"Resource Allocation under Asymmetric Information", joint with Milton Harris, Econometrica 49(1), January 1981: 33-69.

Econometrica, Vol. 49, No. 1 (January, 1981)

RESOURCE ALLOCATION UNDER ASYMMETRIC INFORMATION[1]

BY MILTON HARRIS AND ROBERT M. TOWNSEND

The purpose of this paper is to provide a method for characterizing *efficient allocation processes* and efficient allocations for a large class of environments in which asymmetric information is an important factor. This method is based on a rigorous application of statistical decision theory and makes explicit both the information available to agents ex ante and the way in which information is transmitted during any multistage allocation process.

1. INTRODUCTION

THE PURPOSE OF THIS PAPER is to provide a method for characterizing *efficient allocation processes* and efficient allocations for a large class of environments in which asymmetric information is an important factor.[2] This method is based on a rigorous application of statistical decision theory and makes explicit both the information available to agents ex ante and the way in which information is transmitted during any multistage allocation process.[3] Our approach is illustrated by its application to a principal-agent environment, a public goods environment, and a competitive environment with informed and uninformed traders.

In much of the literature on environments with asymmetric information, efficiency is defined with respect to full or symmetric information.[4] Much of the work on public goods, for example, focuses on constructing mechanisms to achieve Samuelson-Lindahl allocations (e.g., Groves-Ledyard [**9**]). These are allocations which would be Pareto optimal if all agents possessed full information about the *actual* preferences of the agents. We argue that, as an alternative approach to characterizing efficient allocations and mechanisms, one must consider explicitly the restrictions imposed by asymmetric information.[5] In

[1] This paper has formed the basis for presentations at the NBER-CEME Conference on Decentralization (held at the University of Minnesota, April, 1978), the Econometric Society meetings (Chicago, August, 1978), and Northwestern University (October, 1978). We would like to thank the participants at these seminars, and especially J. Jordan, J. Ledyard, E. Maskin, M. Satterthwaite, and C. Wilson for helpful comments. In addition we are grateful for the extremely valuable suggestions of two anonymous referees. Financial support from the National Science Foundation (under grant SOC-7826262) and the Federal Reserve Bank of Minneapolis is gratefully acknowledged. We assume full responsibility for any errors as well as the views expressed here.

[2] Such environments have recently been the object of considerable and growing attention. Examples include public goods (Groves and Ledyard [**9**]), signaling and screening (Spence [**29**], Riley [**22**]), agency (Spence and Zeckhauser [**30**], Shavell [**28**]), insurance (Rothschild and Stiglitz [**25**], C. Wilson [**31**]), competitive markets with some informed traders (Green [**7**], Grossman [**8**]), and auction markets (Holt [**14**], R. B. Wilson [**32**]). By *asymmetric* information, we mean a situation in which certain agents believe other agents to be *better* informed about some aspect of the economic environment. This is to be distinguished from *differential* information, a situation in which agents have different ex ante information, but no agent believes any other agent to be better informed than himself.

[3] Since writing an earlier version of this paper we have discovered Myerson [**18**] whose approach is similar to our own in several crucial aspects. We also wish to acknowledge the influence of the seminal work of Hurwicz on resource allocation mechanisms (e.g., see Hurwicz [**15**]). Reiter pioneered in this area as well (e.g., see Reiter [**21**]).

[4] Myerson [**18**] is a notable exception.

[5] Such restrictions have proved quite useful in developing positive theories of economic phenomena; see Prescott-Townsend [**19**].

particular we argue that in environments with asymmetric information, uncertainty about preferences, technology, and endowments may be characterized by probability distributions over unknown shocks or parameters. Then one must take account of this information structure in developing an equilibrium theory of allocation processes or mechanisms and in defining efficient allocations. *This results in a different theory of allocation processes and a different definition of optimal allocations than those which would obtain under full information.* In particular, both the normative and the positive implications of our approach differ, for some environments, from those of the traditional approach.

The approach which we here propose consists of five steps. First, we carefully specify the economic environment including the information structure. Second, we define the concept of an allocation mechanism for an environment and specify the class of available mechanisms. Third, we define the concepts of equilibrium of a mechanism and equilibrium allocations of a mechanism. Fourth, we define preferences of the agents over mechanisms and the concept of an efficient mechanism. Fifth, and last, we characterize efficient mechanisms and their allocations. Each aspect of the approach is now discussed briefly.

The general class of environments considered in this paper (see Section 3) includes arbitrary sets of agents, production, and a rich set of possible information structures comprising both public and private information. An environment consists of a specification of three objects. First is the set of agents and the technology of production and exchange available to each coalition of agents. Second, the preferences of each agent are specified. These are assumed to depend on a vector of exogenously determined parameters. The information structure is then specified as the set of parameters which is *observed* by each agent together with a prior distribution for each agent describing his information about the parameters he cannot observe. More than one agent may observe the same parameter, i.e., some information may be public. These observations are the sole source of all information asymmetries.

Once an environment is specified, the concept of an allocation mechanism for the environment may be defined. Intuitively, an allocation mechanism for a given environment is simply a set of rules which specify the game to be played by the agents in allocating resources and a specification of how the allocation is determined, given the "plays" of the agents during the game. In Section 4, this concept is defined formally, together with the class of all available mechanisms.

Next we must have a theory of the outcome of any given mechanism, i.e., we must define the concept of equilibrium strategies for a given mechanism. In this context, a *strategy* of an agent is a mapping which determines his play or signal at a given stage of a mechanism as a *function* of the parameters initially observed by him and the sequence of past signals of all agents. Optimality of a strategy is defined relative to the information the agent has at the time the strategy is used. This information will be different from the initial information if the signals of other agents convey information. In particular, agents are assumed to update their beliefs about unobserved parameters using Bayes' rule and the sequence of past signals of other agents. Using this definition of optimal strategies, we define a *perfect Bayesian equilibrium* concept for any mechanism. This concept is related to Harsanyi's [**13**] "Bayesian equilibrium" and Selten's [**27**] "perfect equilibrium."[6] It also draws on the work of Prescott and Visscher [**20**], Kydland and Prescott [**17**], and E. Green [**6**]. Once the equilibrium strategies are determined, the equilibrium allocation of a mechanism is easily defined. The equilibrium concept is discussed more fully in Section 5.

As can be seen from the above discussion, equilibrium allocations of mechanisms are allocations which depend on the actual, realized values of the parameters of the environment (since equilibrium strategies depend on these values). Such allocations will be referred to as *parameter-contingent* (*or p.c.*) *allocations*. Preference orderings of agents *over mechanisms* then are naturally defined using *expected utilities*, based on the initial information structure, over these equilibrium p.c. allocations. Efficient mechanisms are defined in an obvious way, e.g., a core mechanism is one which cannot be improved upon by any coalition of agents, a Pareto optimal mechanism is one which cannot be improved upon by the set of all agents, and so on. These concepts are defined formally in Section 7.

Our general results provide a fairly simple way of executing the final step of the approach, namely that of characterizing efficient allocation mechanisms and efficient allocations. Our first major result (see Section 6) is that the equilibrium p.c. allocations of any mechanism must satisfy certain self-selection properties. That is, any agent with private information must prefer (in the sense of expected utility relative to his initial information) the equilibrium allocation which is to result under the actual value of the parameters to the equilibrium final allocation which would result if the parameters observed by him alone took on some other of their possible values. This is because the actual (equilibrium) strategies must be preferred to all other strategies, including the strategy of acting as if one's observed parameters took on some other value.

Our first result thus shows that certain technically feasible p.c. allocations, namely those which do not satisfy self-selection properties, cannot in fact be achieved.[7] Conversely, a second major result (also in Section 6) is that any p.c. allocation which does satisfy self-selection properties can be achieved under a mechanism. Moreover, it is established that one can restrict attention to a mechanism of a particularly simple form, *a direct mechanism* under which each agent reports on the value of each parameter he observes.[8]

These two results suggest that one can define notions of efficiency directly in the space of p.c. allocations which exhibit the self-selection property. This we do in Section 8 where the equivalence between efficient allocations (in a certain sense) and efficient mechanisms is established. Our third major result is that an allocation

[6] The Bayesian equilibrium concept is used in Arrow [**1**], d'Aspremont and Gerard-Varet [**5**], and Dasgupta, Hammond, and Maskin [**4**] (D-H-M), among others. It is closely related to the rational expectations equilibrium (see Prescott and Townsend [**19**]). We would like to thank L. Hurwicz, J. Jordan, and E. Maskin for calling our attention to some of this literature.

[7] Myerson derives similar results for one-stage mechanisms.

[8] The terminology "direct mechanism" is borrowed from the game theory literature (see, e.g., D-H-M [**4**]).

is efficient if and only if it is associated with an efficient mechanism. This, then, allows one to convert a problem of characterizing efficient mechanisms into one of characterizing efficient allocations, a problem which can usually be solved using standard mathematical programming techniques. Using the second result (outlined in the previous paragraph), one can then easily characterize an efficient direct mechanism which supports the efficient allocations.

An additional result, which is somewhat less general than the three results mentioned above, concerns the relationship between optimal allocations in our sense and full-information (FI) optimal allocations. In particular, we show that if there are only two agents and one is fully informed while the other is not, then any p.c. allocation which is FI optimal for every parameter value and satisfies self-selection is also optimal in our sense (see Section 8). That this result cannot be generalized (without adding other restrictions) is shown via an example (see Section 9).

In addition to the general development of Sections 3–8, outlined above, we also present three examples. (A fourth example concerned with risk sharing in a pure exchange environment is discussed in Harris-Townsend [**12**], where a slightly different approach was taken.[9]) The three examples described in Sections 2, 9, and 10 respectively are (i) a principal-agent environment, (ii) a public goods environment, and (iii) a competitive, pure exchange environment.

The agency problem presented in Section 2, before the general development, introduces several concepts and results central to that development. This environment is essentially the one analyzed by Hurwicz-Shapiro [**16**] in which the agent has better information about some parameters of a production process than the principal. The example illustrates four important points: (i) an allocation is achievable by a mechanism if and only if it satisfies certain self-selection conditions, (ii) not all allocations which are optimal in a full information sense are achievable (i.e., some fail to satisfy the self-selection conditions), (iii) there are allocations which are optimal in our sense but which are *not* full information optimal, and (iv) which allocations are optimal in our sense depend on prior beliefs of the agents.

In Section 9, after the general development, we examine a simple public goods environment in which each agent knows only his own preferences for public and private goods. The main point of this example, aside from further illustrating the power of our approach, is to establish that achievable full information optimal allocations (namely achievable Samuelson-Lindahl allocations) need not be optimal in our sense.[10] Thus if one takes our approach seriously, full information

[9] Further applications are contained in Harris and Raviv [**11**, **12**].

[10] This complements the results of D'Aspremont and Gerard-Varet [**5**] (hereafter D-G-V) and Arrow [**1**] who establish that in some public goods environments there is at least one full information optimal p.c. allocation which can be achieved as a Bayesian equilibrium. A related question is whether a full information optimal p.c. allocation can be achieved under a mechanism satisfying a budget balance requirement under a dominant strategy (strong Nash) equilibrium, for as D-H-M [**4**] point out, any p.c. allocation which can be attained as a dominant strategy equilibrium can also be attained as a Bayesian equilibrium. It is well known that there does not exist one mechanism with this property for a fairly rich class of utility functions. (Of course a given mechanism may work for a particular set of utility

optimal p.c. allocations are not necessarily desirable from a normative point of view.

Finally, in Section 10 we note, by way of contrast, that in a pure exchange economy with informed and uninformed traders (informed traders are all informed about the same thing) competitive equilibrium allocations in which prices fully reveal all information are core allocations in our sense. This result turns on the standard definition of the core and the fact that with no private information there are no self-selection constraints. Section 11 contains some concluding remarks.

2. A PRINCIPAL-AGENT ENVIRONMENT WITH TECHNOLOGICAL UNCERTAINTY

We consider in this section a principal-agent model motivated by Hurwicz-Shapiro [**16**]. A principal, p, has sole access to a technology for transforming labor effort z into output y of the single consumption good of the model in accordance with a production function f. Let $y=f(z,\theta)=\theta z$. Here the parameter θ is drawn from a known distribution. In particular, θ can take on one of two values, θ^1 or θ^2, with $\theta^1<\theta^2$ and φ denotes the probability that $\theta=\theta^1$. The agent (worker), denoted a, knows the actual realization of θ; the principal (manager, landlord) does not. The principal takes the probability that $\theta=\theta^1$ to be φ. The agent alone can supply labor, and such effort cannot be observed by the principal.

