sum \pi_i(a^*)v_i - V(C_{FB}(a^*)) \geq \lambda \sum \pi_i(a'')v_i + (1-\lambda)\sum \pi_i(a')v_i - V(C_{FB}(a^*))$$

$$= \lambda\Big(\sum\pi_i(a'')v_i - V(C_{FB}(a''))\Big) + (1-\lambda)\Big(\sum \pi_i(a')v_i - V(C_{FB}(a'))\Big) \ .$$

which contradicts (3.13) and (3.14).

To prove the result for A infinite, one again proceeds by way of a finite approximation. Q.E.D.

So far we have considered only the monotonicity of the optimal incentive scheme.[9/] One would also like to know when the optimal incentive scheme is <u>progressive</u>, i.e. $(I_{i+1} - I_i)/(q_{i+1} - q_i)$ is decreasing in i, or <u>regressive</u>, i.e. $(I_{i+1} - I_i)/(q_{i+1} - q_i)$ is increasing in i. To get results about this, one needs considerably stronger assumptions, as the following proposition indicates.

<u>Proposition 8</u>. Assume (A1) or (A1'), (A2), (A3), V strictly concave and differentiable. Suppose that MLRC and CDFC hold and that $(q_{i+1} - q_i)$ is independent of i, $1 \leq i \leq n-1$. Then a second-best optimal incentive scheme will be regressive if

(3.15) $(1/V'(I))$ is concave in I and $a, a' \in A$, $C_{FB}(a') < C_{FB}(a)$,

$$\text{implies that } \frac{\pi_{i+1}(a')}{\pi_{i+1}(a)} - \frac{\pi_i(a')}{\pi_i(a)} \text{ is decreasing in } i.$$

It will be progressive if

(3.16) $(1/V'(I))$ is convex in I and $a, a' \in A$, $C_{FB}(a') < C_{FB}(a)$,

$$\text{implies that } \frac{\pi_{i+1}(a')}{\pi_{i+1}(a)} - \frac{\pi_i(a')}{\pi_i(a)} \text{ is decreasing in } i.$$

Proof: Assume first that A is finite. Let a^* be a second-best optimal action. Let a' maximize $C_{FB}(a)$ subject to $C_{FB}(a) < C_{FB}(a^*)$, i.e. a' is the next most costly action after a^*. Consider the cost minimizing way of implementing a^* given that a' is the only other action that the agent can choose. Using the same concavity argument as in the proof of Proposition 7, we can show that the resulting incentive scheme $(I_1,\ldots,I_n)$ also implements a^* when the agent can choose from all of A. Hence $(I_1,\ldots,I_n)$ is an optimal incentive scheme.

By (3.10),

$$\frac{1}{V'(I_i)} = h'(v_i) = \lambda - \mu\frac{\pi_i(a')}{\pi_i(a^*)} ,$$

and so

$$\frac{1}{V'(I_{i+1})} - \frac{1}{V'(I_i)} = -\mu\left(\frac{\pi_{i+1}(a')}{\pi_{i+1}(a^*)} - \frac{\pi_i(a')}{\pi_i(a^*)}\right) .$$

(3.15) and (3.16) now follow immediately. Q.E.D.

Note that $\frac{1}{V'}$ is linear if $V = \log I$; is concave if $V = -e^{-\alpha I}$, $\alpha > 0$, or $V = I^{\alpha}$, $0 < \alpha < 1$; is convex if $V = -I^{-\alpha}$, $\alpha > 1$.

Let us summarize the results of this section. We have shown that an optimal incentive scheme will not be declining everywhere, but that only under quite strong assumptions (SC or MLRC plus concavity) will it be increasing everywhere. We have also shown that it is not optimal for the agent's marginal remuneration for an extra pound of profit to exceed one everywhere, although it may exceed one sometimes. Finally, we have obtained sufficient conditions for the incentive scheme to be progressive or regressive.

In the next section, we show that considerably stronger results can be proved when $n = 2$. We also provide a simple algorithm for computing optimal incentive schemes when $n = 2$.

4. The Case of Two Outcomes

When $n = 2$, we will refer to q_1 as the "bad" outcome and $q_2 > q_1$ as the "good" outcome. In this case, the agent's incentive scheme can be represented simply by a fixed payment w and a share of profits s, where $w + sq_1 = I_1$, $w + sq_2 = I_2$, i.e. $s = (I_2-I_1)/(q_2-q_1)$. (Note that in principle s can be negative or can exceed one.) In addition, we know from (2.2) that, at an optimum,

$$(4.1) \qquad \underset{a\varepsilon A}{\mathrm{Max}}\{\pi_1(a)V(w + sq_1) + \pi_2(a)V(w + sq_2) - V(C_{FB}(a))\} = 0$$

The expression in the brackets is strictly increasing in w. Therefore, given s, there is exactly one value of w satisfying (4.1). It follows that, when $n = 2$, an incentive scheme is completely characterized by the value of s.

Proposition 4 of the last section showed that I_i cannot always be declining. When $n = 2$, this means that $s \geq 0$.[10/] Similarly, the proposition shows that $s < 1$ when $n = 2$. This has a number of interesting implications.

Definition. Let $n = 2$. We say that $a \varepsilon A$ is efficient if there does not exist $a' \varepsilon A$ satisfying $C_{FB}(a') \leq C_{FB}(a)$ and $\pi_2(a') \geq \pi_2(a)$, with at least one strict inequality.

In other words, an action is efficient if the probability of a good outcome can only be increased by incurring greater cost.

Proposition 9. Assume (A1) or (A1'), (A2), (A3) and V strictly concave. Let $n = 2$. Then every second-best optimal action is efficient.

Proof: Since $s \geq 0$, it is not in the interest of the agent to choose an inefficient action. Furthermore, if $s = 0$ and the agent is indifferent between an efficient action and an inefficient one, then the principal will prefer the efficient one. Q.E.D.

Proposition 4, for the case $n = 2$, has a second interesting implication. Suppose that we start off in a situation where the agent has access to a set of actions A, and now some additional actions become available, so that the new action set is $A' \supset A$. Then, if the new actions are all higher cost actions for the agent than those in A -- in the sense that their C_{FB}'s are higher -- the principal cannot be made worse off by such a change.

Proposition 10. Assume (A1) or (A1') and (A2). Let $n = 2$. Suppose that $A' \supset A$ and that $a \in A$, $a' \in A' \setminus A \Rightarrow C_{FB}(a') \geq C_{FB}(a)$. Assume that (A3) holds for both A and A'. Then $\underset{a \in A'}{\text{Max}} [B(a) - C'(a)] \geq \underset{a \in A}{\text{Max}} [B(a) - C(a)]$, where C' is the second-best cost function under A'.

