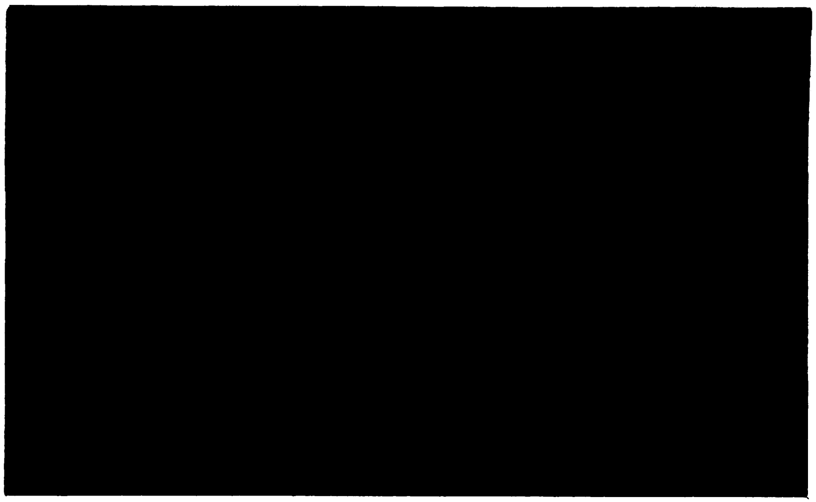
C-27

Center for Research on Economic and Social Theory
Research Seminar in Quantitative Economics

Discussion Paper

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109



Incentive Contracting with Asymmetry of Precontractual Beliefs

by

David Sappington*
Department of Economics
University of Michigan
Ann Arbor, Michigan 48109

C-27

The interaction between a risk-neutral principal and agent is explored in which the agent's information is better than that of the principal both before and after a contract is agreed upon. It is shown that, in contrast to the case in which precontractual information is symmetric, the final contract between the principal and the agent will usually not be ex post Pareto efficient. The properties of the optimal (set of) contract(s) is derived in detail, and it is shown that the qualitative such properties may vary depending upon whether the random state of nature follows a continuous or a discrete distribution.

Revised November 1981

1. INTRODUCTION

The "principal-agent" literature explores the interaction between two economic parties, a principal and an agent. The former owns a productive technology which requires as an input the effort of the latter. The salient feature of most models in this literature is that the principal can observe neither the level of effort expended by the agent in production nor the realization of a random state of nature, θ, which affects the agent's productivity. Thus, any (non-renewable, single-period) contract between the two parties must specify payments to the agent as a function of the only observable variable, the value of the output which ensues from production.

One striking result in the literature holds that if a principal and risk-neutral agent initially share the same beliefs about the distribution of θ, and if the agent (alone) can observe the actual realization of θ before choosing his level of effort, then, despite the principal's limited information he can (and, in order to maximize his expected utility, will) design a contract which induces the agent to act so as to realize an outcome which is Pareto efficient in the particular state of nature that does occur, whatever that state might be. (See, for example, Harris and Raviv [1979], Holmstrom [1979], and Shavell [1979].)

It is the purpose of this research to demonstrate that this result depends crucially upon the assumption of precontractual symmetry of beliefs. To this end, a broad class of informational environments are presented in which the principal's optimal strategy will induce a risk-neutral agent who possesses better information (in the sense of Blackwell [1951]) than the principal to act so as to realize an inefficient outcome in some states of nature. In other words, the "optimal" contract (from the point of view of the principal) will not be an efficient one.

The analysis begins in Section 2 with a formal statement of the model that is analyzed here. Special emphasis is awarded the nature of the information asymmetry between principal and risk-neutral agent. Some remarks concerning the relation of this model to others in the literature are also offered in this section. Then in Section 3 it is proved that the final contract between principal and agent may not be efficient, whether the random state of nature follows a continuous or discrete distribution. However, a specification of the conditions which are necessary and sufficient to guarantee that the two parties do agree to an efficient contract is also offered in this section. The conditions are shown to differ systematically depending upon whether θ follows a continuous or a discrete distribution.

In the following section, stochastic dominance of distributions is defined, and its economic importance discussed. Furthermore, it is shown that when only two distributions of θ are possible and they stand in a relation of stochastic dominance, inefficient contracts between principal and risk-neutral agent will be observed almost always. An intuitive explanation of all of these findings is offered in Section 5 before conclusions are drawn in Section 6.

2. DESCRIPTION OF THE MODEL

The interaction between principal and agent to be considered here is as follows. The principal owns a productive technology that requires as an input the effort, a, of a risk-neutral agent. The functional form of this technology, which is known to both principal and agent, is $x = X(a,\theta)$, where x is the value of output produced and θ is the realization of a random state of nature. At no time can the principal observe either a or θ.

The random state of nature may follow any one of D possible distributions. When a contract between principal and agent must be signed, the principal's knowledge of the actual distribution is characterized by a non-degenerate prior defined over these D distributions. The agent, however, has learned of the actual distribution by this time. Thus, the agent's information is better in the sense of Blackwell [1951] and the principal is aware of this fact. Consequently, the work of Harris and Townsend [1981] justifies the conclusion that the optimal strategy for the principal is to design at most D distinct contracts (that specify dollar payments, s, to the agent as a function of the value of output produced) from which the agent is permitted to make a binding choice. The rationale behind this strategy is that if the contracts are designed appropriately, the agent can be induced to use his private information under some distributions to select contracts that the principal would prefer (if he also knew the actual distribution) to the single contract that the principal will design in the absence of any better information about the actual distribution of θ.

After the agent chooses a contract, he observes the actual realization of θ, and then decides how much effort to supply in the production process. His level of effort will be selected to maximize his utility function, defined as $U^A(s,a) = s - v(a)$ where $v'(a) > 0$. Given knowledge of the technology and the realization of θ, the agent's choice of an effort level is equivalent to the choice of a value for x. Hence, the agent's utility function can be rewritten as $U^A(x,s|\theta) = s - w(x,\theta)$ where $w(x,\theta)$ is the dollar value of the disutility to the agent of producing x in state θ.

It is assumed that for all non-negative values of x and θ, $w_\theta(x,\theta) \leqq 0$, $w_x(x,\theta) \geqq 0$, $w_{xx}(x,\theta) \geqq 0$, and $w_{x\theta}(x,\theta) \leqq 0$, where subscripts indicate partial derivatives (which are assumed to be continuous in both x and θ). These inequalities are assumed to hold as strict inequalities for all $x > 0$. It is also assumed that $w_x(0,\theta_1) < 1$ where $\theta_1 > 0$ is the smallest value of θ in each distribution. This lattermost assumption ensures that even in the least productive state, the agent can produce some positive level of output while incurring a personal disutility which is strictly less than the value of that output. The other assumptions suggest that θ can be interpreted as a productivity parameter, where higher values of θ correspond to states in which the agent is more productive and in which his disutility from additional effort increases less rapidly.

The dollar value of any output produced by the agent is defined by the principal's utility function, which is represented as $U^P(\cdot) = x - s$. Hence, the principal, too, is assumed to be risk neutral in order to abstract from the risk-sharing considerations often explored in the principal-agent literature.