The principal cares only about consumption. Thus, letting r denote the reward to the agent (his allotment of the consumption good), the utility function of the principal is assumed to be $U_p(y,r)=y-r$. The agent values both leisure and consumption, and his utility function is assumed to be $u_a(r,z)=r-z^2$. Hence with $y=\theta z$, the agent's utility as a function of output, reward, *and* the productivity parameter θ is $U_a(y,r,\theta)=r-(y/\theta)^2$.

The environment is depicted in Figure 1. The space of feasible allocations is the area below the 45° line through the origin. Indifference curves of the principal are 45° lines with his utility increasing toward the southeast. Indifference curves of the agent depend on θ and are labeled by θ^1 or θ^2. His utility increases toward the northwest. Allocations which would be efficient under full information are shown by the vertical lines labeled C_j (allocations which would be efficient if both parties knew that $\theta=\theta^j$ for $j=1,2$). Thus any point on say C_1 is such that marginal product of labor equals the agent's marginal rate of substitution between effort and reward, if $\theta=\theta^1$.

A parameter-contingent (p.c.) allocation for this environment is an output quota $y(\theta)$ and reward to the agent $r(\theta)$ which depends on the true, realized value

functions.) This has motivated some to consider a weaker notion of equilibrium, namely the (weak) Nash equilibrium. Under this notion of equilibrium there do exist mechanisms which satisfy both the budget balance and full information optimal criteria. (See, for example, the celebrated work of Groves-Ledyard and the "demand revelation" literature.) Yet here one may question how it is that the weak Nash equilibrium is to be attained; in a weak Nash equilibrium an agent's signal will vary with parameters unobserved by him. Both Groves-Ledyard [**9**] and D-H-M [**4**] have noted difficulties with adjustment schemes which are intended to circumvent this problem.

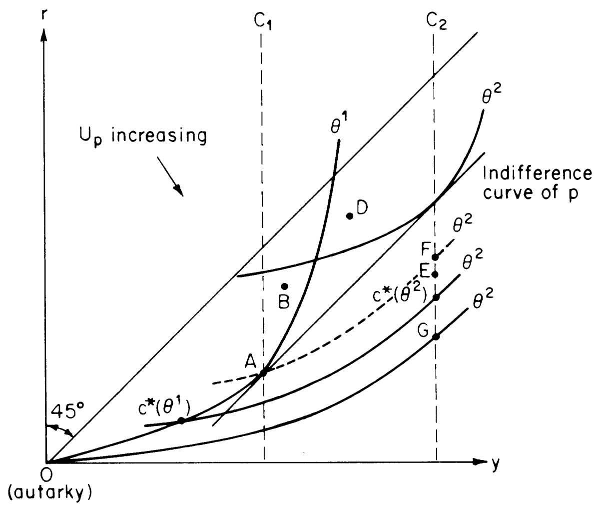

FIGURE 1.—Principal-agent environment.

of the productivity parameter θ. Such a p.c. allocation can be denoted (y_j, r_j), $j = 1, 2$, where $y_j = y(\theta^j)$ and $r_j = r(\theta^j)$. In this environment, a p.c. allocation (y_j, r_j) satisfies self selection if and only if the agent (worker) prefers the allocation (y_1, r_1) in which he must produce y_1 in return for a reward r_1 to (y_2, r_2) if, in fact $\theta = \theta^1$, and vice versa if $\theta = \theta^2$. A p.c. allocation satisfying this property is shown in Figure 1 as points $B = (y_1, r_1)$ and $D = (y_2, r_2)$. Mathematically the condition which (y_j, r_j) must satisfy to be a self-selection p.c. allocation is

$$\text{(SS)}\qquad \begin{aligned} &r_1 - (y_1/\theta^1)^2 \geqslant r_2 - (y_2/\theta^1)^2,\\ &r_2 - (y_2/\theta^2)^2 \geqslant r_1 - (y_1/\theta^2)^2. \end{aligned}$$

The first point we wish to make with this example is that if a p.c. allocation does not satisfy (SS) it cannot be achieved by any mechanism and conversely, if it does satisfy (SS) it can be achieved by a very simple mechanism. In this regard, an allocation is achieved by a mechanism if that allocation is an equilibrium allocation of the mechanism. Although the above statement is true for any mechanism, it will be easier to motivate if we restrict ourselves to direct mechanisms. In a direct mechanism, the agent is asked to name a value for θ (either θ^1 or θ^2) and then some pre-determined allocation $[y(m), r(m)]$ is effected, where m is the value of θ declared by the agent. To see the first part of the statement, suppose we have a p.c. allocation which does *not* satisfy (SS), say $(y_1, r_1) = A$, $(y_2, r_2) = B$ in Figure 1. That is, suppose we wish to achieve the allocation represented by point A (in Figure 1) when $\theta = \theta^1$ and the allocation represented by point B when $\theta = \theta^2$. Now suppose that to achieve this allocation, we choose the direct mechanism with

$$[y(\theta^1), r(\theta^1)] = A,$$
$$[y(\theta^2), r(\theta^2)] = B.$$

Note, however, that A and B do not satisfy self-selection, and in particular the agent prefers B to A regardless of the true value of θ. Thus in the direct mechanism described above, the agent will always *declare* $m = \theta^2$ even if $\theta = \theta^1$. Thus the *equilibrium* allocation of this mechanism will be B for either value of θ. In this case A will not be achieved when $\theta = \theta^1$ by this mechanism. The only other choice of a direct mechanism which has any chance of resulting in A for $\theta = \theta^1$ and B for $\theta = \theta^2$ is

$$[y(\theta^2), r(\theta^2)] = A,$$
$$[y(\theta^1), r(\theta^1)] = B.$$

In this case, the agent will always declare $m = \theta^1$, and the equilibrium allocation will again be B for either value of θ. Consequently the p.c. allocation represented by (A, B) cannot be achieved by a direct mechanism.

For the second part of the statement, suppose we have the p.c. allocation represented by $(y_1, r_1) = B$, $(y_2, r_2) = D$ in Figure 1. Consider the direct mechanism in which the final allocation as a function of the agent's declared value m of θ is given by $[y(m), r(m)] = (y_i, r_i)$ for $m = \theta^i$. Since the pair (B, D) in Figure 1 satisfies (SS), the agent will always declare the true value of θ. Thus if $\theta = \theta^i$, the equilibrium allocation of this mechanism is simply (y_i, r_i). That is, the p.c. allocation represented by (B, D) can be achieved using a simple direct mechanism in which the agent's equilibrium strategy is to "tell the truth." In Sections 3–8, these results are generalized to a large class of environments and mechanisms.

The second point we wish to illustrate in this example is that there are full information optimal p.c. allocations which are not achievable. A full information optimal p.c. allocation is simply a p.c. allocation (y_i, r_i) such that (y_1, r_1) is on C_1 and (y_2, r_2) is on C_2 and both are below the 45° line. For example, consider the p.c. allocation represented by points A and E in Figure 1. Note that the agent prefers A to E no matter the value of θ. Therefore (A, E) does not satisfy (SS) and, as indicated above, cannot be achieved.

The third point we wish to make with this example is that there are p.c. allocations which are optimal in our sense but which are not full information optimal (this will also serve to illustrate our definition of optimality in Section 8). For this environment, a feasible p.c. allocation (y_i, r_i) is Pareto optimal if (i) it satisfies (SS) and (ii) no other feasible p.c. allocation satisfying (SS) can make the principal better off (in the expected utility sense, taking expectations over θ using his prior) without making the agent worse off for at least one value of θ, and (iii) no other feasible p.c. allocation satisfying (SS) can make the agent better off for at least one value of θ without either making him worse off under the other value of θ or making the principal worse off.

Using this definition it is easy to establish that a solution of the following problem is a Pareto optimal p.c. allocation which is best for the principal subject to the constraint that the agent be no worse off for either value of θ than in autarky, when $(y, r) = (0, 0)$ (these individual rationality constraints are labeled "IR" in the problem).

$$\max_{(y_j, r_j), j=1,2} \varphi(y_1 - r_1) + (1-\varphi)(y_2 - r_2)$$

subject to (SS),

(IR) $\quad r_j - (y_j/\theta^j)^2 \geq 0 \quad \text{for} \quad j = 1, 2,$

(Feasibility)

$$0 \leq r_j \leq y_j \quad \text{for} \quad j = 1, 2.$$

In the objective function, φ is the principal's prior probability that $\theta = \theta^1$. Note that since any solution must satisfy (SS), it can be achieved by a direct mechanism as outlined above.

A typical solution (for a particular value of φ) is shown as $(y_j^*, r_j^*) = c^*(\theta^j)$, $j = 1, 2$ in Figure 1. Notice that $c^*(\theta^1)$ is not on C_1, thus showing that there are optimal p.c. allocations which are not full information optimal.

The fourth, and last, point we wish to make in this example is that optimal allocations in our sense depend on prior distributions, whereas full information optimal allocations do not (this fact plays a key role in the public goods example of Section 9). To see this imagine that the probability which the principal places on the event $\theta = \theta^1$, i.e., φ, increases. The principal will then seek to increase his utility in this event. This can be accomplished by moving $c^*(\theta^1)$ along the agent's θ^1 indifference curve toward point A. Since self-selection must hold, $c^*(\theta^2)$ must move up C_2 toward point F, i.e., p must give up utility in the event $\theta = \theta^2$. He is willing to make this tradeoff since his probability on $\theta = \theta^1$ has increased. Thus the original values of $c^*(\theta)$ are no longer optimal when the principal's prior changes.

This concludes our discussion of the agency example. We now proceed to the general development.

3. THE GENERAL ENVIRONMENT

The general economic environment consists of a finite set $T \subset R$ of agents and a collection of production technologies associated with coalitions of agents. The agents are defined by their common consumption possibility set $C \subseteq R_+^l$, where l, the number of commodities, is a fixed integer, $l \geq 2$; by their preferences; and by their initial information.

Preferences are assumed to depend on a vector, θ, of parameters. Let $N = \{1, \ldots, n\}$ be the set of parameter indices; $\Theta_k \subset R$, Θ_k finite, be the set of possible values of the kth parameter, $k \in N$, with typical element θ_k; and $\Theta = \times_{k \in N} \Theta_k$ be the parameter space with typical element θ. Each agent a is assumed to have observed values of the parameters θ_k for k in some subset $N_a \subseteq N$. The set of possible values of parameters observed by agent a is then $\Theta^a = \times_{k \in N_a} \Theta_k$, with typical element θ^a. We assume that each of the n parameters is observed by some agent, i.e. $\bigcup_{a \in T} N_a = N$. For any subset of agents $A \subset T$, let $N_A = \bigcup_{a \in A} N_a$ be the set of parameters observed by agents of A. Let $\Theta_A = \times_{k \in N_A} \Theta_k$.

The preferences of each agent are represented by a utility function U_a which depends on his consumption bundle and the parameter vector, i.e. $U_a : C \times \Theta \to R$.

We assume that each agent's beliefs concerning parameters not observed by him can be summarized by a nondegenerate prior distribution. This prior for agent a is a mapping ρ_a from $\Theta \times \Theta^a$ into the interval $[0, 1]$ where $\rho_a(\xi|\theta^a)$ is interpreted as a's prior probability that the value of the true parameter vector is ξ given that he observes θ^a.[11] Note that ξ specifies a value for the parameter vector observed by agent a; thus if $\xi^a \neq \theta^a$, $\rho_a(\xi|\theta^a) = 0$. We also assume that $\rho_a(\xi|\theta^a) > 0$ for all ξ such that $\xi^a = \theta^a$.

An essential ingredient in defining the economic environment is a specification of the allocations achievable by any coalition of agents. Thus, with any subset A of agents of T we associate a nonempty set $\lambda(A) \subseteq C^A$.[12] A typical element of $\lambda(A)$, say c_A, consists of a consumption bundle $c_a \in C$ for each agent $a \in A$. We may think of λ as a mapping which associates with each subset of agents of T its set of achievable allocations. The *technology* is a specification of λ.