Proof: Suppose (I_1, I_2) is an optimal second-best incentive scheme when the action set is A. Let the principal keep this incentive scheme when the new actions $A' \setminus A$ are added. The only way that the principal can be made worse off is if the agent now switches from $a \in A$ to $a' \in A' \setminus A$. But a' must then provide higher utility for the agent. Since $C_{FB}(a') \geq C_{FB}(a)$, this can only be the case if $\pi_1(a')v_1 + \pi_2(a')v_2 > \pi_1(a)v_1 + \pi_2(a)v_2$, which implies, since $v_2 \geq v_1$ by Proposition 4, that $\pi_2(a') \geq \pi_2(a)$. But it follows that the principal's expected profits $\pi_1(q_1 - I_1) + \pi_2(q_2 - I_2)$ rise (or stay the same) when the agent moves from a to a' since, again by Proposition 4, $s < 1$, i.e. $q_2 - I_2 > q_1 - I_1$. Q.E.D.

As a final implication of Proposition 4, when $n = 2$, consider a manager-entrepreneur who initially owns 100% of a firm, i.e. $\tilde{w} = 0$, $\tilde{s} = 1$. In the absence of any risk sharing possibilities the manager will choose a to maximize $\pi_1(a)U(a, q_1) + \pi_2(a)U(a, q_2)$. Let $\tilde{a}$ be a solution to this. Clearly $\tilde{a}$ is efficient. Now suppose a risk neutral principal appears with whom the manager can share risks. We know from Proposition 3 that at the new optimum $s < 1 = \tilde{s}$. Therefore, by Lemma 4, $\pi_2(a^*) \leq \pi_2(\tilde{a})$. In addition, $C_{FB}(a^*) \leq C_{FB}(\tilde{a})$ by Proposition 9. Thus, the existence of risk-sharing possibilities leads the agent to choose a less costly action with a lower probability of a good outcome.

We may use Propositions 9 and 10 to develop a method for computing a second-best optimal incentive scheme when $n = 2$. We will consider two cases: (1) where A is finite; (2) where A is such that $\{C_{FB}(a) | a \in A\}$ is an interval of the real line.

Recall that Proposition 5 states that when A is finite, the agent will be indifferent between a^* and some less costly action. This fact makes the computation of an optimal incentive scheme fairly straightforward for the case $n = 2$ and A finite. We know from Proposition 9 that it is never optimal to get the agent to choose an inefficient action. Hence we can assume w.l.o.g. that $C_{FB}(a_1) < C_{FB}(a_2) < \ldots < C_{FB}(a_m)$ and $\pi_2(a_1) < \pi_2(a_2) < \ldots < \pi_2(a_m)$. The computation of $C(a_1)$ is easy: by Proposition 1, it is just $C_{FB}(a_1)$. To compute $C(a_k)$, $k > 1$, we use Proposition 5. For each action a_j, $j < k$, find I_1, I_2 so that the agent is indifferent between a_k and a_j and the agent's expected utility is $\bar{U}$. This means solving

$$(4.2) \qquad \begin{aligned} \pi_1(a_k)v_1 + \pi_2(a_k)v_2 &= V(C_{FB}(a_k)) \quad , \\ \pi_1(a_j)v_1 + \pi_2(a_j)v_2 &= V(C_{FB}(a_j)) \quad , \end{aligned}$$

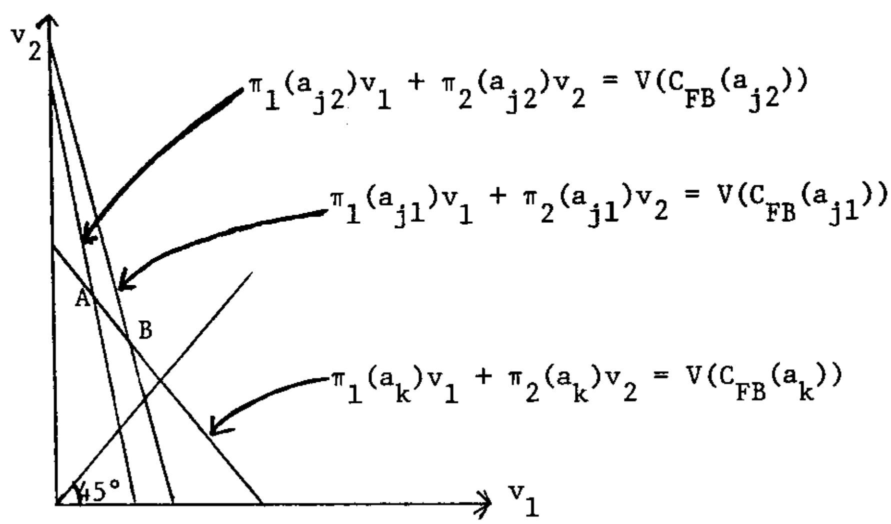
which yields

$$(4.3) \qquad \begin{aligned} v_1 &= \frac{V(C_{FB}(a_j)) - V(C_{FB}(a_k)) + \pi_1(a_j)V(C_{FB}(a_j)) - \pi_1(a_k)V(C_{FB}(a_k))}{\pi_1(a_j) - \pi_1(a_k)} \\ v_2 &= \frac{V(C_{FB}(a_j)) - V(C_{FB}(a_k))}{\pi_1(a_j) - \pi_1(a_k)} \quad . \end{aligned}$$

We then set $I_1 = h(v_1)$, $I_2 = h(v_2)$. Note that $v_1 < v_2$ in (4.3), so that $I_1 < I_2$.

Doing this for each $j = 1,\ldots,k-1$ yields $(k-1)$ different (v_1,v_2) (and (I_1,I_2)) pairs, each with $v_1 < v_2$. This is illustrated in Figure 1 for the case $k = 3$, where the (v_1,v_2) pairs are at A,B.

Figure 1



We know from Proposition 5 that one of these pairs is the solution to (2.2). In fact, the solution must occur at the (v_1,v_2) pair with the <u>smallest</u> v_1 (and hence, by (4.2), with the largest v_2) -- denote this pair by $(\hat{v}_1,\hat{v}_2)$. To see this, suppose that the agent is indifferent between a_k and a_j under $(\hat{v}_1,\hat{v}_2)$. Consider the expression

$$(4.4) \qquad \pi_1(a_k)v_1 + \pi_2(a_k)v_2 - \pi_1(a_j)v_1 - \pi_2(a_j)v_2$$
$$= (\pi_1(a_k) - \pi_1(a_j))v_1 + (\pi_2(a_k) - \pi_2(a_j))v_2 \quad .$$

When $v_1 = \hat{v}_1, v_2 = \hat{v}_2$, this expression equals $V(C_{FB}(a_k)) - V(C_{FB}(a_j))$. Suppose now that $v_1 > \hat{v}_1$, $v_2 > \hat{v}_2$. Then (4.4) falls since $\pi_1(a_k) < \pi_1(a_j)$. Hence the agent now prefers a_j to a_k and so a_k is not implemented.

In Figure 1, the solution is at A. Note that it is possible that the $(\hat{v}_1, \hat{v}_2)$ picked in this way does not lie in $\mathcal{V} \times \mathcal{V}$; i.e. $h(\hat{v}_1)$ or $h(\hat{v}_2)$ may be undefined. In this case, the constraint set of (2.2) is empty and so $C(a_k) = \infty$. If $(\hat{v}_1, \hat{v}_2) \in \mathcal{V} \times \mathcal{V}$, then the principal's minimum expected cost of getting the agent to pick a_k is $\pi_1(a_k)h(\hat{v}_1) + \pi_2(a_k)h(\hat{v}_2)$. Expected net benefits are $B(a_k) - C(a_k)$. Finally, the overall optimum is determined by finding the a which maximizes $B(a_k) - C(a_k)$.