Given the nature of the informational asymmetry, the principal's problem can be stated formally once the distributions of θ are characterized. If each distribution is continuous and has strictly positive support on the interval $[\theta_1,\theta_n]$, the principal's problem, (PA-C), is to

$$\underset{s(x)}{\text{Maximize}} \quad \sum_{d=1}^{D} \phi^d \int [x^d(\theta) - s^d(x^d(\theta))]f^d(\theta)d\theta$$

subject to: (PA-C)

(IR) $$\int [s^d(x^d(\theta)) - w(x^d(\theta),\theta)]f^d(\theta)d\theta \geqq 0$$
$$\forall d = 1,\dots,D$$

(SSB) $$\int [s^d(x^d(\theta)) - w(x^d(\theta),\theta)]f^d(\theta)d\theta \geqq \int [s^r(x^r(\theta)) - w(x^r(\theta),\theta)]f^d(\theta)d$$
$$\forall r = 1,\dots,D$$

(SSW) $$x^d(\theta) = \underset{x'}{\text{argmax}}\ s^d(x') - w(x',\theta)$$

where ϕ^d = probability that the actual distribution is $f^d(\theta)$,

$x^d(\theta)$ = value of output produced by the agent in state θ under contract $s^d(x)$, and

where all integrals are defined over the interval $[\theta_1,\theta_n]$.

If, on the other hand, θ can take on any one of n possible values $\theta_1 < \dots < \theta_n$ under each of D possible distributions, the principal's problem, (PA-D), is to

Maximize$_{x,s}$ $$\sum_{d=1}^{D} \phi^d \sum_{i=1}^{n} p_i^d [x_i^d - s_i^d]$$

subject to: (PA-D)

(IR) $$\sum_{i=1}^{n} p_i^d [s_i^d - w(x_i^d,\theta_i)] \geqq 0 \qquad \forall d = 1,\dots,D$$

(SSB) $$\sum_{i=1}^{n} p_i^d [s_i^d - w(x_i^d,\theta_i)] \geqq \sum_{i=1}^{n} p_i^d [s_i^r - w(x_i^r,\theta_i)]$$
$$\forall r = 1,\dots,D$$

(SSW) $$s_i^d - w(x_i^d,\theta_i) \geqq s_j^d - w(x_j^d,\theta_i)$$
$$\forall i,j = 1,\dots,n \text{ for each } d = 1,\dots,D$$

where p_i^d = probability that $\theta = \theta_i$ under distribution d,

x_i^d = value of output produced by the agent in state θ_i under contract $s^d(x)$, and

s_i^d = payment to the agent for producing x_i^d under contract $s^d(x)$.

The individual rationality (IR) constraints guarantee that any contract selected by the agent provides him with a level of expected utility that (weakly) exceeds his (known) reservation level, which is normalized at zero. The (SSB) constraints guarantee self-selection between contracts, i.e., they are sufficient to ensure that the agent selects contract $s^d(x)$ when he observes the d-th distribution of θ. The (SSW) constraints ensure self-selection within any contract $s^d(x)$, i.e., they guarantee that the agent will produce $x^d(\theta)$ in state θ.[1]

Finally, a definition of the following concepts is important for the ensuing analysis:

Definition: An efficient contract is one that induces the agent to produce the value of output $x^*(\theta)$ that is (ex post Pareto) efficient in the particular state that is realized, whatever that state might be.[2]

Definition: A pooling contract is one in which the agent is induced to produce the same value of output in more than one state.

Definition: A non-revealing contract is one that, when selected by the agent, reveals no information about the true distribution of θ.

Definition: A set of contracts is revealing if the principal can infer the true distribution of θ by the agent's choice of a contract from this set.

Before the solutions to (PA-C) and (PA-D) are explored in Section 3, these models are briefly compared to others in the principal-agent literature that are concerned with information asymmetry. Harris and Townsend [1981] present a fairly general formulation of the interaction among agents who are asymmetrically informed about various aspects of their environment. Although their analysis of the principal-agent model assumes that the principal knows (only) the true distribution of θ while the agent knows the actual realization of θ before contracting, their results concerning the optimality of "direct mechanisms" are sufficiently general to permit the conclusion that, in the model considered in this paper, it is indeed (Pareto-) optimal for the agent to choose from among a set of contracts offered by the principal.

The form of information asymmetry examined by Holmstrom [1979] is post-contractual in nature, since, when a contract is agreed upon in his model, the principal and agent share symmetric beliefs about the distribution of θ. It is only after the contract is signed (but before the agent chooses an effort level) that the agent (only) observes a signal which is correlated with the true state of nature that will ultimately be realized. Holmstrom's model, then, is similar to those of Harris and Raviv [1979] and Sappington [1982]. None of these models incorporates precontractual information asymmetry.

The works of Green [1979] and Green and Stokey [1980a, b], though, do consider the effects of precontractual information asymmetry in some detail. In their models, the principal and agent initially share symmetric beliefs concerning the distribution of θ, but the agent (only) observes a signal which provides him with better knowledge (in the sense of Blackwell [1951]) of the true distribution before a contract is signed. Aware of this fact, the principal announces what (possibly random) action he will take for any signal that the agent claims to have observed. After the agent's report is received and before the state of nature is observed, the principal undertakes his promised action.

The important differences between the model considered here and those of Green [1979] and Green and Stokey [1980b] warrant brief elaboration. First, the latter models are more general in that they permit the principal to pursue a mixed strategy, while the principal is assumed to follow a pure strategy in the present analysis. It can be shown though, that the major conclusions of this research are unaltered if the principal pursues a mixed strategy. (For details, see Sappington [1980b]).

Second, in the model analyzed in this paper, the principal's choice of payment schedules (contracts) is unrestricted, since any arbitrary function of x may serve as a contract between principal and agent. In the other models, however, the principal's choice of actions must be from among a finite set of actions which is specified exogenously. Consequently, the sensitivity of the final equilibria to the presumed form of the principal's action space (which is noted and explored by Green [1979] and Green and Stokey [1980a, b]) is avoided in this paper.

One other difference between the analyses is the fact that the model considered here explicitly incorporates both precontractual and post-contractual information asymmetry. The latter asymmetry arises because the agent (but not the principal) is able to observe the state of nature before an action is chosen. Hence, questions regarding the ex post efficiency of outcomes are readily addressed in the present model.

Finally, it should be noted that the example described by Sappington [1980a] is a special case of (PA-C). The findings in Section 3 extend and generate his results, and indicate that additional analytic and conceptual difficulties are introduced into the analysis when the distribution of θ is discrete rather than continuous.

3. PROPERTIES OF THE SOLUTIONS TO (PA-C) AND (PA-D)

There are a number of important differences between the solution to (PA-C) and the solution to (PA-D). However, there is at least one feature that the two solutions share, as Theorem 1 indicates.

Theorem 1. Whether the distributions of θ are continuous or discrete, it will never be optimal for the principal to offer to the agent a non-revealing contract that is not efficient.

Proof of Theorem 1.

Suppose the optimal non-revealing contract, $S^{NR}(x)$, is not efficient, and let k^d represent the agent's expected utility under this contract when the actual distribution of θ is $f^d(\theta)$.

Consider that distribution, $f^r(\cdot)$, for which

$m(d) \equiv k^{*d} - k^d$ is maximized over d,

where $k^{*d} \equiv \int_{\underline{\theta}}^{\bar{\theta}} [x^*(\theta) - w(x^*(\theta),\theta)] f^d(\theta) d\theta.$

Also, consider the efficient contract $s^r(x) = x - k^{*r} + k^r$.

It is straightforward to show that the pair of contracts $\{s^r(x)$ and $s^{NR}(x)\}$ provide the principal with strictly more expected utility than does $s^{NR}(x)$ alone, because the total surplus is increased under $f^r(\theta)$, the agent will choose contract $s^r(x)$ under distribution $f^r(\theta)$, the agent's expected utility is not increased above k^r under distribution $f^r(\theta)$, and, by construction, the agent will not prefer contract $s^r(x)$ to contract $s^{NR}(x)$ under any distribution other than $f^r(\theta)$.