We are now ready to define the economic environment. *An environment* E is a vector $[T, l, N, (N_a)_{a \in T}, \Theta, \lambda, U_T, \rho_T, \theta]$, where T is the set of agents, l is the number of goods, N is the set of parameter indices, N_a is the set of parameters observed by agent a, U_a is the utility function of agent a, ρ_a is agent a's prior, and θ is the actual value of the parameter vector. A *sub-environment* E_A of E for a subset of agents $A \subset T$ is a vector $[A, l, N_A, (N_a)_{a \in A}, \Theta_A, \lambda, U_A, \rho_A, \theta_A]$, where λ is restricted to A.

Crucial to our analysis is a specification of what each agent knows before any trade takes place, i.e., the *initial information structure.* Given an environment E, each agent in T is assumed to know everything except the values of certain parameters, i.e., everything except θ_j for $j \notin N_a$. It should be emphasized that if agent a knows θ_k while agent b does not, then both know that agent a is better informed about parameter k.

This completes our description of the general class of environments. We now turn to developing the concept of a resource allocation mechanism and proving some general results.

4. MECHANISMS

Our concept of a mechanism is simply a set of rules which define an extensive form game to be played by the agents. These rules are constraints on what

[11] For most of our analysis the parameters *may* be thought of as being drawn from some exogenously specified joint distribution known to all agents with agents' priors being conditional distributions given their observed parameters. We do not require this consistency, however.

[12] In general, we use the following notation for Cartesian products of sets. If $A \subseteq T$ and $\{Y_a | a \in A\}$ is any collection of sets, then we denote the product $\times_{a \in A} Y_a$ by Y^A. An element of Y^A is denoted by $(y_a)_{a \in A}$ or by y_A.

proposals the agents may make at each stage of a multistage bargaining process and a rule for determining the final allocation as a function of the sequence of these proposals.[13]

The elements of a *mechanism* for a sub-environment E_A for a set of agents $A \subseteq T$ with technology λ (or, for short, a mechanism for A) are as follows. First is an integer, τ, which is interpreted as the number of stages in the mechanism, i.e., signals are sent by the agents sequentially in stages $1, 2, \ldots, \tau$. At each stage, each agent is assumed to have observed all previous signals. Second is a set $\mathcal{S}$ which is the set of all potential signals and an element s^0 which is the null signal (hereafter taken to be the word "pass") introduced mainly for notational convenience. The mechanism is assumed to begin with all agents sending the null signal (at stage 0). Thereafter an agent may send the null signal only if required to do so. Thus observing the null signal provides no new information. Third is a sequence $S_A = \{S_{at} | a \in A,\ t = 1, \ldots, \tau\}$ of constraint sets which define the set of feasible signals for each agent at each stage. Each of these subsets of $\mathcal{S}$ may depend on all past signals, i.e. each S_{at} is a correspondence which associates with each sequence of past signals, s_A^{t-1}, a subset $S_{at}(s_A^{t-1})$ of the signal space $\mathcal{S}$. Fourth is a function F which determines the final allocation in $\lambda(A)$ as a function of the entire sequence of signals, s_A^{τ}.[14]

A mechanism for a set of agents A will be denoted by M_A or by an explicit listing $[\tau, S_A, F]$ of its elements (where $\mathcal{S}$ can be taken to be any set containing all signals consistent with S_A). Note that the only aspect of the environment on which a mechanism depends is the technology, λ.

We denote by $\mathcal{M}_A$ the *class of all mechanisms* for the set of agents A.

5. EQUILIBRIUM OF A MECHANISM

In this section we define the concept of an equilibrium for a mechanism. The concept we adopt is closely related to Harsanyi's [**13**] Bayesian equilibrium and Selten's [**27**] perfect equilibrium; consequently we call it perfect Bayesian equilibrium. Speaking loosely, we use the term equilibrium of a mechanism to mean a specification of a strategy and a posterior distribution for each agent at each stage of the mechanism such that two properties hold: (i) the specified strategy of each agent at each stage is maximal for him relative to his specified contemporary

[13] The definition of a mechanism used here is a straight-forward extension of the one used in our earlier paper (Harris-Townsend [**12**]). The reader is referred to that paper for a fuller discussion of the issues involved in this concept and some examples which motivate the definition.

[14] The formal definition is as follows: Given $A \subseteq T$ and λ_A a *mechanism* for the set of agents A, M_A, is: (i) an integer $\tau \geq 1$; (ii) a set, $\mathcal{S}$, a designated element, $s^0 \in \mathcal{S}$, and a sequence

$$S_A = \{S_{at} : (\mathcal{S}^A)^{t-1} \to 2^{\mathcal{S}} | a \in A, t = 1, \ldots, \tau\} \quad \text{where}$$

$$(\mathcal{S}^A)^t = \{s_A^0\} \times \bigtimes_{r=1}^{t} \mathcal{S}^A \text{ for } t = 1, \ldots, \tau,\ s_A^0 = (s^0)_{a \in A}$$

and $(\mathcal{S}^A)^0 = \{s_A^0\}$. If $s^0 \in S_{at}(s_A^{t-1})$ for some a, t, and s_A^{t-1}, then $S_{at}(s_A^{t-1}) = \{s^0\}$; (iii) a function, $F: (\mathcal{S}^A)^{\tau} \to \lambda(A)$.

posterior distribution, taking as given the past signals, specified contemporary strategies of the other agents, and specified future strategies of all agents, and (ii) the specified posterior distribution of each agent at each stage is consistent with Bayes' rule and the specified past strategies of all other agents.

We now make this definition more precise. For what follows we consider a given mechanism $[\tau, S_A, F]$, for a set of agents $A \subseteq T$. A *strategy* at stage t for each agent a is a function σ_{at} mapping previous signals, s_A^{t-1}, and the vector of parameters observed by agent a, θ^a, into a signal in $\mathscr{S}$.[15]

It is clear that a specification of strategies for each agent of A determines entirely the sequence of signals which will be sent under the given mechanism, as a function of the parameters observed by agents of A. In particular, given some stage r, a sequence of future strategies $\sigma_A^{r+1,\tau}$ from stages $r+1$ through τ for each agent of A, a sequence of past and present signals s_A^r, and a parameter draw θ_A, we may calculate the sequence of induced signals from stages $r+1$ through τ (see the Appendix and footnote 16 below for a more formal development).

Now consider the decision problem confronting agent a at stage r. Agent a takes as given the past signals s_A^{r-1}, the contemporary strategies of others, $(\sigma_{br}(s_A^{r-1}, \xi^b))$, $b \in A-\{a\}$, and the future strategies of all agents, $\sigma_A^{r+1,\tau}$. He may then calculate the contemporary signals of the other agents and all future signals *as functions of the unobserved* (*by agent a*) parameters and his own present signal. Finally, using this calculation, agent a may calculate his expected utility for the final allocation determined by this sequence of signals as a function of his current signal, i.e., $E_{ar}\{U_a[F_a(s_A^\tau), \xi]|s_A^{r-1}, \sigma_{\bar{a}r}, \sigma_A^{r+1,\tau}, \theta^a\}$, where F_a is the allocation assigned by F to agent a, and the expectation is taken using his *current* beliefs about ξ, $\rho_{ar}(\xi|\theta^a, s_A^{r-1})$.[16] Here s_A^{r-1}, $\sigma_{\bar{a}r}=(\sigma_{br})$, $b \in \bar{a} = A-\{a\}$, and $\sigma_A^{r+1,\tau}$ are taken as given. Agent a then chooses his current signal, s_{ar}, to maximize this expected utility over the set of feasible signals at stage r, $S_{ar}(s_A^{r-1})$.

Agent a's current beliefs about ξ are, however, not arbitrary. Rather, agent a takes as given the strategy of every other agent b at each previous stage j, $\sigma_{bj}(s_A^{j-1}, \delta^b)$, and eliminates from consideration any parameter vector δ which is not consistent with previously observed signals. The posterior at any stage r is thus

[15] More formally we have *a strategy at stage t for agent a*, $1 \le t \le \tau$, $a \in A$, which is a function $\sigma_{at}: (\mathscr{S}^A)^{t-1} \times \Theta^a \to \mathscr{S}$. We assume that strategies and the allocation rule, F, are deterministic. In general this is inconsistent with the spirit of our approach, because the constraint that a proposal not be chosen in a random way is, in general, unenforceable. In many applications, however, this restriction is not binding, i.e., any mechanism whose equilibrium final allocations are random, because either the strategies or the allocation rule are random, can be dominated by one in which they are deterministic (see Harris and Townsend [**12**]).

[16] Note the recursive procedure used to define $\Psi_A^\tau(\theta_A|\sigma_A^\tau)$ in the Appendix. Then starting at stage r, given past signals s_A^{r-1}, one may similarly define $\Psi_A^{r,\tau}(\theta_A, s_{ar}|s_A^{r-1}, \sigma_{\bar{a}r}, \sigma_A^{r+1,\tau})$, the sequence of signals which will be sent under the given mechanism as a function of the parameter vector θ_A and the contemporary signal s_{ar} of agent a, *given* past signals, contemporary strategies of others, and future strategies of everyone. Thus

$$E_{ar}\{U_a[F_a(s_A^\tau), \xi]|s_A^{r-1}, \sigma_{\bar{a}r}, \sigma_A^{r+1,\tau}, \theta^a\}$$

$$= \sum_{\xi \in \Theta} \rho_{ar}(\xi|s_A^{r-1}, \theta^a) U_a\{F_a[s_A^{r-1}, \Psi_A^{r,\tau}(\xi_A, s_{ar}|s_A^{r-1}, \sigma_{\bar{a}r}, \sigma_A^{r+1,\tau}), \xi\}.$$

the initial prior normalized in this way. We are thus led to the following formal definition of an equilibrium.[17]

DEFINITION 1: Let A be any subset of agents, and let M_A be any mechanism for A. A *perfect Bayesian equilibrium* of M_A is a sequence of strategies $\sigma_A^{*\tau} = \{\sigma_{Ar}^*|r=1,\ldots,\tau\}$ for each agent in A, and a set of distribution functions $\rho_A^{*\tau} = \{\rho_{Ar}^*|r=1,\ldots,\tau\}$ for each agent in A, where $\rho_{ar}^*: \Theta \times \Theta^a \times (\mathscr{S}^A)^{r-1} \to [0,1]$, which satisfy the following two properties:

(i) For any agent $a \in A$, any stage $r = 1, \ldots, \tau$, any parameter vector $\theta^a \in \Theta^a$ and history s_A^{r-1}, and given strategies $\sigma_A^{*\tau}$, let $\Delta_{ar}(s_A^{r-1})$ denote the set of parameter values which are consistent with the observed past signals s_A^{r-1} of agents other than a, given those agents' equilibrium strategy functions. That is,

$$\Delta_{ar}(s_A^{r-1}) = \{\xi \in \Theta | \text{for every } b \in \bar{a}, \text{ and } 1 \leq j \leq r-1, \sigma_{bj}^*[s_A^{j-1}, \xi^b] = s_{bj}\}.$$

Then let $K_{ar}(s_A^{r-1}, \theta^a)$ denote the total prior probability assigned by a to such parameter values given that he observes values θ^a for parameters he can observe. That is,

$$K_{ar}[s_A^{r-1}, \theta^a] = \sum_{\delta \in \Delta_{ar}(s_A^{r-1})} \rho_a(\delta|\theta^a).$$

Then $\rho_{ar}^*(\xi|\theta^a, s_A^{r-1})$ is assumed to satisfy

$$\rho_{ar}^*(\xi|\theta^a, s_A^{r-1}) = \begin{cases} \rho_a(\xi|\theta^a)/K_{ar}(s_A^{r-1}, \theta^a); & \xi \in \Delta_{ar}(s_A^{r-1}) \neq \varnothing; \\ 0; & \xi \notin \Delta_{ar}(s_A^{r-1}) \neq \varnothing; \\ \rho_a(\xi|\theta^a); & \Delta_{ar}(s_A^{r-1}) = \varnothing. \end{cases}$$

(ii) For any agent $a \in A$, stage $r = 1, \ldots, \tau$, parameter vector $\theta^a \in \Theta^a$ and history s_A^{r-1}, the strategy $\sigma_{ar}^*(s_A^{r-1}, \theta^a)$ is assumed to solve

$$\max_{s_{ar} \in S_{ar}(s_A^{r-1})} E_{ar}^*\{U_a[F_a(s_A^\tau), \xi]|s_A^{r-1}, \sigma_{\bar{a}r}^*, \sigma_A^{*r+1,\tau}, \theta^a\},$$

where the expectation E_{ar}^* is taken with respect to $\rho_{ar}^*(\xi|\theta^a, s_A^{r-1})$ taking $\sigma_{\bar{a}r}^*$ and $\sigma_A^{*r+1,\tau}$ as given.