Remark 5. In computing the cost of implementing a_k, we have ignored actions which are more costly for the agent than a_k. This means that the cost function which we have computed is not the true cost function $C(a)$ but a modified cost function $\tilde{C}(a)$. Clearly, $\tilde{C}(a) \leq C(a)$ for each a since more actions can only make implementation more difficult. On the other hand, Proposition 10 tells us that $\underset{a \varepsilon A}{\text{Max}} [B(a) - \tilde{C}(a)] \leq \underset{a \varepsilon A}{\text{Max}} [B(a) - C(a)]$. Combining these yields $\underset{a \varepsilon A}{\text{Max}} [B(a) - C(a)] = \underset{a \varepsilon A}{\text{Max}} [B(a) - \tilde{C}(a)]$, which means that we are justified in working with $\tilde{C}(a)$ instead of $C(a)$.

Another case where computation is quite simple is when $\{C_{FB}(a) | a \varepsilon A\}$ is an interval $[\underline{c}, \bar{c}]$ of the real line. In this case, it is more convenient to regard C_{FB} rather than a as the agent's choice variable and to express π_1, π_2 as functions of C_{FB}. This is legitimate since, by Proposition 9, inefficient actions can be ignored. Proposition 9 implies also that π_2 is strictly increasing in C_{FB}. We will assume in addition that π_2 is differentiable when $C_{FB} \varepsilon (\underline{c}, \bar{c})$, is differentiable to the left at $\bar{c}$, and that $d\pi_2/dC_{FB} > 0$. We also assume that V is differentiable on an open set containing $(\underline{c}, \bar{c})$. To ease the notation, we replace C_{FB} by c, and let $C(c)$ denote the least cost of getting the agent to choose an action a for which $C_{FB}(a) = c$.

If the principal wishes the agent to pick a cost level $\underline{c} < c^* < \bar{c}$, the following conditions must be satisfied:

$$(4.5) \qquad \begin{aligned} \pi_1(c^*)v_1 + \pi_2(c^*)v_2 &= V(c^*) \\ \pi_1'(c^*)v_1 + \pi_2'(c^*)v_2 &= V'(c^*) \end{aligned} \quad ,$$

where $\pi_i'(c) = (d\pi_i/dc)$. The second part of (4.5) follows from the fact that, by (2.2), $\pi_1(c)v_1 + \pi_2(c)v_2 - V(c)$ achieves a maximum at c^*. Using the fact that $\pi_1' + \pi_2' = 0$, we get

$$(4.6) \qquad \begin{aligned} v_1 &= \frac{-\pi_2(c^*)V'(c^*) + \pi_2'(c^*)V(c^*)}{\pi_2'(c^*)} \quad , \\ v_2 &= \frac{-\pi_2(c^*)V'(c^*) + \pi_2'(c^*)V(c^*) + V'(c^*)}{\pi_2'(c^*)} \quad . \end{aligned}$$

Notice that $v_2 = v_1 + (V'(c^*)/\pi_2'(v^*)) > v_1$. Now v_1, v_2 are determined not by the condition that the agent is indifferent between c^* and some $c' \neq c^*$ but by the fact that the agent is indifferent between c^* and $c^* + dc$.11/

One problem which can arise is that, having computed v_1 and v_2 in (4.6), we may find that c^* does not maximize $\pi_1(c)v_1 + \pi_2(c)v_2 - V(c)$. This can happen because we have used only the agent's first-order condition in (4.6). In this case, the principal is unable to implement c^* and we must set $C(c^*) = \infty$. Another case where we must set $C(c^*) = \infty$ is if the solution (v_1, v_2) to (4.6) $\notin \mathcal{V} \times \mathcal{V}$. In all other cases, we set $C(c^*) = \pi_1(c^*)h(v_1) + \pi_2(c^*)h(v_2)$. (We do not, as in the finite case, ignore c's greater than c^* -- see Remark 5. In the continuous case, it appears to be more convenient to consider all the c's.)

It is in fact quite easy to determine when c^* maximizes $\pi_1(c)v_1 + \pi_2(c)v_2 - V(c)$. For $v_2 = v_1 + (V'(c^*)/\pi_2'(c^*))$, and so

$$\pi_1(c)v_1 + \pi_2(c)v_2 - V(c) = v_1 + \pi_2(c)\frac{V'(c^*)}{\pi_2'(c^*)} - V(c) \quad .$$

This is maximized at $c = c^*$ if and only if

$$\pi_2(c)\,\frac{V'(c^*)}{\pi_2'(c^*)} - V(c) \;\le\; \pi_2(c^*)\,\frac{V'(c^*)}{\pi_2'(c^*)} - V(c^*) \quad \text{for all } c ,$$

i.e. if and only if

(4.7) $$\pi_2(c) - \pi_2(c^*) \;\le\; \frac{\pi_2'(c^*)}{V'(c^*)}\,(V(c) - V(c^*)) \quad \text{for all } c .$$

Condition (4.7) can be expressed in a more illuminating way if we regard π_2 as a function of $V(c)$ rather than of c; that is, write $\tilde{\pi}_2(v) = \pi_2(h(v))$ for $V(\underline{c}) \le v \le V(\bar{c})$. Then $\tilde{\pi}_2'(v) = \pi_2'(c)/V'(c)$ and (4.7) becomes:

(4.8) $$\tilde{\pi}_2(v) - \tilde{\pi}_2(V(c^*)) \le \tilde{\pi}_2'(V(c^*))(v-V(c^*)) \quad \text{for all } V(\underline{c}) \le v \le V(\bar{c}) \quad .$$

(4.8) is certainly satisfied if $\tilde{\pi}_2(v)$ is concave in v. Hence when $\tilde{\pi}_2$ is concave, we need only worry about the first-order conditions (4.5). This is not surprising since $\pi_1(c)v_1 + \pi_2(c)v_2 - V(c) = \tilde{\pi}_1(v)v_1 + \tilde{\pi}_2(v)v_2 - v = v_1 + \tilde{\pi}_2(v)(v_2-v_1)$ and, since $v_2 > v_1$, this is a concave function if $\tilde{\pi}_2(v)$ is.[12/] If $\tilde{\pi}_2(v)$ is not concave, however, then (4.8) may well not hold. In this case, we must set $C(c^*) = \infty$.