The proof for the discrete case is exactly analogous, and the possibility of non-uniqueness of r poses no problems for the solution technique.

Q.E.D.

The approach to the problem that is undertaken here is to distinguish between the case in which the principal will choose to offer a single non-revealing contract to the agent, and all other cases. Theorem 1 suggests that in order to completely characterize the former case, it need only be determined when it will be optimal for the principal to offer a non-revealing efficient contract. It is in this regard that important differences between the continuous and discrete analyses emerge. These differences are made evident in Theorems 2 and 3.

Theorem 2. When the distributions of θ are continuous, it is optimal for the principal to offer an efficient non-revealing contract to the agent only when the expected surplus from efficient production is the same under all distributions.[3]

The proof of Theorem 2, which is an extension of the work of Sappington [1980a] is presented in Appendix A.

Corollary. When the distributions of θ are continuous, the set of contracts from which the agent is permitted to select will consist wholly of efficient contracts only when the expected surplus from efficient production is the same under all distributions.

The corollary (proved for a special case in Sappington [1980a]) follows from Theorem 2 and the fact that it is impossible to design efficient revealing contracts since efficient contracts must be of the form $s(x) = x - k$ (k a constant) when θ has a continuous distribution. This lattermost fact is proved in Appendix B.

Recall that when the agent is risk-neutral and the principal and agent share the same precontractual information, an outcome of the agency relationship that is ex post Pareto-efficient is guaranteed. Theorem 2 and its corollary indicate that such an efficient outcome is guaranteed only under a very restrictive condition when precontractual beliefs differ and when θ follows a continuous distribution. Theorem 3 suggests that the conditions under which an efficient outcome will be guaranteed even in the presence of asymmetric precontractual information are less restrictive when the distribution of θ is discrete.

<u>Theorem 3.</u> When the distributions of θ are discrete, it may be optimal for the principal to offer an efficient non-revealing contract to the agent even when the expected surplus from efficient production is not identical under all distributions.

<u>Proof of Theorem 3.</u>

An example suffices to prove the theorem. To this end, let $w(x,\theta) = (x/\theta)^2$, $D = 2$, $n = 3$, $\theta_1 = 1$, $\theta_2 = \sqrt{5/2}$, $\theta_3 = 2$, $p_1^1 = p_2^1 = 3/8$, $p_3^1 = 1/4$, $p_1^2 = p_2^2 = 1/8$, $p_3^2 = 3/4$.

It is straightforward, although tedious, to show that the contract

Page Missing in Original Volume

Proof of Theorem 4.

1. If the optimal non-revealing contract is not efficient, then, by Theorem 1, the principal will offer revealing contracts to the agent. It is shown here that the optimal non-revealing contract is not efficient (almost always) when the agent's expected utility cannot be held to zero under both distributions with a single efficient contract. The analogous result for the case of revealing contracts can then be proved using identical techniques. The proof of this latter result is omitted.

2. If the agent's expected utility cannot be held to zero under both distributions with an efficient contract, then either the optimal non-revealing contract holds the agent's expected utility to zero under both distributions and is inefficient (in which case the proof is complete), or the agent's expected utility under the optimal contract is strictly positive under at least one distribution. Suppose this latter possibility is the relevant one.

3. The form of the optimal contract is derived from the solution to (PA-C) with d = 1,2. If the solution is a pooling contract (which is inefficient because there is a unique efficient outcome associated with each state of nature), then, again, the proof is complete. Suppose, then, that the optimal contract is not a pooling contract.

4. Let λ^d and λ_{ij} be the Lagrange multipliers associated with constraints (IP) and (SCW) respectively. Then, after some manipulation and proof by contradiction, it can be shown that $\forall i\ \lambda_{ij} = 0\ \forall j > i + 1$ and $\forall j < i - 1$, and that if $\lambda_{ij} > 0$, $\lambda_{ji} = 0\ \forall i,j$, and also that the necessary conditions for a maximum include

$$\lambda_{i,i-1} - \lambda_{i-1,i} = [\phi^1 - \lambda^1]h_i \qquad \forall i = 2,\ldots,n \qquad . \qquad (4.1)$$

where $h_i \equiv \sum_{j=i}^{n}(p_j^1 - p_j^2)$, and

$$\sum_{d=1}^{2}\phi^d p_i^d[1 - w_x(x_i,\theta_i)] + \sum_{j=i-1}^{i+1}\lambda_{ji}[w_x(x_i,\theta_j) - w_x(x_i,\theta_i)] = 0 \qquad (4.2)$$

$\forall i = 1,\ldots,n$ where $\lambda_{n,n+1} = \lambda_{n+1,n} = \lambda_{10} = \lambda_{01} = 0$.

5. There are four distinct cases to consider:

CASE 1: $h_i > 0$ and $h_{i+1} > 0$.

CASE 2: $h_i > 0$ and $h_{i+1} \leqq 0$.

CASE 3: $h_i \leqq 0$ and $h_{i+1} > 0$.

CASE 4: $h_i \leqq 0$ and $h_{i+1} \leqq 0$.

6. Consider CASE 1 and assume that it is when $d = 1$ that the agent's expected utility is strictly positive so that $\lambda^1 = 0$. Then, from (4.1)

$$\lambda_{i,i-1} = \phi^1 h_i; \qquad \lambda_{i+1,i} = \phi^1 h_{i+1}; \quad \text{and} \quad \lambda_{i-1,i} = \lambda_{i,i+1} = 0.$$

Hence, from (4.2),

$$\sum_{d=1}^{2}\phi^d p_i^d[1 - w_x(x_i,\theta_i)] = \phi^1 h_{i+1}[w_x(x_i,\theta_i) - w_x(x_i,\theta_{i+1})]$$

so that $w_x(x_i,\theta_i) < 1$, or equivalently $x_i < x_i^*$ because the right hand side of this equation is strictly positive since $w_{x\theta}(\cdot,\cdot) < 0$.

7. Similar calculations show that when $\lambda^1 = 0$, in:

Case 2, $x_i = x_i^*$;

Case 4, $x_i \geqq x_i^*$ with equality if and only if $h_i = 0$;

Case 3, x_i is determined by the equation

$$p_i^2[1 - w_x(x_i,\theta_i)] = \phi^1 h_i[w_x(x_i,\theta_{i-1}) - w_x(x_i,\theta_{i+1})] - \phi^1[p_i^1 - p_i^2][1 - w_x(x_i,\theta_{i+1})]. \quad ($$

8. By convention, $p_{\ell}^2 > p_{\ell}^1$ in the least productive state, θ_{ℓ}, in which probabilities are not equal under both distributions. Hence, x_{ℓ} is determined according to Case 3, so that if $\lambda^1 = 0$, $x_{\ell} \neq x_{\ell}^*$ almost always. It is only when equation (4.3) holds at $x_{\ell} = x_{\ell}^*$ (which will fail to be the case almost always when distributions of θ are generated randomly) that the optimal non-revealing contract can be efficient.

9. Similar arguments reveal that when $\lambda^2 = 0$, $x_{\ell+1} \neq x_{\ell+1}^*$ almost always, so that the proof is complete.

Q.E.D.