Thus our equilibrium concept is Bayesian in the sense that each agent evaluates strategies and outcomes relative to that agent's prior (or posterior) distribution over unobserved parameters, and perfect in the sense that strategies are required to be optimal *for all possible previous histories*.

In general, a mechanism may have more than one equilibrium. *In what follows, we assume that one particular equilibrium prevails and refer to "the equilibrium" of a mechanism.* For the specific mechanisms considered below, the equilibrium assumed to prevail will be clear from the context.

[17] Note that in the definition, $\rho_{ar}^*(\xi|\theta^a, s_A^{r-1}) = \rho_a(\xi|\theta^a)$ for any history s_A^{r-1} which is inconsistent with the equilibrium strategies of agents other than a, evaluated at any possible parameter values, i.e., for s_A^{r-1} for which $\Delta_{ar}(s_A^{r-1}) = \varnothing$. Also note that for $r = 1$, $\rho_{a1}^*(\xi|\theta^a, s_A^0) = \rho_a(\xi|\theta^a)$ for any ξ and θ^a and $\Sigma_{\xi \in \Theta}\, \rho_{ar}^*[\xi|\theta^a, s_A^{r-1}] = 1$ for any θ^a, s_A^{r-1}.

Given any mechanism M_A for a set of agents A, the equilibrium strategies for M_A determine an equilibrium sequence of signals, $s_A^{*\tau}(\theta_A)$, for each possible value θ_A of the vector of parameters observed by agents of A. This equilibrium sequence of signals, in turn, determines the *equilibrium allocation*, $F[s_A^{*\tau}(\theta_A)]$ for M_A as a function of θ_A, which we denote by $c_A(M_A, \theta_A)$. If a mechanism has multiple equilibria, we could denote each one explicitly by including the equilibrium sequence of strategies as an argument in c_A. This notation would allow us to make explicit which equilibrium sequence of strategies is assumed to prevail. This would, however, make the notation more cumbersome. In what follows we will rarely need to make explicit which equilibrium sequence of strategies is assumed to prevail. We have therefore chosen to suppress the equilibrium strategies in the notation.

At this point the reader may well wonder why we have not restricted our analysis a priori to games in normal as opposed to extensive form and used the (regular) Bayesian equilibrium notion. This would simplify the notation and proofs considerably. Our answer is that games which are sequential and use the perfect equilibrium notation are of intrinsic interest and seem natural in many economic contexts.[18] Thus, for us, the real question is whether games in extensive form with the perfect Bayesian equilibrium notation are equivalent in some sense to games in normal form with the (regular) Bayesian equilibrium. In this regard we agree with Selten [**27**] that "... in the transition from the extensive form to the normal form some important information is lost." In particular, in the asymmetric information environments of this paper, the (regular) Bayesian equilibrium concept for normal form games allows neither an *explicit* representation of learning nor the concept of perfectness. That the latter concept has content is well illustrated by noting that the perfect Bayesian equilibrium for our extensive form games admits far fewer equilibria than does the (regular) Bayesian equilibrium for their normal form counterparts.[19]

6. SELF-SELECTION THEOREMS

In this section we demonstrate two closely related results. The first result characterizes allocations which result from mechanisms. It is shown that the equilibrium allocation of any mechanism must satisfy certain self-selection properties.[20] The second result is that, conversely, any allocation which satisfies these self-selection properties can be achieved by a mechanism.

[18] See, for example, Kydland and Prescott [**17**], Prescott and Visscher [**20**], and E. Green [**6**].

[19] We are indebted to Jim Jordan for providing us with an example. The example is omitted for the sake of brevity but is obtainable from the authors on request. It may well be, however, that every perfect Bayesian equilibrium of an extensive form game (mechanism) is one of the many Bayesian equilibria of its normal form counterpart. We might well have set out to prove this so that the self-selection theorem, Theorem 1, would be more easily established (for normal form games). Here we hazard the opinion that such a proof would be similar in many respects to the proof of Theorem 1. In any event we would still need to consider extensive form games and endure the consequent notational burden.

[20] The concept of self-selection has been used extensively in insurance and screening literatures (see, e.g., Rothschild-Stiglitz [**25**], C. Wilson [**31**], Salop and Salop [**26**]). This concept is closely related to incentive compatibility as discussed by Hurwicz [**15**].

These results are important and require some elaboration. First, note that mechanisms result in parameter-contingent (p.c.) allocations, i.e., allocations which depend on the realized values of the parameters. This would suggest that if one is interested in characterizing optimal allocations for this class of environments, one could ignore the mechanisms by which these allocations are achieved and simply search over the set of all technically feasible p.c. allocations. Our first result shows that certain technically feasible p.c. allocations, namely those that do not satisfy the self-selection properties, cannot, in fact, be achieved. This implies that, at least, one must eliminate from consideration all p.c. allocations which do not satisfy these properties. Our second result then shows that only these allocations need be eliminated. This latter result is established by showing that any p.c. allocation satisfying the self-selection properties can be achieved by a mechanism of a *certain, simple form*. This result itself is important since it implies that the class of mechanisms that need be considered can be reduced to those of this form.

Having emphasized the crucial importance of self-selection properties, it remains to define them precisely and show why they arise. To motivate these, consider an agent, say a, who has private information about, say, θ_1 (i.e., only this agent observes the value of θ_1). Whatever mechanism is used, say M, agent a may behave as though $\theta_1=\delta$ even if he has observed θ_1 to be γ (γ and δ are any two elements of Θ_1). That is, he may use the functional form σ_a^* of the equilibrium strategy to generate signals, but evaluate this strategy at the counterfactual realization $\theta_1=\delta$ instead of at $\theta_1=\gamma$. This strategy is an alternative to the equilibrium strategy. Since agent a is the only agent who knows θ_1, this alternative strategy will generate the *equilibrium* allocation for the *counterfactual* $\theta_1=\delta$, namely $c_A(M, \delta, \theta_2, \ldots, \theta_n)$. Of course equilibrium strategies are preferred (maximizing) so agent a must prefer the allocation $c_A(M, \gamma, \theta_2, \ldots, \theta_n)$ where his parameter draw is $\theta_1=\gamma$.

It should be pointed out that if more than one agent observes a parameter, i.e., if the information is public, then the above argument does not hold. This is due to the fact that the adoption of the Bayesian equilibrium concept for mechanisms precludes collusion. In general, the existence of public information does not constrain the allocations which may be achieved by mechanisms (this point is illustrated by example in Section 10).

We now proceed with a formal treatment of self-selection. For the rest of this section let A be a subset of agents, let M_A be a mechanism for A with $M_A=[\tau, S_A, F]$, and let $(\sigma_A^{*\tau}, \rho_A^{*\tau})$ be an equilibrium of M_A.

Parameters observed only by one agent are private to that agent, and as noted above, play a special role in the analysis. Accordingly, for any vector of parameters θ and for any agent a, we write $\theta=(\theta_P^a, \theta_H^a)$, where θ_P^a consists of those components of θ which *only a* observes and θ_H^a consists of the remaining components of θ. The vector θ_P^a is called the *private parameter vector of agent a.*

The first main result of this section (discussed above) is as follows:

THEOREM 1 (Self-Selection for Mechanisms): *For any mechanism M_A, any* *agent a, and any value of agent a's observed parameters* θ^a,

$$\sum_{\xi \in \Theta} U_a[c_a(M_A, \xi_A), \xi]\rho_a(\xi|\theta^a) \geq \sum_{\xi \in \Theta} U_a[c_a(M_A, \delta_P^a, \xi_H^a), \xi]\rho_a(\xi|\theta^a)$$

for any other possible value of agent a's private parameter vector, δ_P^a.

PROOF: See Appendix.

As discussed above, a mechanism for the set of agents A results in a p.c. allocation which is feasible relative to the technology λ, i.e., mappings $\tilde{c}_A: \Theta_A \to \lambda(A)$. Clearly, we can discuss such p.c. allocations regardless of whether they are the result of some mechanism (hereafter the term p.c. allocation will be taken to mean a p.c. allocation which is feasible relative to the appropriate technology). Accordingly, for an arbitrary p.c. allocation, $\tilde{c}_A$, for any agent a and any value of his observed parameters θ^a, define a's expected utility for $\tilde{c}_a$ by

$$V_a(\tilde{c}_a, \theta^a) = \sum_{\xi \in \Theta} U_a[\tilde{c}_a(\xi_A), \xi]\rho_a(\xi|\theta^a).$$

The p.c. allocation $\tilde{c}_A$ is said to *satisfy self-selection* (SS) if for every $a \in A$,

$$V_a(\tilde{c}_a, \theta^a) \geq \sum_{\xi \in \Theta} U_a[\tilde{c}_a(\delta_P^a, \xi_H^a), \xi]\rho_a(\xi|\theta^a)$$

for any θ^a and any alternative value of a's private parameter vector δ_p^a. Thus Theorem 1 can be interpreted as saying that all p.c. allocations achievable by a mechanism must satisfy SS. (An allocation is achievable by a mechanism if it is an equilibrium allocation of the mechanism.) Theorem 2 (below) states that the converse is also true, i.e., any p.c. allocation which satisfies SS is achievable by a mechanism. Moreover, the mechanism which can be used to achieve the allocation (referred to as the direct mechanism, D, below) is of the following simple form: there is only one stage in which each agent reports a value for each parameter he observes; the final allocation is given by a pre-specified function of the values reported by each agent. If all information is private, this allocation function is the p.c. allocation to be achieved.

More formally, *a direct mechanism* is any mechanism in which $\tau = 1$, $S_{a1}(s_A^0) = \Theta^a$, for each agent a.

THEOREM 2: *For any p.c. allocation* $\tilde{c}_A$ *which satisfies SS, there exists a direct mechanism* D_A *such that* $\tilde{c}_a(\theta_A) = c_A(D_A, \theta_A)$ *for any parameter vector* θ_A *and such that the equilibrium strategies under* D_A *are truth-telling. Furthermore, if all parameters are privately observed, i.e.,* $\theta^a = \theta_P^a$ *for each agent a, then the allocation rule F for* D_A *can be assumed to be the same as the p.c. allocation* $\tilde{c}_A$.

PROOF: See Appendix.

For the case in which all parameters are privately observed, the direct mechanism D_A of Theorem 2 is *Bayesian incentive compatible* in the sense of d'Aspremont and Gérard-Varet [**5**]. Thus the SS condition generalizes this notion for mechanisms to environments with public information.

7. PREFERENCES OVER MECHANISMS AND EFFICIENCY

We assume that at the time when each agent a must express a preference over mechanisms he already knows the value of the parameter vector θ^a. Thus preferences over mechanisms are defined using expected utilities, based on the initial information structure, for the final allocation of an equilibrium of the mechanism. Again if the equilibrium is not unique, one particular equilibrium must be selected. The definitions below rely heavily on this assumption.[21]

More precisely, given a set of agents A, an environment E_A, and a mechanism M_A with equilibrium $(\sigma_A^{*\tau}, \rho_A^{*\tau})$, the expected utility of agent $a \in A$ for the mechanism M_A is defined as

$$W_a[M_A, \theta^a] = \sum_{\xi \in \Theta} U_a\{c_a[M_A, \xi], \xi\}\rho_a(\xi|\theta^a).$$

Given these preferences for mechanisms, we can define the concepts of blocking a mechanism and core mechanism. *A coalition $B \subseteq A$ can block a mechanism M_A* if there exists a mechanism M_B' such that $W_b[M_B', \xi^b] \geq W_b[M_A, \xi^b]$ for every $b \in B$, $\xi^b \in \Theta^b$ with strict inequality for some $b \in B$, $\xi^b \in \Theta^b$. Note that in this definition, we require that the expected utility for M_B' for each b be at least as great as his expected utility for M_A *for each possible draw of b's observed parameters.*[22] The reason for this specification is to ensure that efficient mechanisms (defined below) will be independent of the actual parameter draw. That is, the mechanism which we denote as efficient may well depend on the technology, preferences, and general information structure, but we do not want our selection to depend on particular values of parameters. The elements of the environment on which we allow dependence will be referred to as the partial environment. More precisely, the *partial environment* relative to E_A, denoted E_A^0, includes all elements of E_A except the specific parameter draw, θ_A.