We can get a further necessary condition for overall optimality in the continuous case if $C(c)$ is finite in a neighborhood of the optimal c^*. Then since $C(c) = \pi_1(c)h(v_1(c)) + \pi_2(c)h(v_2(c))$,

$$C'(c) = \pi_1'(c)h(v_1) + \pi_2'(c)h(v_2) + \pi_1(c)h'(v_1)\,\frac{dv_1}{dc} + \pi_2(c)h'(v_2)\,\frac{dv_2}{dc} \quad .$$

Using (4.6) and $\pi_1' = -\,\pi_2'$, we can simplify this to get

(4.9) $$C'(c) = \pi_2'(c)(h(v_2) - h(v_1)) - \tilde{\pi}_2''\left[\frac{(V(c))\pi_1(c)\pi_2(c)}{\pi_2'(V(c))^2}\,(h'(v_2) - h'(v_1))\right] V'(c)$$

as long as $\tilde{\pi}_2$ is twice differentiable. A necessary condition for overall optimality is that c^* maximizes $\pi_1(c)q_1 + \pi_2(c)q_2 - C(c)$, i.e.

$$(4.10) \qquad C'(c^*) = \pi_1'(c^*)q_1 + \pi_2'(c^*)q_2 \quad .$$

Note, however, that this condition is not generally sufficient.

Unfortunately, the computational techniques presented above do not appear to generalize in a useful way to the case $n > 2$. In order to compute an optimum when $n > 2$, in the finite action case, it seems that we must, for each $a \in A$, solve the convex problem in (2.2) and then, by inspection, find the $a \in A$ which maximizes $B(a) - C(a)$. If A is infinite, one takes a finite approximation. These steps can be carried out quite straightforwardly on a computer, although the amount of computer time involved when the number of elements of A is large may be considerable.

One case where a considerable simplification can be achieved is when MLRC and CDFC hold. Then the solution to (2.2) has the property that (1) if A is finite, the agent is indifferent only between a^*, the action the principal wants to implement, and a', where a' maximizes $C_{FB}(a)$ subject to $C_{FB}(a) < C_{FB}(a^*)$, i.e. a' is the next most costly action after a^* (see the proof of Proposition 8); (2) if A is convex, then a^* is the unique maximizer of $\sum \pi_i(a)V(I_i) - V(C_{FB}(a))$, so that $(d/da)(\sum \pi_i(a)V(I_i) - V(C_{FB}(a))) = 0$ is a necessary and sufficient condition for the agent to pick a^*. For more on the latter case, see Mirrlees [1979].

One may ask also whether Propositions 9 and 10 generalize to the case $n > 2$. The answer is no. Second-best optimal actions may be inefficient; i.e. there may exist lower cost actions which dominate the optimal action in the sense of first degree stochastic dominance.[13/] Also the addition of actions costlier than the second-best optimal action may make the principal worse off (in Example 1, the principal's expected profits increase if action a_3 becomes unavailable to the agent). Finally, as Shavell [1979a] has noted, the agent may choose a higher cost action under the second-best rather than under the first-best.

5. What Determines How Serious the Incentive Problem Is?

In previous sections, we have studied the properties of an optimal incentive scheme. We turn now to a consideration of the factors which determine the magnitude of L, i.e. the size of the loss to the principal from being unable to observe the agent's action.

One feels intuitively that the worse is the quality of the information about the agent's action that the principal obtains from observing any outcome, the more serious will be the incentive problem. This idea can be formalised as follows. Suppose that we start with an incentive problem in which the agent's action set is A, his utility function is U, his reservation utility is $\bar{U}$, the probability function is π and the vector of outputs is $q = (q_1,\ldots,q_n)$. We denote this incentive problem by $(A, U, \bar{U}, \pi, q)$. Consider the new incentive problem $(A, U, \bar{U}, \pi', q')$ where $\pi'(a) = R\pi(a)$ for all $a \in A$ and R is an (n×n) stochastic matrix (here $\pi(a), \pi'(a)$ are n dimensional column vectors and the columns of R sum to one). Below we show that $C'(a) \geq C(a)$ for all $a \in A$, where unprimed variables refer to the original incentive problem and primed variables to the new incentive problem.

The transformation from $\pi(a)$ to $R\pi(a)$ corresponds to a decrease in informativeness in the sense of Blackwell (see, e.g. Blackwell & Girshick [1954]).[14/] That is, if we think of the actions $a \in A$ as being parameters with respect to which we have a prior probability distribution, then an experimenter who makes deductions about a from observing $q_1,\ldots,q_n$ would prefer to face the function π than the function $R\pi$.

Proposition 11. Consider the two incentive problems $(A, U, \bar{U}, \pi, q)$, $(A, U, \bar{U}, \pi', q')$ and assume that (A1) or (A1'), (A2) and (A3) hold for both. Suppose that $\pi'(a) = R\pi(a)$ for all $a \in A$, where R is an $(n\times n)$ stochastic matrix. Then $C'(a) \geq C(a)$ for all $a \in A$. Furthermore, if V is strictly cave and $R >> 0$,15/ then $C_{FB}(a^*) > \min_{a\in A} C_{FB}(a)$ and $C(a^*) < \infty \Rightarrow C'(a^*) > C(a^*)$.

The proof of Proposition 11 is very simple. Let $(I_1',\ldots,I_n')$ be the cost minimizing way of implementing a in the primed problem. Then in the unprimed problem, the principal can offer the agent the following incentive scheme: for each i, if q_i is the outcome, I will throw an n-sided die where the probability of side j coming up is r_{ji}, the (j,i)th element of R $(j = 1,\ldots,n)$. If side j then comes up, you get I_j.

With this incentive scheme, the agent is as in the primed problem. Therefore he will choose a. Furthermore, the principal's expected costs are the same as in the primed problem. This shows that the principal can implement a at least as cheaply in the unprimed problem as in the primed problem, i.e. $C'(a) \geq C(a)$ for all $a \in A$. The last part of the proposition follows from the fact that the principal can do better by offering the agent the perfectly certain utility level $v_i = \sum_{j=1}^{n} r_{ji}V(I_j')$ if the outcome is q_i rather than the above lottery.

Remark 6. This argument shows that it is never desirable under our assumption for the principal to offer the agent an incentive scheme which makes his payment conditional on a particular outcome a lottery rather than a perfectly certain income. A related result has been established by Holmstrom [1979].

Corollary 1. Make the hypotheses of Proposition 11. If, in addition, q' is such that $q'R = q$, we have $L' \leq L$.

Proof: Obvious since $B'(a) = q'\pi'(a) = q'R\pi(a) = q\pi(a) = B(a)$.

In the case $n = 2$, the transformation $\pi \to \pi' = R\pi$ is easy to interpret. Take any two actions $a_1, a_2 \in A$, and consider the likelihood ratio vector $\left[\frac{\pi_1(a_1)}{\pi_1(a_2)}, \frac{\pi_2(a_1)}{\pi_2(a_2)}\right]$. Assume without loss of generality that $\pi_1(a_1)/\pi_1(a_2) \leq \pi_2(a_1)/\pi_2(a_2)$. Then it is easy to show that

$$(5.1) \qquad \left[\frac{\pi_1'(a_1)}{\pi_1'(a_2)}, \frac{\pi_2'(a_1)}{\pi_2'(a_2)}\right] \in \left[\frac{\pi_1(a_1)}{\pi_1(a_2)}, \frac{\pi_2(a_1)}{\pi_2(a_2)}\right] .$$

In other words, the likelihood ratio vector becomes less variable in some sense when the stochastic transform R is applied. In fact the converse to this is also true: if (5.1) holds, then there exists a stochastic matrix R such that $\pi' = R\pi$ (see Blackwell & Girshick [1954]). When $n > 2$, a simple characterisation of this sort does not seem to exist, however.