Theorem 4 suggests that in order to determine when the optimal set of contracts will not consist wholly of efficient contracts, it is sufficient to determine when it is impossible for the principal to deprive the agent of any expected surplus from production with efficient contracts. To this end, Lemma 1 explores the properties of that efficient contract which holds the agent to zero expected surplus under one distribution of θ and minimizes his expected surplus under the other distribution. Since the value of output produced in any state is fixed (at its efficient level) under this contract, it is only the payoff differentials, $s_i - s_{i-1}$, that are subject to control by the principal. The optimal manner in which to set these differentials (subject to the self-selection (SSW) constraints) depends in an important way (which is analyzed in detail in Section 5) upon the value of both cummulative distributions of θ in any particular state.

<u>Lemma 1</u>. The efficient contract $\{(x_1^*, s_1), \ldots, (x_n^*, s_n)\}$ that awards the agent zero expected utility under distribution d(= 1 or 2) and also awards him the least possible expected utility under distribution r ($\neq$ d, r = 1 or 2) has the following structure:

(a) $s_i - s_{i-1} = w(x_i^*, \theta_i) - w(x_{i-1}^*, \theta_i)$ for all i such that

$$\sum_{j=i}^{n} p_j^r \geq \sum_{j=i}^{n} p_j^d$$

(b) $s_i - s_{i-1} = w(x_i^*, \theta_{i-1}) - w(x_{i-1}^*, \theta_{i-1})$ for all i such that

$$\sum_{j=i}^{n} p_j^r < \sum_{j=i}^{n} p_j^d.$$

Proof of Lemma 1.

Properties (a) and (b) characterize the solution to the following problem:

$$\underset{S}{\text{Minimize}} \quad \sum_{i=1}^{n} p_i^r \, [s_i - w(x_i^*, \theta_i)]$$

$$\text{subject to:} \quad \sum_{i=1}^{n} p_i^d \, [s_i - w(x_i^*, \theta_i)] = 0, \text{ and}$$

$$s_i - w(x_i^*, \theta_i) \geq s_j - w(x_j^*, \theta_i) \qquad \forall \; i,j = 1, \ldots, n.$$

The techniques employed to derive the solution to this problem are very similar to those outlined in the proof of Theorem 4, and are omitted here.

Q.E.D.

The conclusions of Lemma 1 can be used to derive a simple test which determines for any pair of distributions of θ whether the optimal set of contracts will consist wholly of efficient contracts. The test is essentially an examination of the level of expected surplus that the agent receives from the contracts described in Lemma 1 under each of the two possible distributions. This test is stated as Theorem 5 and it notes that if the minimum level of expected utility to which the agent can be held under, say distribution 2, with an efficient contract that grants the agent zero expected surplus under distribution 1 is positive (i.e., if $U^{min}(2/1) > 0$), then the principal will offer the agent at least one inefficient contract almost always (by Theorem 4). If, on the other hand, the minimum level of expected utility is non-positive for both distributions (i.e., if $U^{min}(2/1) \leqq 0$ and $U^{min}(1/2) \leqq 0$), then the principal can extract all of the surplus from the agent with efficient contracts, and will do so. The analytic conditions (expressed in terms of the probability distributions and the agent's utility function) that distinguish these possible cases are stated here as Theorem 5.

Theorem 5. (Analytic Test for Optimality of Inefficient Contracts)

Define
$$U^{min}(2/1) \equiv \sum_{j=2}^{n} [p_j^2 - p_j^1] \sum_{z=2}^{j} q^1(z)$$

and
$$U^{min}(1/2) \equiv \sum_{j=2}^{n} [p_j^1 - p_j^2] \sum_{z=2}^{j} q^2(z)$$

where
$$q^1(z) = g(h_z)\,[w(x_z^*, \theta_{z-1}) - w(x_z^*, \theta_z)] + [1 - g(h_z)]\,[w(x_{z-1}^*, \theta_{z-1}) - w(x_{z-1}^*, \theta_z)],$$

$$q^2(z) = g(h_z)\,[w(x_{z-1}^*, \theta_{z-1}) - w(x_{z-1}^*, \theta_z)] + [1 - g(h_z)]\,[w(x_z^*, \theta_{z-1}) - w(x_z^*, \theta_z)],$$

$$g(h_z) = \begin{cases} 1 & \text{if} \quad h_z > 0 \\ 0 & \text{if} \quad h_z \leqq 0 \,; \end{cases}$$

and
$$h_z \equiv \sum_{i=z}^{n} [p_i^1 - p_i^2] \,.$$

Then, if:

(1) $U^{min}(1/2) > 0$ or $U^{min}(2/1) > 0$,

The optimal set of contracts will contain at least one contract that is not efficient almost always;

(2) $U^{min}(1/2) = 0$ or $U^{min}(2/1) = 0$,

the principal will offer the agent an efficient non-revealing contract;

(3) $U^{min}(1/2) < 0$ and $U^{min}(2/1) < 0$,

the principal will offer the agent a pair of revealing contracts, both of which are efficient.

The proof of Theorem 2 is outlined in Appendix C.

Theorems 1, 4, and 5 characterize completely the optimal strategy for the principal given any pair of discrete distributions of θ and the agent's utility function. To demonstrate that the conditions under which inefficient contracts will be observed are both likely in practice and of economic importance, the conclusions of Section 4 are presented.

4. INEFFICIENT CONTRACTS AND STOCHASTIC DOMINANCE

In this section, it is shown that there is at least one important class of pairs of discrete distributions under which the principal will offer the agent at least one inefficient contract almost always. In order to specify this class precisely, the following definition is essential.

Definition: A distribution (1) stochastically dominates another distribution (2) if and only if

$$\sum_{i=j}^{n} p_i^1 \geq \sum_{i=j}^{n} p_i^2 \qquad \text{for all } j = 1, \ldots, n.$$

Note that with the convention that $p^2 > p^1$ in the lowest state (θ_ℓ) in which probabilities are not identical under both distributions (which implies that $h_{\ell+1} > 0$ where $h_z = \sum_{i=z}^{n} [p_i^1 - p_i^2]$), distribution 2 cannot stochastically dominate distribution 1.

Before the main relationship between inefficient contracts and distributions that stand in a relation of stochastic dominance is stated in Theorem 6, a note on the economic significance of stochastic dominance is in order. In its role as a "productivity parameter", θ could be regarded as an indication of the random "output/error" ratio in production. Thus, the advent of a technological improvement which systematically reduces the probability of low output/error ratios and increases the probability of high ones is just one example of a situation in which one distribution of θ (reflecting superior technology) may stochastically dominate another.

<u>Theorem 6</u>. In an informational environment in which one distribution of θ stochastically dominates the other, the set of contracts that the principal will offer the agent will contain at least one inefficient contract almost always.

<u>Proof of Theorem 6</u>.

1. Suppose distribution 1 stochastically dominates distribution 2. Then

$$h_z \equiv \sum_{i=z}^{n} [p_i^1 - p_i^2] \geq 0 \qquad \forall\ z = 1, \ldots, n.$$

2. It can be shown that the conclusion of Theorem 5 is unchanged if $g(m_z)$ is defined as

$$g(h_z) = \begin{cases} 1 & \text{if} \quad h_z \geq 0 \\ 0 & \text{if} \quad h_z < 0. \end{cases}$$

Therefore, when distribution 1 stochastically dominates distribution 2, $g(h_z) = 1 \; \forall \; i$, so that from Theorem 5,

$$U^{min}\ (2/1) = \sum_{j=2}^{n} [p_j^2 - p_j^1] \sum_{z=2}^{j} b^1(z), \text{ and } \quad U^{min}\ (1/2) = \sum_{j=2}^{n} [p_j^1 - p_j^2] \sum_{z=2}^{j} b$$

where $\qquad b^2(z) = [w(x^*_{z-1},\theta_{z-1}) - w(x^*_{z-1},\theta_z)] > 0, \quad$ and

$$b^1(z) = [w(x^*_z,\theta_{z-1}) - w(x^*_z,\theta_z)] > 0 \quad \forall \; z.$$

3. Expanding these expressions and recombining terms, it can be shown that

$$U^{min}\ (2/1) = -\sum_{z=2}^{n} b^1(z)h_z < 0, \text{ and } \quad U^{min}\ (1/2) = \sum_{z=2}^{n} b^2(z)h_z > 0.$$

4. The proof is then complete by condition (1) in Theorem 5.

Q.E.D.