There are three concepts of *efficient mechanisms* which we wish to consider: First, define a mechanism M_A^* for A to be a *core mechanism* (relative to the partial environment E_A^0) if no coalition B can block M_A^*. Second, define a mechanism M_A^* to be a *Pareto optimal mechanism* (relative to E_A^0) if A cannot block M_A^*.

[21] An alternative approach would be to define blocking (see below) in terms of the equilibrium allocations of mechanisms instead of the mechanisms themselves. Then we could define a core mechanism to be one with at least one equilibrium which could not be blocked. The results which follow would not be affected by this alteration of the definition.

[22] An alternative formulation would be to suppose that preferences over mechanisms are expressed prior to the revelation of θ^a to agent a, but the mechanism must be "played out" (as before) subsequent to this revelation. In this case each agent may be imagined to have a prior distribution over the entire parameter vector (say a common joint distribution). Expected utility of each agent for a mechanism would then involve integration relative to this prior.

Third, define a mechanism M_A^* to be an *individually rational Pareto optimal mechanism* (relative to E_A^0) if M_A^* is a Pareto mechanism and no single agent in A can block M_A^*. We use the term "efficient mechanism" to refer to any mechanism which satisfies at least one of the above efficiency concepts.

8. EFFICIENT MECHANISMS AND EFFICIENT ALLOCATIONS

The results of Section 6 suggest that there is an equivalence between p.c. allocations which satisfy SS and allocations generated by mechanisms. This would lead one to conjecture that the search for efficient mechanisms can be accomplished by searching for p.c. allocations which are efficient in some sense. In this section we define a notion of efficient p.c. allocations and establish the equivalence between such allocations and allocations generated by efficient mechanisms.

For this section E_A^0 is any fixed partial environment. *A coalition $B \subseteq A$ can block a p.c. allocation $\tilde{c}_A$* satisfying SS if there is a p.c. allocation for B, $\tilde{x}_B$, such that $\tilde{x}_B$ satisfies SS and $V_b(\tilde{x}_b, \theta^b) \geq V_b(\tilde{c}_b, \theta^b)$ for every $b \in B$, $\theta^b \in \Theta^b$ with strict inequality for at least one $b \in B$, $\theta^b \in \Theta^b$. There are again three concepts of efficiency we consider. A p.c. allocation $\tilde{c}_A$ which satisfies SS is a core allocation if for any $B \subseteq A$, B cannot block $\tilde{c}_A$, a *Pareto optimal allocation* if A cannot block $\tilde{c}_A$, and an *individually rational Pareto optimal allocation* if $\tilde{c}_A$ is a Pareto optimal allocation and for each $a \in A$, $\{a\}$ cannot block $\tilde{c}_A$.

THEOREM 3: *If $\tilde{c}_A$ is a core allocation, then there is a direct core mechanism M_A such that the equilibrium allocation of M_A coincides with $\tilde{c}_A$, i.e. $\tilde{c}_A(\theta_A) \equiv c_A(M_A, \theta_A)$ for all θ_A. Conversely, if M_A is a core mechanism, then $c_A(M_A, \cdot)$ is a core allocation. The theorem also holds if "core" is replaced by "Pareto optimal" or "individually rational Pareto optimal."*

PROOF: See Appendix.

Theorem 3 has the important implication that one need only characterize efficient allocations in order to characterize efficient mechanisms. *This is useful because an efficient allocation can often be found as the solution of a constrained optimization problem.* Moreover Theorem 3 implies that the direct mechanism (in which agents signal values of parameters observed by them) is always efficient if its allocation function, F, is constructed to generate an efficient allocation. Thus the problem of finding an efficient mechanism can be reduced to finding an efficient allocation, then constructing F to generate this allocation. In the special case in which all information is private, F can be taken to be the efficient p.c. allocation. A method for constructing F in other cases is given in the proof of Theorem 2.

We would like to conclude this section by defining full information (FI) optimal allocations in general and stating a limited result on their relation to optimal allocations in our sense. Intuitively, a p.c. allocation is FI optimal if it would be Pareto optimal in the usual sense for any value of the parameter vector. More formally, a p.c. allocation $\tilde{c}_A$ for a set of agents A is *FI optimal* if for any parameter value θ there is no other feasible allocation c_A such that

$$U_a[c_a, \theta] \geqslant U_a[\tilde{c}_a(\theta_A), \theta]$$

for each agent $a \in A$ with strict inequality for at least one $a \in A$.

Our result is that if there are only two agents and one is fully informed while the other is not, then any FI optimal allocation which satisfies the self-selection condition is also optimal in our sense. This is stated formally as Theorem 4.

THEOREM 4: *If $T = \{a, b\}$, $N_a = N$, $N_b \neq N$, and $\tilde{c}_T^*$ is a FI optimal p.c. allocation which satisfies SS, then $\tilde{c}_T^*$ is a Pareto optimal p.c. allocation.*

PROOF: See Appendix.

Note that this theorem applies to the principal-agent example of Section 2. Thus, in that example, while there are optimal p.c. allocations which are *not* FI optimal, any FI optimal allocation which satisfies SS is also optimal in our sense. That the result cannot be generalized to more complicated information structures is shown by example in the section on public goods which follows.

9. A PUBLIC GOODS ENVIRONMENT

Here we describe one of the simplest public goods environments we could think of and the results of applying the general approach of this paper to that environment. There are two agents labeled a and b, one private good and one public good. The public good can be produced on a one-for-one basis from the private good, but only integer amounts can be produced.[23] Each agent has an endowment of $e < 1$ units of the private good and no public good. Thus neither agent can produce any public good on his own, and together they can produce either zero or one unit. Each agent i has a linear utility function of the form $U_i(x_i, y, \theta_i) = \theta_i x_i + y$, where x_i represents consumption of the private good, y consumption of the public good, and θ_i is i's marginal utility for the private good, $i = a, b$.

Initially, each agent knows only his own marginal utility for the private good, i.e., agent i knows θ_i but not θ_j for $j \neq i$. Moreover, each agent knows that the other agent's parameter (marginal utility) can be one of two possible values, θ^1 or θ^2 with $0 < \theta^1 \leqslant 1$, $\theta^2 > 2$ (the reason for this assumption will be clear shortly). Agent i has a prior probability distribution on the value of agent j's parameter which may depend on his own parameter value, θ_i. We denote by α_k agent a's prior probability that $\theta_b = \theta^1$, given that $\theta_a = \theta^k$ for $k = 1, 2$. Naturally, agent a's prior that $\theta_b = \theta^2$, given $\theta_a = \theta^k$ is $1 - \alpha_k$. Similarly, β_k is agent b's prior that

[23] This integer assumption is made only to avoid the subtle issue of how to define "blocking" and the core in a public goods environment. This issue is considered by Rosenthal [**23, 24**], Champsaur, Roberts, and Rosenthal [**3**], and others. It is peripheral to the points we wish to make here.

$\theta_a = \theta^1$, given that $\theta_b = \theta^k$ for $k = 1, 2$.[24] Thus, typically, knowledge of one's own tastes for public and private goods allows one to infer something about the tastes of the other agent. For example if α_1 were near 1 and α_2 were near 0, this would reflect agent a's belief that agent b is likely to be similar in tastes to himself. We assume that $0 < \alpha_k < 1$ and $0 < \beta_k < 1$ for $k = 1, 2$. Finally, each agent knows the production technology for the public good, the endowments of both agents, and the utility functions of both agents (except for the parameter value of the other agent).

This environment is very similar to that studied by Arrow [**1**] and d'Aspremont and Gerard-Varet [**5**], but, as will be seen below, our results are somewhat different from theirs. Applying Theorem 1 (Section 6) we note first that only p.c. allocations which satisfy certain self-selection (SS) conditions are achievable (as equilibrium final allocations of mechanisms). For this environment there are two parameters, θ_a and θ_b, where $\theta_i \in \Theta_i = \{\theta^1, \theta^2\}$ for $i = a, b$. Thus, we can represent a p.c. allocation by (y_{ij}, a_{ij}, b_{ij}), $i, j = 1, 2$, where y_{ij} is the total output of public good if $\theta_a = \theta^i$ and $\theta_b = \theta^j$, a_{ij} is agent a's allocation of private good if $\theta_a = \theta^i$ and $\theta_b = \theta^j$, and b_{ij} is agent b's allocation of private good if $\theta_a = \theta^i$ and $\theta_b = \theta^j$. For convenience, we define Y to be the 2×2 matrix whose (i, j) entry is y_{ij} and similarly for A and B. The SS conditions state that if $\theta_a = \theta^1$, agent a's expected utility for (y_{1j}, a_{1j}), $j = 1, 2$ (where the expectation is taken with respect to a's prior on θ_b given $\theta_a = \theta^1$, i.e., using the probabilities α_1 and $1 - \alpha_1$) must be greater than his expected utility for (y_{2j}, a_{2j}) and vice-versa if $\theta_a = \theta^2$. Also included in (SS) are similar conditions for agent b. Using the matrix notation mentioned above, a p.c. allocation $[Y, A, B]$ satisfies SS if

$$\text{(SS.1)} \quad (1, 0) \cdot [Y + \theta^1 A]\begin{pmatrix} \alpha_1 \\ 1 - \alpha_1 \end{pmatrix} \geqslant (0, 1) \cdot [Y + \theta^1 A]\begin{pmatrix} \alpha_1 \\ 1 - \alpha_1 \end{pmatrix},$$

$$\text{(SS.2)} \quad (0, 1) \cdot [Y + \theta^2 A]\begin{pmatrix} \alpha_2 \\ 1 - \alpha_2 \end{pmatrix} \geqslant (1, 0) \cdot [Y + \theta^2 A]\begin{pmatrix} \alpha_2 \\ 1 - \alpha_2 \end{pmatrix},$$

$$\text{(SS.3)} \quad (\beta_1, 1 - \beta_1) \cdot [Y + \theta^1 B]\begin{pmatrix} 1 \\ 0 \end{pmatrix} \geqslant (\beta_1, 1 - \beta_1) \cdot [Y + \theta^1 B]\begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$\text{(SS.4)} \quad (\beta_2, 1 - \beta_2) \cdot [Y + \theta^2 B]\begin{pmatrix} 0 \\ 1 \end{pmatrix} \geqslant (\beta_2, 1 - \beta_2) \cdot [Y + \theta^2 B]\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Second, we note using Theorem 2 (Section 6) that any p.c. allocation satisfying SS, say $[Y, A, B]$, can be achieved using a direct mechanism. In a direct mechanism for this environment, each agent names a value for his parameter from

[24] In order that these priors be consistent, i.e., be conditional distributions derived from the same joint distribution on (θ_a, θ_b), it must be the case that

$$\frac{\alpha_2(1 - \alpha_1)}{\alpha_1(1 - \alpha_2)} = \frac{\beta_2(1 - \beta_1)}{\beta_1(1 - \beta_2)}.$$

Independence of priors (a stronger condition than consistency) would require that $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$. In general, we do *not* assume consistency or independence of these priors.

the set $\{\theta^1, \theta^2\}$ (not necessarily the true value). This is done by both agents simultaneously. The two messages, say m_1 and m_2, are then mapped into a final allocation by a function $F(\cdot, \cdot)$. The allocation $[Y, A, B]$ can be achieved simply by setting $F(\theta^i, \theta^j) = (y_{ij}, a_{ij}, b_{ij})$.