One might ask whether a converse to Proposition 11 holds. That is, suppose $C'(a) \geq C(a)$ for all $a \in A$ and all concave utility functions V. Does it follow that $\pi'(a) = R\pi(a)$ for all $a \in A$, for some stochastic R? A converse along these lines can in fact be established when $n = 2$. Whether it holds for $n > 2$, we do not know.

Corollary 1 gives us a simple way of generating worse and worse incentive problems: repeatedly apply stochastic transforms to π. Suppose that we do this using always the same stochastic transform R, where $R >> 0$ and is invertible. That is, we consider a sequence of incentive problems 1,2,..., where in the mth problem $\pi_m(a) = R^{m-1}\pi(a)$ for all $a \in A$, and the gross profit vector q_m satisfies $q_m R^{m-1} = q$ (this has a solution since R is invertible). We know from Corollary 1 that L_m will be increasing in m. The next proposition says that in the limit the loss from not being able to observe the agent reaches its maximal level.

Definition. Let $L^* = \underset{a \in A}{\text{Max}} (B(a) - C_{FB}(a)) - \text{Max} \{B(a') - C_{FB}(a') | a'$ minimizes $C_{FB}(a)$ on $A\}$.

Since $C(a') = C_{FB}(a')$ if a' minimizes $C_{FB}(a)$, L^* is an upper limit on the loss to the principal from being unable to observe the agent. The next proposition shows that as the information q reveals about a gets smaller and smaller, the principal loses control over the agent, i.e. the agent chooses the least-cost action.

Proposition 12. Consider a sequence of incentive problems $(A, U, \bar{U}, \pi_m, q_m)$, $m = 1,2,\ldots$, where $\pi_m(a) = R^{m-1} \pi_1(a)$ for all $a \in A$, $q_m R^{m-1} = q_1$ for some invertible stochastic matrix $R >> 0$. Assume (A1) or (A1'), (A2) and $\pi_{1i}(a) > 0$ for all $i=1,\ldots,n$ and $a \in A$. Then if V is not a linear function, $\lim_{m\to\infty} L_m = L^*$.

Proof It suffices to show that $\lim_{m\to\infty} C(a^*) = \infty$ for all a^* with $C_{FB}(a^*) > \min_{a \in A} C_{FB}(a)$. Suppose not for some such a^*. Let $(I_{m1},\ldots,I_{mn})$ be the cost minimising way of implementing a^* in problem m. Then $\sum_i \pi_{mi}(a^*) I_{mi}$ and $\sum_i \pi_{mi}(a^*)V(I_{mi})$ are both bounded in m. It follows from Bertsekas [1974] that the (I_{mi}) are bounded. Hence w.l.o.g we may assume $I_{mi} \to I_i$ for each i. It is easy to show that, since R is a strictly positive stochastic matrix, $\lim_{m\to\infty} R^{m-1} = R^*$ where R^* has the property that all of its columns are the same. Therefore $\lim_{m\to\infty} \pi_m(a) = R^*\pi_1(a) = \bar{\pi}$ is independent of a. But this means that $\lim_{m\to\infty} \sum_i \pi_{mi}(a^*)V(I_{mi}) = \sum_i \bar{\pi}_i V(I_i) = \lim_{m\to\infty} \sum_i \pi_{mi}(a)V(I_{mi})$ for all $a \in A$. Hence the agent will prefer actions a with $C_{FB}(a) < C_{FB}(a^*)$ to a^*. This contradicts the assumption that the incentive scheme implements a^*. Q.E.D.

We turn now to a consideration of another factor which influences L: the agent's degree of risk aversion. Since no incentive problem arises when the agent is risk neutral, but an incentive problem does arise when the agent is risk averse, one is led to ask whether L increases as the agent becomes more risk averse. One difficulty in answering this question is the following. The way one makes the agent more risk averse is to replace his utility function $U(I,a)$ by $H(U(I,a))$ where H is a real-valued, increasing, concave function. However, this transformation will in many cases destroy the additive or multiplicative separability upon which our analysis is based. To get round this difficulty, we will not attempt to deal with this question in general, but will confine our attention to the case where A is a subset of the real line and $V(I) = -e^{-kI}$, $G(a) = e^{ka}$, i.e. the agent's utility function is $U(a,I) = -e^{-k(I-a)}$, where $k > 0$. Assume also that $\bar{U} = -e^{-k\alpha}$. An increase in risk aversion can then be represented simply by an increase in k.

Note that if the agent's utility function is $-e^{-k(I-a)}$ and $\bar{U} = -e^{-k\alpha}$, then $C_{FB}(a) = a + \alpha$, which is independent of k. Hence first best profits are also independent of k.

Proposition 13. Consider the incentive problem $(A, U, \bar{U}, \pi, q)$ where A is a subset of the real line, $U(a,I) = -e^{k(I-a)}$, $\bar{U} = -e^{k\alpha}$, and $k > 0$. Assume (A3). Write the loss from being unable to observe the agent as $L(k)$. Then

$\lim_{k \to 0} L(k) = 0$, $\lim_{k \to \infty} L(k) = L^*$.

Proof To show that $\lim_{k\to\infty} L(k) = L^*$, it suffices to show that $\lim_{k\to\infty} C(a^*,k) = \infty$ for all a^* with $C_{FB}(a^*) > \min_{a \in A} C_{FB}(a)$. Suppose not for some such a^*, and let $C_{FB}(a) < C_{FB}(a^*)$. Then if $(I_1,\ldots,I_n)$ implements a^*, we must have

$$-\left(\sum_i \pi_i(a^*)e^{-kI_i}\right) e^{ka^*} \geq -\left(\sum_i \pi_i(a)e^{-kI_i}\right) e^{ka}$$

$(I_1,\ldots,I_n$ of course depend on k.)

Therefore,

$$(5.2) \qquad e^{k(a^*-a)} \leq \sum_i \pi_i(a)e^{-kI_i} \Big/ \sum_i \pi_i(a)e^{-kI_i} .$$

Now let $k \to \infty$. The LHS of (5.2) $\to \infty$. Therefore so must the RHS. We may assume w.l.o.g., however, that $I_1 = \min_i I_i$. Then

$$\frac{\sum_i \pi_i(a)e^{-kI_i}}{\sum_i \pi_i(a^*)e^{-kI_i}} = \frac{\sum_i \pi_i(a)e^{k(I_1-I_i)}}{\sum_i \pi_i(a^*)e^{k(I_1-I_i)}} ,$$

which is bounded since the demoninator $\geq \pi_1(a^*)$. Contradiction.