It can be shown that when U^{min} (2/1) < 0 and U^{min} (1/2) > 0, the set of contracts that constitute the optimal strategy for the principal will hold the agent's expected utility level to zero under distribution 2, and grant the agent an expected return which is strictly positive under distribution 1. Thus, when one distribution (1) stochastically dominates another (2), the agent's expected utility will be held to zero when the actual distribution is 2, but will be strictly positive when the actual distribution is 1. It should be noted that the conclusions cited in Theorem 6 are also true if the distributions of θ are continuous, because if one continuous distribution stochastically dominates another, the expected surplus from efficient production under the former must always exceed the latter.

Before an explanation of the foregoing findings is offered in Section 5, a special case of stochastic domination is explored below in the corollary to Theorem 6.

Corollary. When the distributions of θ are such that under each one, only two different states of nature may occur, the optimal strategy for the principal will always be to induce an inefficient outcome in some states of nature.

Proof of Corollary.

The corollary follows from Theorem 6 because, by convention, $p_1^2 > p_1^1$. Hence, $h_1 \geqq 0$ and $h_2 > 0$, so distribution 1 stochastically dominates distribution 2 whenever each distribution places positive probability on only the same two states of nature. Using techniques similar to those outlined in proof of Theorem 4, it can also be shown that the corollary is true as stated, and need not be expressed as a "probability-one" statement, as in the case with Theorems 4, 5, and 6.

Q.E.D.

5. AN EXPLANATION OF THE RESULTS

Ideally, the principal would like to hold the agent's expected utility to zero with an efficient contract under every distribution. This strategy would guarantee that the agent was granted the smallest possible share of the largest possible surplus under each distribution. However, since the principal does not know the actual distribution, he cannot always accomplish this goal. Theorems 2 through 4 indicate that the conditions under which it will be possible to do so are more restrictive when θ has a continuous rather than a discrete distribution. The fundamental reason for this result is that the set of contracts that are efficient when the distribution of θ is continuous is a proper subset of the corresponding set when θ has a discrete distribution.

When θ is distributed continuously, an efficient contract must be of the form $s(x) = x - k$ (k a constant), as is proved in Appendix B. Consequently, the only way in which the principal can ensure that the total expected surplus is as large as possible under each distribution is to offer the non-revealing contract $s(x) = x - k_{min}$ where $k_{min} = \underset{d}{\text{minimum}}\ k(d)$ and $k(d) \equiv \int_{\theta_1}^{\theta_n} [x^*(\theta) - w(x^*(\theta),\theta)]f^d(\theta)d\theta$. This contract, though, will grant the agent strictly positive expected utility under all distributions for which the expected surplus from efficient production ($k(d)$) exceeds k_{min}. Theorem 2 reveals that as long as it is not the case that $k(d) = k_{min}\ \forall d = 1,\ldots,D$, it will not be optimal for the principal to pursue this strategy. Instead, the principal should sacrifice some of the total surplus under at least one distribution in order to obtain more favorable terms of trade with a different contract under at least one other distribution (as indicated in Theorem 1).[4]

When the distributions of θ are discrete, however, the principal gains sufficient flexibility in the design of contracts to admit the possibility that the total surplus can be maximized and the agent's expected utility held to zero under every distribution of θ, even when the expected surplus is not identical under each distribution. The additional flexibility stems from the fact that the self-selection constraints (SSW) are, in an important sense, less restrictive under discrete distributions than under continuous ones. The self selection contraints in (PA-D) require that under an efficient contract,

$$w(x_i^*,\theta_i) - w(x_{i-1}^*,\theta_i) \leq s_i - s_{i-1} \leq w(x_i^*,\theta_{i-1}) - w(x_{i-1}^*,\theta_{i-1}) \qquad (5.1)$$

in any two contiguous states, θ_i and θ_{i-1}. Thus, the self selection constraints impose upper and lower bounds on the allowable difference in compensation that can be awarded the agent for efficient production in contiguous states. As θ_i approaches the level of θ_{i-1} (the limiting case being characteristic of a continuous distribution) the upper and lower bounds converge so that the principal's freedom to manipulate payoff differentials for different levels of performance disappears.[5]

The importance of being able to manipulate payoff differentials more freely is that it permits the principal to render a particular contract more or less attractive to the agent under one distribution without sacrificing the total surplus or changing the agent's expected utility from that contract under another distribution. This is the essence of the example described in the proof of Theorem 3. In that example, a lump-sum type contract that held the agent's expected utility to zero under the first distribution "1" (consisting of allocations A_1, A_2, and A_3 in Figure 1) would provide the agent with a strictly positive level of expected utility under the second distribution "2". However, by increasing the payoff to the agent for producing x_1^* and reducing the compensation awarded the agent for producing x_2^* and x_3^* so as to leave unchanged the agent's expected utility under distribution 1, the efficient contract (A_1', A_2', and A_3' in Figure 1) becomes less attractive to the agent in states more productive than θ_1, states which are relatively more likely to occur under distribution 2 than under 1.

To examine the principal's optimal strategy under more general conditions when the pair of possible distributions are discrete, consider, initially, the situation in which $x_1, \ldots, x_n$ have somehow been determined and it is the principal's task to choose $s_1, \ldots, s_n$ optimally. Furthermore, suppose that the principal chooses $s_1, \ldots, s_n$ such that, for some s_i and s_{i-1},

$s_i - s_{i-1} < w(x_i,\theta_{i-1}) - w(x_{i-1},\theta_{i-1})$ where $h_i = \sum_{j=i}^{n} (p_j^1 - p_j^2) < 0$. This specification cannot be optimal for the principal; for suppose the principal were to increase s_i (by Δs_i) without changing $[s_j - s_{j-1}]$ $\forall$ $j \neq i$, i.e., suppose $[s_i - s_{i-1}]$ is increased by Δs_i without otherwise altering the structure (or form) of the contract. (Note that this perturbation requires that all s_j $j = i, \ldots, n$ in the original contract be increased by Δs_i.) Then the agent's expected utility from this new contract under distribution 2 exceeds the corresponding measure under the original contract by $T_2 = \sum_{j=1}^{n} p_j^2 \Delta s_i$. Also the agents expected utility from the new contract is increased by $T_1 = \sum_{j=i}^{n} p_j^1 \Delta s_i$ under distribution 1. Therefore, if this new contract were amended by reducing the payoff to the agent in every state of nature by T_2, the agent's expected utility under this new amended contract would still be zero under distribution 2. However, under distribution 1, the agent's expected utility from this same contract would fall short of the corresponding level under the original contract by $T_2 - T_1 > 0$. Consequently, since the proposed modifications of the original contract reduce the agent's expected utility (without altering the total expected surplus), the original contract could not have been optimal for the principal.