Finally, it follows from Theorem 3 (Section 8) that a Pareto optimal p.c. allocation and a Pareto optimal mechanism can be found by solving

$$\max_{[Y,A,B]} \lambda_1(1,0)\cdot[Y+\theta^1 A]\begin{pmatrix}\alpha_1\\1-\alpha_1\end{pmatrix}+\lambda_2(0,1)\cdot[Y+\theta^2 A]\begin{pmatrix}\alpha_2\\1-\alpha_2\end{pmatrix}$$

$$+\lambda_3(\beta_1, 1-\beta_1)\cdot[Y+\theta^1 B]\begin{pmatrix}1\\0\end{pmatrix}+\lambda_4(\beta_2, 1-\beta_2)\cdot[Y+\theta^2 B]\begin{pmatrix}0\\1\end{pmatrix}$$

subject to (SS.1)–(SS.4) and

$$Y+A+B \leqslant E = \begin{bmatrix}2e & 2e\\2e & 2e\end{bmatrix},$$

$$A, B \geqslant 0, \quad y_{ij} \in \{0, 1\}, \quad i, j = 1, 2.$$

In this problem, the objective function is a weighted sum (λ_i's are the weights) of the expected utilities of each agent for each possible marginal utility (θ_i). This guarantees that no agent can be made better off for some possible value of his parameter without making him or another agent worse off for some parameter value. The constraints are the self-selection conditions (SS.1)–(SS.4) and the technological feasibility constraints. In the remainder of this section, we consider the relationship between these Pareto optimal allocations (i.e., solutions of the above problem) and full-information Pareto optimal allocations.

Among the full-information (FI) optimal allocations for this environment are those that satisfy the familiar Samuelson-Lindahl (SL) conditions. Because of the linearity of utility functions and the production technology, these SL conditions are that the public good should be built ($y_{ij} = 1$) if and only if the sum of the marginal rates of substitution of private for public good exceed the marginal cost of building the public good (namely 1), i.e.,

$$y_{ij} = 1 \quad \text{if and only if} \quad 1/\theta^i + 1/\theta^j > 1.$$

These allocations are then FI optimal *for any feasible allocation of the private good in this environment.* Given our assumption that $0 < \theta^1 \leqslant 1$, $\theta^2 > 2$, these SL conditions reduce to $y_{11} = y_{12} = y_{21} = 1$, $y_{22} = 0$, i.e., any feasible p.c. allocation $[Y, A, B]$ is an SL allocation so long as

$$Y = \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}.$$

In addition to these Samuelson-Lindahl (or SL) allocations, it is sometimes optimal not to build the public good even when $1/\theta^i + 1/\theta^j > 1$. Such an allocation is optimal when the agent with low value of θ (high marginal utility for public good) has too little of the private good to compensate the other agent. Consistent with much of the literature on public goods, however (e.g., Groves-Ledyard [**9**] and Brock [**2**]) we shall focus attention on SL allocations, rather than on such corner solutions.

We would like to use the remainder of this section to focus on a question which is not addressed in the other examples, namely, is it possible for an SL allocation to be achievable but *not* Pareto optimal in our sense? The answer is yes, thus suggesting that SL allocations, even when achievable, may not be desirable from a normative viewpoint. From the positive viewpoint, the affirmative answer means that one might be able to explain public goods allocations which do not fit the SL definition of optimality by using our approach.

To answer the above question in the affirmative, one need only provide an example. Before presenting such an example, however, we would like to motivate the result.[25]

An SL allocation requires (only) that the allocation of resources be Pareto optimal for each possible specification of the parameters (θ_a, θ_b). An SL allocation does not take into account any trade-off between utility under one parameter specification and utility under another parameter specification, whereas all that matters for an optimum in our sense is expected utility of each agent across parameter values he cannot distinguish (i.e., those of the other agent). Consequently it may be possible to increase expected utility for example by giving agent a more utility and b less utility for some specification of (θ_a, θ_b) while doing the opposite for another specification of (θ_a, θ_b).

To make this verbal argument a bit more explicit, suppose we have an initial SL allocation which we want to improve. Moreover, suppose we restrict ourselves to a reallocation which involves *only* the private good (i.e., consider only reallocations which are themselves SL allocations). Let Δa_{ij} be the change in a's allocation of private good in the event that $(\theta_a, \theta_b) = (\theta^i, \theta^j)$, $i, j = 1, 2$. Since we are not changing the allocation of public good, the change in b's allocation of private good is given by $-\Delta a_{ij}$. The corresponding changes in expected utility are given by

$$\Delta U_i^a = \alpha_i \theta^i \Delta a_{i1} + (1-\alpha_i)\theta^i \Delta a_{i2}, \quad i = 1, 2,$$

$$\Delta U_j^b = -\beta_j \theta^j \Delta a_{1j} - (1-\beta_j)\theta^j \Delta a_{2j}, \quad j = 1, 2,$$

where ΔU_i^a is the change in a's expected utility when $\theta_a = \theta^i$ and similarly for ΔU_j^b. Ignoring the fact that the new allocation must satisfy SS, manipulation of the above expressions for ΔU_i^a and ΔU_j^b reveals that we can accomplish a Pareto improvement only if

$$\frac{\alpha_2(1-\alpha_1)}{\alpha_1(1-\alpha_2)} \neq \frac{\beta_2(1-\beta_1)}{\beta_1(1-\beta_2)}.$$

This nonequality will hold if, for example, $\alpha_2 < \alpha_1$ and $\beta_2 > \beta_1$. In this case, agent a believes that agent b tends to have preferences for the public good which are similar to his own while agent b believes that agent a tends to have preferences

[25] The intuition which follows is due largely to the suggestion of a referee.

which are opposite to his own. Thus the improvement exploits the fact that the agents hold different beliefs about the probabilities of the same events.[26]

The above discussion does not constitute a proof that there are achievable SL allocations which can be improved upon. The reasons that it is not a proof are that (i) we have not shown that there are achievable SL allocations in the first place, (ii) the construction of the Δa_{ij}'s ignores the requirement that the new allocation must be achievable, and (iii) we restricted ourselves to changing only the private good allocation, i.e., we considered only reallocations which were also SL allocations. As mentioned, however, the proof requires only an example. One such example is the following.[27] The SL allocation

$$Y = Y^S = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0.44399 & 0.22898 \\ 0.5 & 0.72487 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.05601 & 0.27102 \\ 0.0 & 0.77513 \end{bmatrix}$$

is achievable and can be strictly dominated by the achievable SL allocation

$$Y = Y^S, \quad A = \begin{bmatrix} 0.5 & 0 \\ 0 & 1.5 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0 & 0.5 \\ 0.5 & 0.0 \end{bmatrix}$$

when $\alpha_1 = 0.9$, $\alpha_2 = 0.1$, $\beta_1 = 0.1$, $\beta_2 = 0.9$, $\theta^1 = 1$, $\theta^2 = 2.1$.

The above example shows that an achievable SL allocation can be dominated by another SL allocation. We have also computed an example in which an achievable SL allocation is dominated by an achievable non-SL-allocation (this example is omitted for the sake of brevity but is available on request).

10. COMPETITIVE ALLOCATIONS AND PUBLIC INFORMATION

In this section we present one final example which illustrates the important point that *when information is public (or common) the set of feasible allocations is not restricted by any self-selection conditions.* As a consequence, in the environment of this section, a pure exchange economy, the standard result of welfare economics that competitive allocations are core allocations continues to hold when the term "core" is used in the sense defined in the previous section.

Consider a pure exchange economy with at least three agents and $l \geq 2$ commodities. The consumption set is $C = R^l_+$. Each agent t has an initial endowment $e_t \in C$. For any $A \subseteq T$, $\lambda(A) = \{c_A \in C^A | \Sigma_{a \in A} c_a = \Sigma_{a \in A} e_a\}$. There is a single parameter, $\theta_1 \in \Theta_1 = \{\theta^1, \theta^2\}$, and given θ_1, agent t has preferences defined by

[26] Note that this is precisely the condition that the agents' priors be *inconsistent* (see footnote 23, above). If we had more general utility functions, say of the form $\theta u(x) + y$ where $u'' \neq 0$, then the condition which the priors must satisfy in order to be able to improve on some SL allocation would involve marginal utilities like $u'(a_{ij})$ as well as the probabilities α_k, β_k. In this case a Pareto improvement could be possible even with consistent priors.

[27] We are grateful to Jeffrey Branman for use of his mixed integer-linear computer code and for assistance in implementing it to solve the maximization problem presented in the text. The example presented below was generated using this code.

$U_t(c, \theta_1)$ for $c \in C$ where U_t is concave and monotone increasing in c. There is assumed to be a set of informed traders, T_1 (with at least two members), who observe θ_1, i.e., $N_t = \{1\}$ for each $t \in T_1$. Let $T_2 = T - T_1$ be the set of uninformed traders, i.e., $N_t = \emptyset$ for $t \in T_2$ (we assume $T_2 \neq \emptyset$). We assume that all *informed* traders have the same endowment and preferences. We further assume that all *uninformed* traders have the same preferences, endowments, and the same prior distribution on θ_1.[28]

We restrict our attention to competitive allocations which result from the competitive equilibrium concepts discussed in Green [7] and Grossman [**8**] in which all information is revealed in equilibrium. Suppose $c_T(\theta^j)$ is a competitive allocation in the above sense when $\theta_1 = \theta^j$ for $j = 1, 2$. Therefore $c_T(\theta^j)$ is a core allocation relative to preferences defined by $U_t(\cdot, \theta^j)$, for $j = 1, 2$. The proof that the parameter contingent allocation $c_T(\cdot)$ is also a core allocation in the sense of Section 8 is by contradiction. Suppose $c_T(\cdot)$ is not a core allocation. Then there is a coalition B and a parameter-contingent allocation $c'_B(\cdot)$ such that for each $b \in B \cap T_2$,

$$(1) \qquad \sum_{j=1}^{2} \varphi_j U_b[c'_b(\theta^j), \theta^j] \geq \sum_{j=1}^{2} \varphi_j U_b[c_b(\theta^j), \theta^j]$$

where φ_j is the prior probability of all uninformed agents that $\theta_1 = \theta^j$ ($j = 1, 2$) and for each $b \in B \cap T_1$,

$$(2) \qquad U_b[c'_b(\theta^j), \theta^j] \geq U_b[c_b(\theta^j), \theta^j] \quad \text{for} \quad j = 1, 2,$$

with at least one inequality strict. Since all uninformed traders have identical concave preferences, endowments, and priors, we may assume, without loss of generality, that $c'_b(\cdot)$ is the same for all $b \in B \cap T_2$. Now suppose (2) is strict for some $b \in B \cap T_1$ and $j = 1$. Since $c_T(\theta^1)$ cannot be blocked relative to θ^1 preferences, we must have $U_b[c'_b(\theta^1), \theta^1] < U_b[c_b(\theta^1), \theta^1]$ for all $b \in B \cap T_2$. But by (1), this implies that $U_b[c'_b(\theta^2), \theta^2] > U_b[c_b(\theta^2), \theta^2]$ for all $b \in B \cap T_2$. But since (2) holds for $j = 2$, this implies that B can block $c_T(\theta^2)$ relative to θ^2 preferences. This is a contradiction. Clearly, the same argument applies if (2) is strict for some $b \in B \cap T_1$ and $j = 2$. The only case left to consider is when (2) holds as an equality for all $b \in B \cap T_1$ and $j = 1, 2$. In this case, (1) must hold strictly for all $b \in B \cap T_2$. Therefore $U_b[c'_b(\theta^j), \theta^j] > U_b[c_b(\theta^j), \theta^j]$ for at least one j for all $b \in B \cap T_2$. This again contradicts the fact that $c_T(\theta^j)$ is a core allocation relative to θ^j preferences for this j. This completes the proof.

By Theorem 3, the above core allocation $c_T(\cdot)$ can be generated using a direct mechanism. To see how this mechanism operates in this environment, consider the following simple example. There are only two informed agents, $T_1 = \{1, 2\}$. Suppose there are two goods and the competitive allocations for the informed agents are given by $c_1(\theta^1) = c_2(\theta^1) = (10, 10)$ and $c_1(\theta^2) = c_2(\theta^2) = (5, 5)$. Note that

[28] This is consistent with the notion that priors reflect the information of the agents and would be different across agents only if information varied across these agents. Since we are assuming that the information structure is fully described by $N_t = \{1\}$ for $t \in T_1$, $N_t = \emptyset$ for $t \in T_2$, there is no reason to assign different prior beliefs to different uninformed agents.

both informed traders prefer the allocation associated with $\theta_1 = \theta^1$ regardless of the true value of θ_1. In the first stage of the mechanism, each informed trader signals an element of $\{\theta^1, \theta^2\}$. The final allocation as a function of the two signals is described by Table I (only the bundles assigned to the informed agents are shown).