We show now that $\lim_{k \to o} L(k) = 0$. Let $I_i = q_i - F$. Then the agent maximises

$$(5.3) \qquad E(-e^{-k(I-a)}) = -E(1-k(I-a) + \frac{k^2}{2}(I-a)^2 + \ldots\ldots)$$

$$= -1 + k(\Sigma\pi_i(a)q_i - F - a) - \frac{k^2}{2} E(I-a)^2 + \ldots\ldots$$

It follows that the agent maximises

$$(\Sigma\pi_i(a)q_i - F - a) - \frac{k}{2} E(I-a)^2 + \ldots\ldots$$

which means that in the limit $k \to 0$ the agent maximises $B(a) - C_{FB}(a)$, i.e. chooses a first best action. Furthermore, setting (5.3) equal to $-e^{-k\alpha} = -1+k\alpha + \ldots\ldots$, we see that in the limit $k \to 0$,

$$\max_{a \in A} (\sum_i \pi_i(a)q_i - a) - F = \alpha,$$

so that the the principal's expected profit equals $F = \max_{a \in A} (\sum_i \pi_i q_i - a) - \alpha = \max_{a \in A} (B(a) - C_{FB}(a)) =$ first-best profit.

Q.E.D.

Proposition 13 tells us about the behaviour of $L(k)$ for extreme values of k. It would be interesting to know whether $L(k)$ is increasing in k. We do not know the answer to that question except for the case $n = 2$, A finite.

Proposition 14. Make the same hypotheses as in Proposition 13. Assume in addition that $n = 2$ and A is finite. Then $L(k)$ is increasing in k.

Proof It suffices to show that $C(a,k)$ is increasing locally in k for each $a \in A$ whenever $C(a,k)$ is finite. Let $\tilde{k} = \lambda k, \lambda \geq 1$. Assume that (I_1, I_2) is the cost minimising way of implementing a given $\tilde{k}$. Then, by the results of Section 4, e.g. eq. (4.2),

$$\pi_1 w_1 + \pi_2 w_2 = \frac{1}{e^{\tilde{k}(a+\alpha)}} ,$$

(5.4)

$$\pi_1' w_1 + \pi_2' w_2 = \frac{1}{e^{\tilde{k}(a'+\alpha)}} ,$$

where $w_1 = e^{-\tilde{k}I_1}$, $w_2 = e^{-\tilde{k}I_2}$, $\pi_1 = \pi_1(a)$, $\pi_2 = \pi_2(a)$, $\pi_1' = \pi_1(a')$, $\pi_2' = \pi_2(a')$, $a' \in A$, $a' < a$. Furthermore we can pick a' so that a' is independent of $\tilde{k}$ for λ close to 1.

(5.4) determines w_1 and w_2 for each value of $\tilde{k}$. The cost of implementing a, $C(a,\tilde{k})$, is then given by

$$(5.5) \qquad C(a,\tilde{k}) = \pi_1 I_1 + \pi_2 I_2 = -\frac{1}{\tilde{k}} (\pi_1 \log w_1 + \pi_2 \log w_2).$$

Differentiating (5.5) with respect to λ we get

$$(5.6) \qquad \left.\frac{\partial C(a,\lambda k)}{\partial \lambda}\right|_{\lambda=1} = \frac{1}{k} (\pi_1 \log w_1 + \pi_2 \log w_2 - \frac{\pi_1}{w_1} \frac{dw_1}{d\lambda} - \frac{\pi_2}{w_2} \frac{dw_2}{d\lambda}).$$

Set $x = e^{-k(a+\alpha)}$, $y = e^{-k(a'+\alpha)}$ in (5.4). Then $e^{-\tilde{k}(a+\alpha)} = x^{\lambda}$,

$e^{-\tilde{k}(a'+\alpha)} = y^{\lambda}$. Hence

$$(5.7)\qquad \begin{aligned} \pi_1 \frac{dw_1}{d\lambda} + \pi_2 \frac{dw_2}{d\lambda} &= x \log x \\ \pi_1' \frac{dw_1}{d\lambda} + \pi_2' \frac{dw_2}{d\lambda} &= y \log y \end{aligned} ,$$

where derivatives are evaluated at $\lambda = 1$. Solving (5.4), (5.7) yields

$$w_1 = \frac{\pi_2' x - \pi_2 y}{\pi_1 \pi_2' - \pi_1' \pi_2} = \frac{\pi_2' x - \pi_2 y}{\pi_2' - \pi_2} ,$$

$$\frac{dw_1}{d\lambda} = \frac{\pi_2' x \log x - \pi_2 y \log y}{\pi_2' - \pi_2} .$$

It follows that $\log w_1 \geq (1/w_1)(dw_1/d\lambda)$. For

$$(5.8)\qquad w_1 \log w_1 - \frac{dw_1}{d\lambda} = \frac{\pi_2' x - \pi_2 y}{\pi_2' - \pi_2} \log \frac{\pi_2' x - \pi_2 y}{\pi_2' - \pi_2} - \frac{\pi_2' x \log x - \pi_2 y \log y}{\pi_2' - \pi_2}$$

$$= \frac{1}{\pi_2' - \pi_2} \left[(\alpha x - \beta y)\log \frac{\alpha x - \beta y}{\alpha - \beta} - \alpha x \log x - \beta y \log y\right] ,$$

where $\alpha = \pi_2'$, $\beta = \pi_2$. However, the RHS of (5.8) ≥ 0 by Lemma 3 below. The same argument shows that $\log w_2 \geq (1/w_2)(dw_2/d\lambda)$. It follows from (5.6) that $(\partial C/\partial \lambda) \geq 0$, i.e. C is increasing locally in $\tilde{k}$.

Lemma 3. Assume $\alpha, \beta, x, y > 0$. Then if $\alpha > \beta$ and $\alpha x > \beta y$, $\alpha x \log x - \beta y \log y < (\alpha x - \beta y) \log \left(\frac{\alpha x - \beta y}{\alpha - \beta}\right)$. On the other hand, if $\alpha < \beta$ and $\alpha x < \beta y$, $\alpha x \log x - \beta y \log y > (\alpha x - \beta y) \log \left(\frac{\alpha x - \beta y}{\alpha - \beta}\right)$.

Proof: Since $z \log z$ is a convex function,

$$\frac{\beta}{\alpha} (y \log y) + \left(\frac{\alpha - \beta}{\alpha}\right) \left(\frac{\alpha x - \beta y}{\alpha - \beta} \log \frac{\alpha x - \beta y}{\alpha - \beta}\right) \geq x \log x .$$

This proves the first part. The second part follows similarly. Q.E.D.

Remark 7. Propositions 13 and 14 tell us how the principal's welfare varies with k. It is also interesting to ask how the shape of the optimal incentive scheme depends on k. Unfortunately, even in the case $n = 2$, very little can be said. In this case, the incentive scheme is characterised by the agent's share s. It is not difficult to construct examples showing that an increase in the agent's risk aversion may increase the optimal value of s, or may decrease it.

We conclude this section by considering how L depends on the agent's incremental costs. Suppose that we write the agent's cost function as $G_\lambda(a) = \alpha + \lambda F(a)$, where $\lambda > 0$. Then, when λ is small, one feels that L will be small since the agent does not require much of a reward to work hard. The fact that $\lim_{\lambda \to 0} L(\lambda) = 0$ has in fact been established by Shavell [1979b]. We prove a somewhat stronger result for the case of additive separability.