Hence, whenever the probability that a state of nature at least as high as θ_i will be realized is greater under distribution 1 than distribution 2, it will be optimal for the principal to set s_i as far above s_{i-1} as possible without violating the self-selection constraints characterized by equation (5.1). Similar arguments explain why, in the optimal non-revealing contract, $[s_i - s_{i-1}]$ will always be set at its minimum level, $w(x_{i-1},\theta_i) - w(x_{i-1},\theta_i)$, whenever $h_i > 0$.

Thus, the conclusions of Lemma 1 and the analytic conditions of Theorem 5 are explained. What remains to be explored is the nature of the benefits which accrue to the principal when inefficient outcomes are induced.

A simple extension of the preceding argument explains why, when the principal can determine both x_i and s_i, $i = 1, \ldots, n$, his expected utility is increased, ceteris paribus, if he chooses these variables so as to maximize the increase (or gain, G_i) in the agent's ex post utility level in state θ_i over that achieved in state θ_{i-1} whenever $h_i < 0$ and to minimize the same gain whenever $h_i > 0$. However, two additional complications are introduced when the principal can determine x_j $j = 1,\ldots,n$ in addition to the associated compensation levels. First, changes in any x_j alter the total available surplus from production. Second, a change in any x_j may affect the magnitude of both G_j and G_{j+1}.

To examine these effects in more detail, consider, first, Case 1 in the proof of Theorem 4, where $h_i > 0$ and $h_{i+1} > 0$. As explained above, for any levels of x_j $j = i-1, i, i+1$, s_i and s_{i+1} should be set in the optimal non-revealing contract such that

$$s_i - s_{i-1} = w(x_i,\theta_i) - w(x_{i-1},\theta_i)$$

and $$s_{i+1} - s_i = w(x_{i+1},\theta_{i+1}) - w(x_i,\theta_{i+1}).$$

Consequently,

$$\begin{aligned} G_i &\equiv [s_i - w(x_i,\theta_i)] - [s_{i-1} - w(x_{i-1},\theta_{i-1})] \\ &= w(x_{i-1},\theta_{i-1}) - w(x_{i-1},\theta_i), \quad \text{and} \end{aligned}$$

$$G_{i+1} = w(x_i,\theta_i) - w(x_i,\theta_{i+1}).$$

These expressions reveal that when the contract is structured optimally, the increase in the agent's level of ex post utility in state θ_i over the corresponding level in state θ_{i-1} is independent of the value of x_i while the corresponding gain in state θ_{i+1} increases with x_i (since $w_{x\theta}(\cdot) < 0$). Consequently, because the first-order effect on the total surplus of a deviation from the efficient contract is zero (see footnote 4), it will always be optimal for the principal to set x_i below its efficient level whenever the distributions of θ are such that both h_i and h_{i+1} are strictly positive.

Similarly, in Case 4 where $h_i \leq 0$ and $h_{i+1} \leq 0$, G_i can be shown to be a (strictly) increasing function of x_i (for $h_i < 0$) when S_i and S_{i+1} are chosen optimally in the non-revealing contract, and G_{i+1} can be shown to be independent of the level of x_i. Thus, in order to maximize both G_i and G_{i+1} (subject to the usual efficiency considerations), the principal will set x_i above x_i^* in Case 4.

In Case 2, where $h_i > 0$ and $h_{i+1} \leq 0$, it can be shown that when the contact is structured optimally, both G_i and G_{i+1} are independent of the magnitude of x_i. Hence, the optimal strategy for the principal is to maximize the total surplus in the event that θ_i occurs, i.e., set $x_i = x_i^*$. (Note that the same logic explains why x_i is set at its efficient level when $h_i = 0$ in Case 4.)

It is only in Case 3, where $h_i \leq 0$ and $h_{i+1} > 0$ that some ambiguity is introduced. Here, the principal will, <u>ceteris paribus</u>, benefit from increases in G_i and decreases in G_{i+1}. However, when the contract is structured optimally, both G_i and G_{i+1} are increasing functions of x_i, so that any increase in x_i will increase both G_i and G_{i+1}. It is only when the expected benefits from increasing G_i (through an increase in x_i) happen to exactly offset the expected costs of increasing G_{i+1} at $x_i = x_i^*$ that an efficient outcome will be realized in state θ_i under the optimal non-revealing contract. When discrete distributions of θ are randomly generated, this event will occur with probability zero.

In summary, the principal may induce an inefficient outcome in some states of nature under the optimal (set of) contract(s) because the inefficiency permits him, ceteris paribus, to increase (or decrease) the ex post level of utility awarded the agent in those states beyond the maximum (or minimum) feasible level under an efficient contract. This added flexibility may be of benefit to the principal because it permits him to reduce the expected surplus awarded the agent under some distributions of θ without increasing the agent's expected surplus under other distributions. However, because the ability of the principal to manipulate actual payoff differentials is greater when θ follows a discrete rather than a continuous distribution (as indicated in equation (5.1)), the principal may not need to resort as often to inefficient outcomes to impose the desired "discipline" on the agent when the distribution of θ is discrete.

6. CONCLUSIONS

A standard result in the principal-agent literature claims that in the absence of limited liability considerations (of the type analyzed by Sappington [1982]), the contract that the principal will offer to a risk-neutral agent is an efficient contract. This result, though, is based on the assumption that the principal and agent share the same precontractual beliefs.

This paper has demonstrated that inefficient contracts are often optimal for the principal in the presence of precontractual information asymmetry between principal and agent. The nature of the expected benefits derived by the principal from the introduction of such inefficiencies was characterized, and relatively simple conditions were derived to determine when an arbitrary member of the class of informational environments considered here would give rise to an inefficient outcome in some states of nature. These conditions were also shown to vary depending upon whether the distributions of θ were continuous or discrete.

In conclusion, a few brief comments on two possible extensions of the model are offered. To begin with, it should be noted that sufficient competition among perfectly-informed agents can eliminate the phenomenon of inefficient outcomes in the presence of information asymmetry. If a large number of identical agents are all aware of the true distribution of θ before a contract must be signed, then the principal need only carry out an auction (of the type discussed by Demsetz [1968]), with the agents bidding for the right to carry out production and retain the entire value of their output. The winning bid among risk-neutral agents will approach the expected surplus from efficient production given the true distribution of θ. Hence, even without any knowledge of the true distribution, the principal can capture the entire (actual) expected surplus with the aid of such an auction.

In addition, it should be noted that the structure of the model considered here incorporates important restrictions on the nature of the information asymmetry between principal and agent. In particular, both parties agree on the set of possible distributions of θ and the extent to which the agent's information is better than that of the principal. Absent such agreement, alternative formulations of the model need be derived and explicit attention need be awarded the incentives of both principal and agent to misrepresent their knowledge of the informational environment.[6] Such phenomena appear worthy of further investigation.