Signal of Agent 2 \ Signal of Agent 1	θ^1	θ^2
θ^1	(10, 10) / (10, 10)	(10, 10) / (5, 5)
θ^2	(5, 5) / (10, 10)	(5, 5) / (5, 5)

The upper half of each box shows the final bundle of agent 1 while the lower half shows the final bundle of agent 2. Thus, for example, if agent 1 signals θ^1 and agent 2 signals θ^2, then agent 2 receives (10, 10) and agent 1 receives (5, 5). We claim that a Bayesian equilibrium pair of strategies is for each informed agent to signal θ^i if the true value of θ_1 is θ^i. To see this, suppose $\theta_1 = \theta^1$. If agent 2 believes that agent 1 will signal θ^1, agent 2 is indifferent as to signalling θ^1 or θ^2 since he receives (10, 10) in either case.[29] He may just as well signal θ^1. Similarly if agent 1 believes that agent 2 will signal θ^1, it is optimal for agent 1 to signal θ^1 as well. An identical argument holds if $\theta_1 = \theta^2$. Note that the (competitive) allocation (5, 5) for both informed agents is achieved when $\theta_1 = \theta^2$ even though both would prefer the competitive allocation corresponding to $\theta_1 = \theta^1$, i.e., (10, 10) for each informed agent. Thus the informed agents would very much like to collude. Such collusion is ruled out *here* by the restriction that there be no communication between players prior to the start of any game and by the adoption of a noncooperative equilibrium concept.

11. CONCLUSION

In this section, rather than summarizing our results, we consider some directions for future research which are suggested by our analysis. These directions include both theoretical extensions and empirical work.

With regard to theoretical extensions, we would like to point out three assumptions whose relaxation may be of interest. The first is the assumption that

[29] The method of filling in the entries of the box here corresponds to the construction in the proof of Theorem 2. It is easy to fill in the entries so that each agent *strictly prefers* to signal the true value of θ, given that the other agent is also signalling the truth. On the other hand, it should be noted that there is an alternative equilibrium for the game in the table. It is easy to check that the strategy "always signal θ^1 regardless of the true value of θ" is also an equilibrium. Moreover both informed agents end up better off in this equilibrium when $\theta = \theta^2$ than in the competitive equilibrium. It turns out, however, that with very mild assumptions on the way preferences depend on θ, it is possible to fill in the entries of the table in such a way that "truth-telling" is the only equilibrium and the competitive allocation the only equilibrium allocation.

there are only a finite number of possible values of each parameter. We would hope that our results continue to hold even when some parameters can take on a continuum of values. In this context, we note, however, that some results of the signalling literature are not robust to this specification (see Riley [**22**]). The second is that the parameters enter the model only through preferences and not through the technology of production and exchange. It is our conjecture that the results will go through without this restriction.[30] Third, we note that our model is essentially static. If the environment involves allocation of resources over time, we could use the Arrow-Debreu device of regarding commodities delivered at different dates as different commodities. The analysis is static, however, in the sense that the mechanism which determines the allocation must be fully played out at the initial date. Thus, although mechanisms may be sequential in nature, we do not allow them to be played out over calendar time with consumption occurring while the mechanism is being played. Relaxing this assumption appears to be difficult and may result in somewhat different conclusions.

With regard to empirical work, we believe that when our general methodology is applied to specific environments modeled in sufficient detail, testable implications will be forthcoming. Since we have not modeled explicitly the process by which mechanisms are chosen, such implications would have to follow from the hypothesis that somehow an "efficient" (core, etc.) mechanism is chosen. Unfortunately, this procedure has two drawbacks. First, there may be constraints which limit the ability of the agents to achieve core mechanisms. If one thinks such limitations important in a specific application, then one must incorporate them into the environment or explicitly model the mechanism choice process. Second, there may be (and usually are) a large number of core mechanisms. This will, of course, reduce the sharpness of any predictions. Explicit modeling of the mechanism choice process may greatly sharpen the predictions. One trivial way of modifying our approach to include a specification of the mechanism choice process is to assume that one agent has the power to choose the mechanism while all other agents are left only with the choice of whether or not to participate. This approach has been taken by Harris and Raviv [**10, 11**] to explain various types of observed monopoly marketing schemes. Obviously this modification of the approach is appropriate only in environments in which one agent has monopoly power. More generally applicable approaches to the issue of the mechanism choice process are clearly needed.[31]

Finally, we would like to focus some attention on the role of prior beliefs in our approach. As the foregoing analysis makes clear, such beliefs play an important

[30] The problem with allowing the parameters to enter through say the production technology instead of through preferences, is that then which allocations are feasible may depend on the particular parameter draw. If so, it may not be possible to specify a mechanism which is independent of the parameters but which still always results in a feasible allocation. In some applications this problem can be finessed, if the model can be formulated so that the set of feasible allocations does not depend on the parameter draw. In the agency model of Section 2, for example, the original specification of the model was such that the relevant parameter was associated with the production technology. In that case, however, it was possible to respecify the model in such a way that the parameter entered through preferences.

[31] Much of the discussion in this paragraph is due to the suggestion of a referee.

role in the analysis: which mechanisms and which allocations are efficient in our sense often depend on prior beliefs of agents concerning parameter values they cannot observe. For the purposes of positive analysis, if priors are not observable by the economist, then the theory may be consistent with a very large class of observed mechanisms and allocations, although some positive implications are robust to the specification of prior beliefs. Thus, in order to get strong testable implications, it may be necessary for the economist to specify particular priors (or classes of priors) as part of his theory. For some applications, this may be no more difficult than specifying other aspects of the environment, however. For normative purposes, the economist may also need specific information on priors in order to suggest an efficient mechanism. Again, this may or may not be difficult relative to the ever present task of modeling the other aspects of an environment. Finally, we would like to point out that for some environments (e.g., the agency model analyzed above), there are mechanisms which are efficient for any prior beliefs.[32] Such mechanisms have great intuitive appeal, probably for this reason. Indeed, a reasonable hypothesis is that observed mechanisms will be such prior-free efficient mechanisms when they exist. An interesting topic for further theoretical and empirical research is under what conditions will such prior-free efficient mechanisms exist.

Carnegie-Mellon University

Manuscript received January, 1979; revision received February, 1980.

APPENDIX: PROOFS OF THEOREMS

Before proving the theorems stated in the text, some further notation and a definition are required. As mentioned in the text, if, for some given mechanism for a set of agents A, one knows the actual parameter values, θ_A, and the strategies to be used by each agent at each stage, σ_A^τ, then one may compute the implied sequence of signals which will be sent under this mechanism. This can be done recursively as follows. The signals sent at stage 1 are

$$s_{A,1}=[\sigma_{a,1}(s_A^0,\theta^a)]_{a\in A}=\sigma_{A,1}(s_A^0,\theta_A).$$

Thus

$$s_A^1=[s_A^0,\sigma_{A,1}(s_A^0,\theta_A)]. \tag{A1}$$

Similarly

$$s_A^2=[s_A^1,\sigma_{A,2}(s_A^1,\theta_A)]$$

when s_A^1 is given in (A1). Clearly we may continue this process until all the signals s_A^τ have been computed. We denote by $\Psi_A^\tau(\theta_A|\sigma_A^\tau)$ this sequence of signals as a function of θ_A given the strategies σ_A^τ. We also denote by $\Psi_A^t(\theta_A|\sigma_A^t)$ the subsequence consisting of signals for each agent from stage 0 through stage t (for $0\leq t\leq\tau$). Note that this sequence of signals depends only on strategies through stage t.

[32] For example, in the agency environment the following mechanism is a core mechanism regardless of the principal's prior. (A proof of this statement for a slightly different model is given in Harris and Townsend [**12**].) The mechanism has two stages. In the first, the principal proposes any *pair* of feasible allocations. In the second, the agent may choose either of these two allocations or autarky. Naturally, the equilibrium allocation of this mechanism will depend on the principal's prior.

We will also need to define precisely the concept of feasibility (relative to some mechanism) for sequences of signals and strategies. A sequence of signals s_A^τ is *feasible* for a mechanism $M_A = [\tau, S_A, F]$ if $s_{at} \in S_{at}(s_A^{t-1})$ for each $a \in A$, $t = 1, \ldots, \tau$. A sequence of strategies σ_A^τ is feasible for M_A if for any $\theta_A \in \Theta_A$, $\Psi_A^\tau(\theta_A|\sigma_A^\tau)$ is a feasible sequence of signals for M_A.

Theorem 1 is proved using the following two lemmas.

LEMMA 1: *Let A be any set of agents, M_A be a mechanism for A with equilibrium $(\sigma_A^{*\tau}, \rho_A^{*\tau})$, and let θ_P^a be any possible value of a's vector of private parameters. Then for any vector of values for those parameters which are not privately observed by a, ξ_H^a, $\Psi_A^\tau(\theta_P^a, \xi_H^a|\sigma_A^{*\tau})$ is a feasible sequence of signals for M_A.*

PROOF: This is obvious since $\Psi_A^\tau(\theta_P^a, \xi_H^a|\sigma_A^{*\tau})$ is an equilibrium sequence of signals for M_A in any environment in which the parameter values are given by (θ_P^a, ξ_H^a). *Q.E.D.*

Lemma 1 shows that it is always feasible for any agent to evaluate his equilibrium strategy at any value for his vector of private parameters even if this value is different from the true value. Lemma 2 will show that, *relative to prior beliefs* about the parameter vector ξ, it is always optimal for any agent to use his equilibrium strategy as opposed to any other strategy which, given the equilibrium strategies of the other agents, results in a feasible collection of strategies.

LEMMA 2: *Suppose A is any set of agents and $M_A = [\tau, S_A, F]$ is a mechanism for A with equilibrium $(\sigma_A^{*\tau}, \rho_A^{*\tau})$. For any agent $a \in A$, if $(\sigma_{\bar a}^{*\tau}, \sigma_a^\tau)$ is a feasible sequence of strategies, then*

$$E_a\{U_a[F_a(\Psi_A^\tau(\xi_A|\sigma_A^{*\tau})), \xi]|\theta^a\} \geq E_a\{U_a[F_a(\Psi_A^\tau(\xi_A|\sigma_{\bar a}^{*\tau}, \sigma_a^\tau)), \xi]|\theta^a\}$$

for any value of the parameter vector θ^a observed by agent a. Here the expectation E_a is taken over the entire parameter vector ξ using a's prior distribution $\rho_a(\xi|\theta^a)$.

PROOF: The proof is by backward induction on the stage. We first show that if agent a has used strategies σ_{at} for $t = 1, \ldots, \tau - 1$, it is optimal (*relative to his prior*) to use $\sigma_{a\tau}^*$ at stage τ. By definition of equilibrium for a mechanism, it is optimal (*relative to his posterior at stage τ*) for agent a to use $\sigma_{a\tau}^*$ at stage τ, *for any previous history*, $s_A^{\tau-1}$. This is true, in particular, if $s_A^{\tau-1}$ is a history generated by the strategies $(\sigma_{\bar a}^{*\tau-1}, \sigma_a^{\tau-1})$ for some parameter vector ξ. Let

$$s_A^{\tau-1}(\xi) = \Psi_A^{\tau-1}(\xi|\sigma_{\bar a}^{*\tau-1}, \sigma_a^{\tau-1}).$$

Also, to streamline the notation, let

$$Q_a^*(s_A^{\tau-1}, \xi) = U_a\{F_a[s_A^{\tau-1}, \sigma_{A\tau}^*(s_A^{\tau-1}, \xi_A)], \xi\},$$
$$Q_a(s_A^{\tau-1}, \xi) = U_a\{F_a[s_A^{\tau-1}, \sigma_{\bar a\tau}^*(s_A^{\tau-1}, \xi_{\bar a}), \sigma_{a\tau}(s_A^{\tau-1}, \xi^a)], \xi\}.$$

Thus $Q_a^*(s_A^{\tau-1}, \xi)$ is agent a's utility if the history through stage $\tau - 1$ is $s_A^{\tau-1}$, the parameter value is ξ, and each agent uses his equilibrium strategy at the last stage. Similarly, $Q_a(s_A^{\tau-1}, \xi)$ is agent a's utility in the same circumstances except that a is assumed to use the alternative strategy $\sigma_{a\tau}$ at the last stage.