Proposition 15. Consider the incentive problem $(A, \pi., V, G_\lambda, q, \bar{U})$, where $G_\lambda(a) = \alpha + \lambda F(a)$ for all $a \in A$, $\lambda > 0$. Assume that (A1), (A2) and (A3) hold for this problem. Assume also that (1) A is an interval of the real line and $\pi(a)$ is twice differentiable and $F(a)$ is differentiable in the interior of A; (2) Every maximiser of $B(a)$ lies in the interior of A. Then $\lim_{\lambda \to 0} (L(\lambda)/\lambda) = 0$.

Proof Suppose first that $B(a)$ has a unique maximiser a^*. Consider the incentive problem with $\lambda = 1$. Then there are a's arbitrarily close to a^* for which $C(a)$ is finite. For let the principal set $v_i = rq_i - k$ where k is chosen so that $v_i \in \mathcal{V}$ for all i. Then the agent will maximize $\sum \pi_i(a)U(a, I_i)$, i.e. $\sum \pi_i(a)q_i - F(a)/r$. By letting $r \to \infty$, we can get the agent to choose an action arbitrarily close to a^*. For such an action, $C(a)$ will be finite.

Consider now an a arbitrarily close to a^*. Let $(v_1,\ldots,v_n)$ be the cost minimizing way of implementing a when $\lambda = 1$. Then it is clear from (2.2) that $(\lambda v_1 + \beta,\ldots,\lambda v_n + \beta)$ will implement a for $\lambda \neq 1$, where

$$\lambda(\sum \pi_i(a)v_i - F(a)) - \alpha + \beta = \bar{U} \quad .$$

It follows that

$$L(\lambda) \leq \sum \pi_i(\hat{a})q_i - h(\bar{U} + \alpha + \lambda F(\hat{a})) - (\sum \pi_i(a)q_i - \sum \pi_i(a)h(\lambda v_i + \beta)) \ ,$$

where $\hat{a}$ maximizes $\sum \pi_i(a)q_i - h(\bar{U} + \alpha + \lambda F(a))$, i.e. $\hat{a}$ is the first-best action in problem λ.

Therefore,

$$\frac{L(\lambda)}{\lambda} \leq [\tfrac{1}{\lambda}\{\Sigma\pi_i(a^*)q_i - h(\bar{U} + \alpha + \lambda F(a^*)) - (\Sigma\pi_i(a)q_i - \Sigma\pi_i(a)h(\lambda v_i + \beta))\}]$$

$$+ [\tfrac{1}{\lambda}\{\Sigma\ \pi_i(\hat{a})q_i - h(\bar{U} + \alpha + \lambda F(\hat{a})) - \Sigma\pi_i(a^*)q_i + h(\bar{U} + \alpha + \lambda F(a^*)\}] \ .$$

It is easy to show by expanding the first-order conditions $\frac{d}{da}\sum \pi_i(a^*)q_i = 0$, $\frac{d}{da}(\Sigma\pi_i(\hat{a})q_i - h(\bar{U} + \alpha + \lambda F(\hat{a})) = 0$ as a Taylor's series that the second square bracket $\to 0$ as $\lambda \to 0$. To see that the first square bracket $\to 0$, note that, since a is arbitrary, we can make a converge to a^* as fast as we like. Therefore we need only show that

(5.9) $$\lim_{\lambda\to 0} \frac{1}{\lambda}\ (\sum \pi_i(a)h(\lambda v_i + \beta) - h(\bar{U} + \alpha + \lambda F(a^*))) = 0 \ .$$

But

$$\sum \pi_i(a)[h(\lambda v_i + \beta) - h(\bar{U} + \alpha + \lambda F(a^*))]$$

$$= \Sigma\pi_i(a)[h(\lambda v_i + \bar{U} + \alpha - \lambda\Sigma\pi_j(a)v_j + \lambda F(a)) - h(\bar{U} + \alpha + \lambda F(a^*))]$$

$$= \Sigma\pi_i(a)[h(\bar{U} + \alpha) + h'(\bar{U} + \alpha)(\lambda v_i - \lambda\Sigma\pi_j(a)v_j + \lambda F(a)) + \ldots\ldots$$

$$- h(\bar{U} + \alpha) - h'(\bar{U} + \alpha)(\lambda F(a^*)) + \ldots.]$$

$$= h'(\bar{U} + \alpha)(\lambda F(a) - \lambda F(a^*)) + \ldots.$$

(5.9) follows.

ı) does not have a unique maximiser, we choose a to be arbitrarily

› set of maximisers. The same argument applies.

Q.E.D.

ɔears that a similar result can be established for the multiplicatively

ıse. Since the proof is more complicated, however, we will not pursue

here.

ınce the marginal product of labour of the agent - that is the

expected profit resulting from an extra pound of expenditure by the

roportional to $\frac{1}{\lambda}$, Proposition 15 can be interpreted as saying that

loss L is of the same order of magnitude as the reciprocal of the

ginal product of labour.[16/]

y

pose of this paper has been to develop a method for analyzing the

ent problem in the case where the agent's utility function is separ-

on and reward. Our method consists of breaking up the principal's p

› a computation of the costs and benefits accruing to the principal

nt takes a particular action. We have used this method to establish

results about the structure of the optimal incentive scheme and about the determinants of the welfare loss resulting from the principal's inability to observe the agent's action. We have shown that it is never optimal for the incentive scheme to be such that the principal's and agent's payoff are negatively related over the whole output range, although such a relationship may be optimal over part of the range. We have found sufficient conditions for the incentive scheme to be monotonic, progressive and regressive. We have shown that a decrease in the quality of the principal's information in the sense of Blackwell increases welfare loss. When there are only two outcomes, welfare loss also increases when the agent becomes more risk averse. Finally, we have discussed how our techniques can be used to compute optimal incentive schemes in particular areas.

While we have talked throughout about "the" principal-agent problem, we have in fact been considering the simplest of a number of such problems. More complicated principal-agent problems arise when not only is the principal unable to monitor the agent, but also the agent possesses information about his environment, i.e. about A, π or $U(a,I)$, which the principal does not. Such problems share a number of features with the preference revalation problems studied in the recent incentive compatibility literature; see, for example, the Review of Economic Studies Symposium [1979]. A start has been made in the analysis of such problems by Harris and Raviv [1979], Holmstrom [1979], and Mirrlees [1979]. It will be interesting to see whether the techniques presented here will also be useful in the solution of these more complicated principal-agent problems.

FOOTNOTES

1/ Recent discussions of the principal-agent problem include Harris and Raviv [1979], Holmstrom [1979], Mirrlees [1975,1976,1979], Radner [1980], Ross [1973], Rubinstein and Yaari [1979], Shavell [1979a,1979b], Spence and Zeckhauser [1971], Stiglitz [1974] and Zeckhauser [1970].