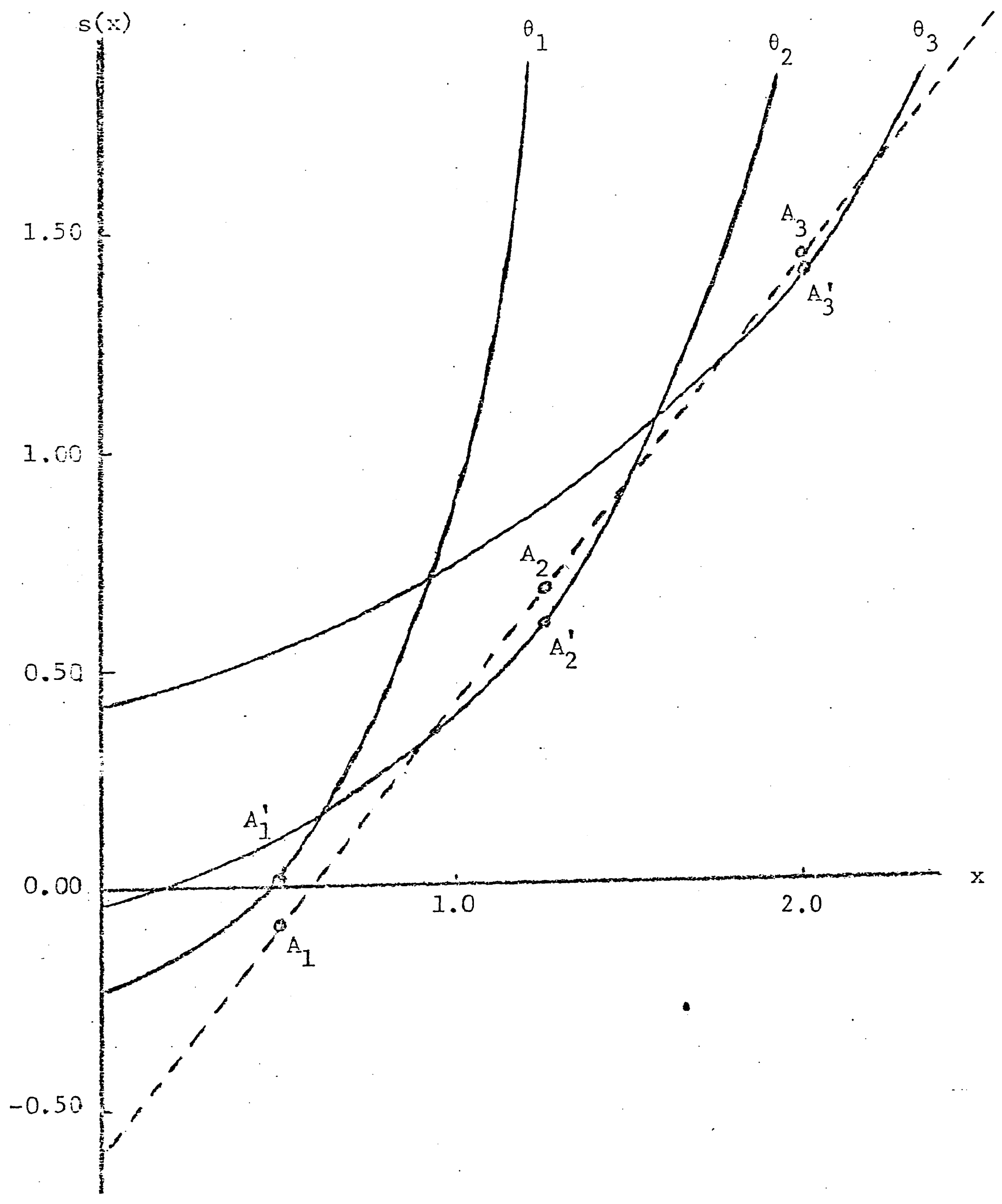


FIGURE 1. Lump-Sum (A_1, A_2, A_3) vs. Optimal (A_1', A_2', A_3') Efficient Contract

Note: θ_i refers to an indifference curve of the agent in state θ_i. The agent's utility increases with movements in a northwesterly direction.

FOOTNOTES

*. The author is especially indebted to Robert D. Willig and Andrew Postlewaite, and also to Edward J. Green for their excellent comments and suggestions. Any errors which may remain are, of course, my own.

1. It is assumed throughout that when the agent is indifferent among two or more contracts (or among two or more levels of effort), he will choose the one most preferred by the principal, were he to share the agent's information.

2. The value of output $x^*(\theta_i)$ that is ex post Pareto efficient in state θ_i is defined by the equation $w_x(x^*(\theta_i),\theta_i) = 1$.

3. The expected surplus from efficient production under continuous distribution $f^d(\theta)$ is defined to be $\int_{\theta_1}^{\theta_n} [x^*(\theta) - w(x^*(\theta),\theta)]f^d(\theta)d\theta$. The expected surplus from efficient production under discrete distribution d is defined to be $\sum_{i=1}^{n} p_i^d[x_i^* - w(x_i^*,\theta_i)]$.

4. This result is due, in part, to the fact that the first-order effect on the total surplus of a deviation from the efficient contract is zero. To see this, note that the total surplus in any state θ_i is $x_i - w(x_i,\theta_i)$, the derivative of which is zero when evaluated at $x_i = x_i^*$.

5. This fact is captured mathematically in the formulation of (PA-C) contained in the proof of Theorem 2. There, the self-selection constraints (SSW) are represented by an equation of motion rather than by upper and lower bounds on allowable payment differentials.

6. For example, if the agent knew that only a single distribution of θ were in fact possible, it would be in his interest to have the principal believe (in error) that an additional distribution that is stochastically dominated by the true distribution might also occur (as suggested in the discussion which follows the proof of Theorem 6).

APPENDIX A

Proof of Theorem 2.

1. The theorem is proved through inspection of the necessary conditions for a maximum of the Hamiltonian function associated with (PA-C), where $x^d(\theta) = x(\theta)\ \forall d = 1,\ldots,D$. The agent's utility function $U(\theta)$ serves as the state variable and $x(\theta)$ is the control variable in the formulation considered here.

The necessary conditions include

$$\phi^d[1 - w_x(x,\theta)]f^d(\theta) - \S(\theta)w_{x\theta}(x,\theta) = 0 \qquad (A.1)$$

$$\dot{U}(\theta) = -w_\theta(x,\theta)$$

$$\dot{\S}(\theta) = \sum_{d=1}^{D} [\phi^d - \lambda^d]f^d(\theta) \qquad (A.2)$$

$$\S(\theta_1) = \S(\theta_n) = 0$$

where λ^d and $\S(\theta)$ are the multipliers associated with the d-th individual rationality constraint and the equation of motion, respectively. Note that the equation of motion accounts for the (SSW) constraints, and the (SSB) constraints are omitted since only a single contract is considered here. Note also that the solution to this problem need not, in general, have $s(\cdot)$ be a differentiable function of x. However, the proposition in Appendix B indicates that if $s(\cdot)$ is not differentiable, then it is not an efficient contract, which is all that matters for this proof.

2. Suppose the optimal non-revealing contract is efficient. Then $w_x(x,\theta) = 1\ \forall\theta$ so that $\S(\theta) = 0\ \forall\theta$ from (A.1).

3. From (A.2), $\S(\theta) = \sum_{d=1}^{D} [\phi^d - \lambda^d] F^d(\theta) \quad \forall\theta$ where $F^d(\theta) = \int_{\theta_1}^{\theta} f^d(\theta)d\theta$. Note also that since $\S(\theta_n) = 0$ by the transversality condition, $\sum_{d=1}^{D} \lambda^d = 1$.

4. Let the first m distributions be those for which the agent's expected utility is held to zero under this efficient contract. Recall $m < D$ by hypothesis. Also, let the first j distributions be those distributions for which $\lambda^d > \phi^d$. (Note that $1 \leq j \leq m$ since by hypothesis, we cannot have $\lambda^d = \phi^d > 0 \; \forall d = 1,\ldots,D$, and $\lambda^d = 0 \; \forall d > m$ by construction.)