2/ The assumption that the principal cannot monitor the agent's actions at all may in some cases be rather extreme. For a discussion of the implications of the existence of imperfect monitoring opportunities, see Harris and Raviv [1979], Holmstrom [1979] and Shavell [1979a,1979b]. See also Remark 3 in Section 2.

3/ This distinguishes our study from the literature on incentive compatibility, see e.g. the recent Review of Economic Studies symposium [1979]. The incentive compatibility literature has been concerned with incentive problems arising from differences in information between individuals rather than with those arising from monitoring problems.

4/ Among other things, Proposition 4 shows that it is not optimal to have $q_1 - I_1 = q_2 - I_2 = \dots = q_n - I_n$. This result has also been established by Shavell [1979b] under stronger assumptions.

5/ We use the term decreasing (resp. strictly decreasing) to mean non-increasing (resp. decreasing).

6/ See Mirrlees [1976] or Holmstrom [1979]. Milgrom [1979] has shown that MLRC, as stated here, implies the differential version of the monotone likelihood condition which is to be found in Mirrlees [1976] or Holstrom [1979].

7/ The function V violates (2) of (A1), but this is unimportant for the example.

8/ To prove the converse, define $a \precsim a'$ if $\pi_i(a')/\pi_i(a)$ is increasing in i. (3.11) implies that $\precsim$ is a complete pre-ordering on A. Furthermore, $\precsim$ is continuous. Since A is compact, there exist $\underline{a}$, $\bar{a} \in A$ such that $\underline{a} \precsim a \precsim \bar{a}$ for all $a \in A$. Given $a \in A$, consider $\lambda(\pi_i(\bar{a})/\pi_i(a)) + (1-\lambda)(\pi_i(\underline{a})/\pi_i(a))$. When $\lambda = 1$, this is increasing in i, and, when $\lambda = 0$, it is decreasing in i. Furthermore, (3.11) implies that it is monotonic in i for all $0 < \lambda < 1$. It follows by continuity that it is independent of i for some $0 < \lambda < 1$.

9/ Another case which ensures monotonicity of the optimal incentive scheme is when the firm's profits can be freely disposed of by the agent; i.e. if the agent can always make a better outcome look like a worse outcome by reducing the firm's profits after the outcome has occurred (possibly even by paying the profits to himself). This case can be analyzed by adding the (linear) constraints $v_1 \leq v_2 \leq \dots \leq v_n$ to the problem (2.2).

10/ Shavell [1979a] also proves that $s \geq 0$ when $n = 2$, but under stronger assumptions.

11/ Mirrlees [1975] computes an example for the case $n = 2$ using equations like (4.6).

12/ The importance of the concavity of the probability functions π_1, π_2 for computational purposes has been noted by Mirrlees [1979]. See also the end of this section.

13/ Let $A = \{a_1, a_2, a_3\}$, $n = 3$. Assume $C_{FB}(a_1) < C_{FB}(a_2) < C_{FB}(a_3)$, and that $\pi(a_1) = (3/4, 1/8, 1/8)$, $\pi(a_2) = (1/3, 1/3, 1/3)$, $\pi(a_3) = (\frac{1}{2}, \frac{1}{2}, 0)$. ((A3) is violated, but this is unimportant.) Then $C(a_1) = C_{FB}(a_1)$ since a_1 is the least cost action, and $C(a_3) = C_{FB}(a_3)$ since a_3 can be implemented by setting $I_1 = I_2$, $I_3 = -\infty$. However, $C(a_2) > C_{FB}(a_2)$ and, in fact, if the agent is very risk averse, $C(a_2)$ will be so big that it is profitable for the principal to implement a_3 rather than a_2 (the effect of risk aversion on $C(a)$ is discussed in Section 5). This is in spite of the fact that a_3 is inefficient relative to a_2.

14/ The possibility of using Blackwell's notion of informativeness to characterize the seriousness of an incentive problem was suggested, but not explored, by Holmstrom [1979].

15/ We use this notation to mean that every element of R is strictly positive.

16/ Proposition 15 does not generalize to the case where the principal is risk averse. In fact, $L(0)$ will generally be positive in this case. The reason is that it is socially optimal for both individuals to bear some risk when they are both risk averse, and there may be a conflict over the type of lottery they should share in even when there is no disutility of effort for the agent. On this, see Ross [1973] and Wilson [1968].

REFERENCES

K.J. Arrow [1971], "Insurance, Risk and Resource Allocation," Essays in the Theory of Risk Bearing, Chicago: Markham.

D. Bertsekas [1974], "Necessary and Sufficient Conditions for Existence of an Optimal Portfolio," Journal of Economic Theory, 8, 235-247.

D. Blackwell & M.A. Girshick [1954], "Theory of Games and Statistical Decisions," John Wiley & Sons, Inc.

K. Borch [1968], The Economics of Uncertainty, Princeton University Press.

G.H. Hardy, J.E. Littlewood & G. Polya [1952], "Inequalities," Cambridge University Press.

M. Harris & A. Raviv [1979], "Optimal Incentive Contracts with Imperfect Information," Journal of Economic Theory, 20, 231-259.

B. Holmstrom [1979], "Moral Hazard and Observability," Bell Journal of Economics, Spring, 74-91.

P.R. Milgrom [1979], "Good News and Bad News: Representation Theorems and Applications," Discussion Paper no. 407, Northwestern University, Illinois, mimeo.

J.A. Mirrlees [1975], "The Theory of Moral Hazard and Unobservable Behaviour -- Part I," Nuffield College, Oxford, mimeo.

J.A. Mirrlees [1976], "The Optimal Structure of Incentives and Authority Within An Organization," Bell Journal of Economics, Spring, 105-131.

J.A. Mirrlees [1979], "The Implications of Moral Hazard for Optimal Insurance," Seminar given at Conference held in honour of Karl Borch, Bergen, Norway, mimeo.

M. Pauly [1968], "The Economics of Moral Hazard: Comment," American Economic Review, LVIII, 531-536.

R. Radner [1980], "Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship," mimeo, Bell Laboratories.

Review of Economic Studies Symposium on Incentive Compatibility, April 1979.

S. Ross [1973], "The Economic Theory of Agency: The Principal's Problem," American Economic Review, 63, 134-139.

A. Rubinstein & M. Yaari, Seminar given at Conference held in honour of Karl Borch, Bergen, Norway, 1979.

S. Shavell [1979a], "On Moral Hazard and Insurance," Quarterly Journal of Economics, XCIII, 4, 541-562.

S. Shavell [1979b], "Risk Sharing and Incentives in the Principal and Agent Relationship," Bell Journal of Economics, Spring, 55-73.

M. Spence & R. Zeckhauser [1971], "Insurance, Information, and Individual Action," American Economic Review, 61, 380-387.

J.E. Stiglitz [1974], "Incentives and Risk Sharing in Sharecropping," Review of Economic Studies, LXI, 219-256.

R. Wilson [1968], "The Theory of Syndicates," Econometrica, 36, 119-132.

R. Zeckhauser [1980], "Medical Insurance: A Case Study of the Trade-Off Between Risk Spreading and Appropriate Incentives," Journal of Economic Theory, 2, 10-26.