5. Therefore, $\sum_{d=1}^{j} \gamma^d F^d(\theta) = \sum_{d=j+1}^{D} \gamma^d F^d(\theta) \; \forall\theta$ where $\gamma^d = |\lambda^d - \phi^d|$. Now, letting the range of all integrals be the interval $[\theta_1, \theta_n]$, define $k_{min} = \int [x^*(\theta) - w(x^*(\theta),\theta)] f^d(\theta)d\theta$ for $d = 1,\ldots,m$. Then it follows that

$$\int \{x^*(\theta) - w(x^*(\theta),\theta) - k_{min}\} f^d(\theta)d\theta =$$

$$\sum_{d=j+1}^{D} \alpha^d \int \{x^*(\theta) - w(x^*(\theta),\theta) - k_{min}\} f^d(\theta)d\theta > 0, \quad \text{where } \alpha^d \equiv \frac{\gamma^d}{\gamma^1},$$

which is a contradiction.

Q.E.D.

APPENDIX B

Lemma. Every efficient contract must be a differentiable function of x when θ has a continuous distribution with strictly positive support on the interval $[\theta_1, \theta_n]$.

Proof of Lemma.

1. The self-selection constraints (SSW) require that for every efficient contract, $s^*(x)$,

$$\frac{w(x_i^*, \theta_i) - w(x_j^*, \theta_i)}{x_i^* - x_j^*} \leqq \frac{s^*(x_i^*) - s^*(x_j^*)}{x_i^* - x_j^*} \leqq \frac{w(x_i^*, \theta_j) - w(x_j^*, \theta_j)}{x_i^* - x_j^*}$$

for any $x_j^* \varepsilon [x_1^*, x_n^*)$ and for all $x_i^* > x_j^*$.

2. By the Mean Value Theorem, there exist x', $x'' \varepsilon [x_j^*, x_i^*]$ such that

$$\frac{w(x_i^*, \theta_i) - w(x_j^*, \theta_i)}{x_i^* - x_j^*} = w_x(x', \theta_i), \quad \text{and}$$

$$\frac{w(x_i^*, \theta_j) - w(x_j^*, \theta_j)}{x_i^* - x_j^*} = w_x(x'', \theta_j)$$

because $w(x, \theta)$ is a continuous function of x for all θ.

3. As $\theta_i \to \theta_j$, $x_i^* \to x_j^*$ because $x^*(\theta)$ is a continuous function of θ. The continuity follows from the fact that $x^*(\theta)$ is the maximum of a continuous function $x - w(x, \theta)$ (see the corollary to Theorem 3 (p. 30) in Hildenbrand [1974]).

Therefore, in the limit as $x_i^* \to x_j^*$,

$$w_x(x_j^*, \theta_j) \leqq S_x^{*+}(x_j^*) \leqq w_x(x_j^*, \theta_j),$$

where

$$s_x^{*+}(x_j^*) \equiv \underset{x_i^* \to x_j^*}{\text{limit}} \left\{ \frac{s^*(x_i^*) - s^*(x_j^*)}{x_i^* - x_j^*} \right\}.$$

4. The same arguments can be employed to show that for all $x_t^* < x_j^*$

$$w_x(x_j^*, \theta_j) \leq s_x^{*-}(x_j^*) \leq w_x(x_j^*, \theta_j),$$

where

$$s_x^{*-}(x_j^*) \equiv \underset{x_t^* \to x_j^*}{\text{limit}} \left\{ \frac{s^*(x_j^*) - s^*(x_t^*)}{x_j^* - x_t^*} \right\}.$$

5. Therefore, the left-hand and right-hand derivatives of $s^*(x)$ exist and are identical for all values of $x \in (x_1^*, x_n^*)$.

Q.E.D.

The following proposition then follows directly from the lemma.

Proposition. Any efficient contract must be of the form $s^*(x) = x - k$ (k a constant) when θ has a continous distribution with strictly positive support on a closed interval.

Proof of Proposition

Since the agent chooses $x(\theta)$ to maximize $s(x(\theta)) - w(x(\theta), \theta)$ in each state, on efficient contract must have $s_x^*(x) = 1$ for all x.

This condition implies that $s^*(x) = x - k$ since the efficient contract is a differentiable function of x.

Q.E.D.

APPENDIX C

Proof of Theorem 5

1. Using Lemma 1, it can be shown that under the contract described there, the increase in utility received by the agent in state θ_z over the level received in state θ_{z-1} is $q^d(z)$, $d = 1, 2$; thus,

$$s_2 - w(x_2^*,\theta_2) = s_1 - w(x_1^*,\theta_1) + q^d(2)$$

and more generally,

$$s_j - w(x_j^*,\theta_j) = s_1 - w(x_1^*,\theta_1) + \sum_{z=2}^{j} q^d(z).$$

2. Using this relationship and the fact that

$$\sum_{j=1}^{n} p_j^d \, [s_j^d - w(x_j^*,\theta_j)] = 0,$$

it can be shown that

$$s_1 - w(x_1^*,\theta_1) = -\sum_{j=2}^{n} p_j^d \sum_{z=2}^{j} q^d(z), \text{ and}$$

$$U^{min}(1/2) \equiv \sum_{j=1}^{n} p_j^1 \, [s_j^2 - w(x_j^*,\theta_j)]$$

$$= \sum_{j=2}^{n} [p_j^1 - p_j^2] \sum_{z=2}^{j} q^2(z).$$

3. Similar analyses can be performed to derive the stated expression for

$$U^{min}(2/1) \equiv \sum_{j=1}^{n} p_j^2 \, [s_j^1 - w(x_j^*,\theta_j)].$$

4. The conclusion of the theorem then follows directly from the interpretations of $U^{min}(1/2)$ and $U^{min}(2/1)$, and from Theorem 4.

Q.E.D.

REFERENCES

Blackwell, D.: "Comparison of Experiments," in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, ed. by J. Neyman. University of California Press, 1951.

Demsetz, H.: "Why Regulate Utilities?" The Journal of Law and Economics, 11 (1968), 55-65.

Green, J. R.: "Statistical Decision Theory Requiring Incentives for Information Transfer," Harvard Institute of Economic Research, Discussion Paper Number 708, May 1979.

Green, J. R., and N. L. Stokey.: "A Two-Person Game of Information Transmission," Harvard Institute of Economic Research, Discussion Paper Number 751, March 1980 (a).

Green, J. R., and N. L. Stokey.: "The Value of Information in the Delegation Problem," Harvard University Mimeo, May 1980 (b).

Harris, M., and A. Raviv.: "Optimal Incentive Contracts with Imperfect Information," The Journal of Economic Theory, 20 (1979), 231-259.

Harris, M., and R. M. Townsend.: "Resource Allocation Under Asymmetric Information," Econometrica, 49 (1981), 33-64.

Hildenbrand, W.: Core and Equilibria of Large Economy. Princeton, New Jersey: Princeton University Press, 1974.

Holmstrom, B.: "Moral Hazard and Observability," The Bell Journal of Economics, 10 (1979), 74-91.

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

Sappington, D.: "Information Asymmetry Between Principal Agent: The Continuous Case," Economics Letters, 5 (1980a), 371-375.

Sappington, D.: "The Inefficiency of Indirect Control Under Information Asymmetry," Unpublished Doctoral Dissertation at Princeton University, October 1980 (b).

Sappington, D.: "Limited Liability Contracts Between Principal and Agent," Journal of Economic Theory, Forthcoming (1982).

Shavell, S.: "Risk Sharing and Incentives in the Principal and Agent Relationship," The Bell Journal of Economics, 10 (1979), 55-73.