Now for $s_A^{\tau-1} = s_A^{\tau-1}(\xi)$, $\Delta_{a\tau}(s_A^{\tau-1}) \neq \emptyset$ since at least $\xi \in \Delta_{a\tau}(s_A^{\tau-1})$. Thus, it follows (from the definition of equilibrium for a mechanism), that

$$\sum_{\delta \in \Delta_{a\tau}[s_A^{\tau-1}(\xi)]} Q_a^*[s_A^{\tau-1}(\xi), \delta]\rho_a(\delta|\theta^a)/K_{a\tau}[s_A^{\tau-1}(\xi), \theta^a]$$
$$\geq \sum_{\delta \in \Delta_{a\tau}[s_A^{\tau-1}(\xi)]} Q_a[s_A^{\tau-1}(\xi), \delta]\rho_a(\delta|\theta^a)/K_{a\tau}[s_A^{\tau-1}(\xi), \theta^a]$$

or, multiplying both sides by $K_{a\tau}[s_A^{\tau-1}(\xi), \theta^a]$,

$$\text{(A2)} \qquad \sum_{\delta \in \Delta_{a\tau}[s_A^{\tau-1}(\xi)]} \{Q_a^*[s_A^{\tau-1}(\xi), \delta] - Q_a[s_A^{\tau-1}(\xi), \delta]\}\rho_a(\delta|\theta^a) \geq 0.$$

But (A2) must hold for each ξ. Moreover, for any two parameter values, ξ and ξ', either $\Delta_{a\tau}[s_A^{\tau-1}(\xi)] = \Delta_{a\tau}[s_A^{\tau-1}(\xi')]$ or the two sets do not intersect at all, i.e., if the two sets intersect, then the two histories $s_A^{\tau-1}(\xi)$ and $s_A^{\tau-1}(\xi')$ must be the same. Therefore, if we sum the relation (A2) over all parameter values, ξ, but eliminate double counting by eliminating one of the values ξ or ξ' whenever

$\Delta_{a\tau}[s_A^{\tau-1}(\xi)]=\Delta_{a\tau}[s_A^{\tau-1}(\xi')]$, we obtain

$$\sum_{\xi\in\Theta} Q_a^*[s_A^{\tau-1}(\xi),\xi]\rho_a(\xi|\theta^a)\geq \sum_{\xi\in\Theta} Q_a[s_A^{\tau-1}(\xi),\xi]\rho_a(\xi|\theta^a).$$

Recalling the definitions of Q_a^* and Q_a, this shows that it is optimal (relative to his prior) for agent a to use strategy $\sigma_{a\tau}^*$ at stage τ if he has used σ_{at} in stages $1,\ldots,\tau-1$, assuming that all other agents are using their equilibrium strategies at all stages. It can then be shown (using the same argument) that it is optimal (relative to his prior) for agent a to use $\sigma_{a,\tau-1}^*$ at stage $\tau-1$ given that he used σ_{at} for $t=1,\ldots,\tau-2$ and that he plans to use $\sigma_{a\tau}^*$ at stage τ. This then shows that (relative to his prior) the sequence of strategies $\sigma_{a1},\ldots,\sigma_{a,\tau-2},\sigma_{a,\tau-1}^*,\sigma_{a\tau}^*$ dominates the sequence σ_a^τ for agent a. If we proceed with this backward induction we obtain that $\sigma_a^{*\tau}$ dominates σ_a^τ for agent a (relative to his prior) as is claimed in the lemma. *Q.E.D.*

We may now prove Theorem 1.

THEOREM 1 (Self-Selection for Mechanisms): *Suppose M_A is a mechanism for a set of agents A. Then for any agent a and any value of agent a's observed parameters θ^a,*

$$\sum_{\xi\in\Theta} U_a[c_a(M_A,\xi_A),\xi]\rho_a(\xi|\theta^a)\geq \sum_{\xi\in\Theta} U_a[c_a(M_A,\delta_P^a,\xi_H^a),\xi]\rho_a(\xi|\theta^a)$$

for any other possible value of agent a's private parameter vector, δ_P^a.

PROOF: Let θ^a be any vector of values for agent a's observed parameters and let δ_P^a by any (other) values for agent a's privately observed parameters. Let M_A be any mechanism for A (where $a\in A$) with equilibrium $(\sigma_A^{*\tau},\rho_A^{*\tau})$. Then

$$c_A(M_A,\delta_P^a,\xi_H^a)=F[\Psi_A^\tau(\delta_P^a,\xi_H^a|\sigma_A^{*\tau})]$$

for any values ξ_H^a for the parameters that agent a does not privately observe. For any $t=1,\ldots,\tau$, any values ξ^a of a's observed parameters, and any history s_A^{t-1}, define a strategy σ_{at} for a by

$$\sigma_{at}(s_A^{t-1},\xi^a)=\sigma_{at}^*(s_A^{t-1},\delta_P^a,\xi_O^a)$$

where ξ_O^a is the sub-vector of ξ^a consisting of those values which a observes but not privately. Thus σ_{at} is the equilibrium strategy of a at stage t but with the values δ_P^a substituted for a's private parameters regardless of their true values. Then for any parameter vector ξ_A observed by agents of A,

$$\Psi_A^\tau(\xi_A|\sigma_{\bar a}^{*\tau},\sigma_a^\tau)=\Psi_A^\tau(\delta_P^a,\xi_H^a|\sigma_A^{*\tau}).$$

Therefore, by Lemma 1, $(\sigma_{\bar a}^{*\tau},\sigma_a^\tau)$ is feasible. The result then follows from Lemma 2 and the definition of $c_A(M_A,\xi_A)$. *Q.E.D.*

THEOREM 2: *For any p.c. allocation $\tilde c_A$ which satisfies SS, there exists a direct mechanism D_A such that $\tilde c_A(\theta_A)=c_A(D_A,\theta_A)$ for any parameter vector θ_A. Furthermore, if all parameters are privately observed, i.e. $\theta^a=\theta_P^a$ for each agent a, then the allocation rule F for D_A can be assumed to be the same as the p.c. allocation $\tilde c_A$.*

PROOF: Let $\tilde c_A$ be any (feasible) p.c. allocation which satisfies SS. Define $\Theta^A=\times_{a\in A}\Theta^a$ with typical elements $\theta^A=(\theta^a)_{a\in A}$. Note that, in general, $\Theta^A\neq\Theta_A$ since an element of Θ^A will contain repeat values of any parameter which is observed by more than one agent. (For example, suppose $A=\{a,b\}$, $N=\{1,2,3\}$, $N_a=\{1,2\}$, $N_b=\{2,3\}$. Then $\theta^a=(\theta_1^a,\theta_2^a)$ and $\theta^b=(\theta_2^b,\theta_3^b)$ so that $\theta^A=(\theta_1^a,\theta_2^a,\theta_2^b,\theta_3^b)$ while θ_A would be of the form $(\theta_1,\theta_2,\theta_3)$.) Let Z be the subset of Θ^A consisting of all those values of θ^A such that whenever two agents, a and b, both observe a parameter, say parameter k, the values θ_k^a and θ_k^b agree, i.e.,

$$Z=\{\theta^A\in\Theta^A|\text{for any } a,b\in A \text{ and any } k\in N_a\cap N_b,\ \theta_k^a=\theta_k^b\}.$$

For any $k\in N$ and any $\theta^A\in Z$, define $z_k(\theta^A)$ to be the common value of θ_k^a for any agent a who observes parameter k, i.e. for any a such that $k\in N_a$. Thus z_k is a function from Z into Θ_k. Finally, let $z=(z_k)k\in N_A$. (In terms of the above example, Z consists of all those values of θ^A for which $\theta_2^a=\theta_2^b$. If θ^A is such a value, then $z(\theta^A)=(\theta_1^a,\theta_2^a,\theta_3^b)=(\theta_1^a,\theta_2^b,\theta_3^b)$.)

We wish to construct a mapping F which associates with each value θ^A a feasible allocation, i.e.

$F: \Theta^A \to \lambda(A)$, such that F satisfies the following two constraints:

(A3) for any $\theta^A \in Z$, $F(\theta^A) = \tilde{c}_A[z(\theta^A)]$

and

(A4) $\hat{V}_a(F, \delta^a, \gamma^a) \leq \hat{V}_a(F, \gamma^a, \gamma^a)$

for any agent a and values δ^a and γ^a of a's observed parameters, where

$$\hat{V}_a(F, \delta^a, \gamma^a) = E_a\{U_a[F_a(\delta^a, \xi^{\bar{a}}), \xi]|\gamma^a\}.$$

Here F_a is the bundle assigned by F to agent a, $\xi^{\bar{a}} = (\xi^b)b \in \bar{a}$, and the expectation is taken with respect to a's prior $\rho_a(\xi|\gamma^a)$.

The mapping F will be the allocation rule of the direct mechanism whose equilibrium allocation is $\tilde{c}_A$. Recall that in this mechanism, there is only one stage in which each agent a reports a value for each parameter he observes. Thus a's signal will be an element θ^a in Θ^a. Condition (A3) requires that if the reported parameter values of all agents, θ^A, agree for jointly observed parameters, i.e. $\theta^A \in Z$, then the rule F should give the same allocation as $\tilde{c}_A$ would if evaluated at the corresponding value $z(\theta^A)$, i.e. after throwing out duplicated parameter values sent in by several agents. Condition (A4) requires that, given F, each agent (weakly) prefers to report the true values of his observed parameters, γ^a, to reporting any other values δ^a, provided that all other agents report their true values. That is, condition (A4) requires that F be such that telling the truth are Nash strategies against F. From this discussion, it is clear that if we can find such an F, then the direct mechanism with allocation rule F will be the mechanism required in the statement of the theorem and that the identity maps on Θ^a for each agent a will constitute a set of (Nash) equilibrium strategies for this mechanism. The remainder of the proof consists of constructing such a rule F using the assumption that $\tilde{c}_A$ satisfies SS.

First consider the case where all parameters are privately observed, i.e., $N_a \cap N_b = \emptyset$ for any $a \neq b$. In this case $Z = \Theta^A = \Theta_A$ so any allocation rule, F, for the direct mechanism is just a p.c. allocation. Therefore, for any such rule, F, the above definitions reduce to

$$\hat{V}_a(F, \delta^a, \gamma^a) = \sum_{\xi \in \Theta} U_a[F_a(\delta^a_P, \xi^a_H), \xi]\rho_a(\xi|\gamma^a) \quad \text{(note } \delta^a_P = \delta^a \text{ in this case)}$$

and

$$\hat{V}_a(F, \gamma^a, \gamma^a) = V_a(F_a, \gamma^a). \quad \text{(See the definition of } V_a \text{ in Section 8.)}$$

Moreover, if we take $F = \tilde{c}_A$, then F satisfies (A3) (since $Z = \Theta_A$ and z is the identity in this case). Also $F = \tilde{c}_A$ satisfies (A4) by definition of SS for p.c. allocations (see Section 8). This proves the theorem for this case.

For the case in which some parameters are observed by more than one agent, define F as follows. For $\theta^A \in Z$, let

$$F(\theta^A) = \tilde{c}_A[z(\theta^A)]$$

so that (A3) is satisfied by definition. We must now define F on $Z' = \Theta^A - Z$ such that (A4) holds. From the definition of $\hat{V}_a$, it is clear that we need only be concerned with elements of Z of the form $(\delta^a, \xi^{\bar{a}})$ in which all signals agree for jointly observed parameters except that one particular agent (a) may disagree with the other agents. Accordingly, for any $\theta^A \in Z'$ of the form $(\delta^a, \xi^{\bar{a}})$, define

$$F(\delta^a, \xi^{\bar{a}}) = \tilde{c}_A(\delta^a_P, \xi^a_H).$$

Thus F is defined to be $\tilde{c}_A$ evaluated at values reported by a for his privately observed parameters and at values reported by the other agents (assuming they agree on jointly observed parameters) for all other parameters.

For elements of Z' not of the above form, let F be any feasible allocation, i.e. any element of $\lambda(A)$. This completes the construction of F. Given the remarks following the statement of conditions (A3) and (A4), it remains only to show that F satisfies (A4). But

$$\begin{aligned}\hat{V}_a(F, \delta^a, \gamma^a) &= \sum_{\xi \in \Theta} U_a[\tilde{c}_a(\delta^a_P, \xi^a_H), \xi]\rho_a(\xi|\gamma^a) \quad \text{by construction of } F,\\ &\leq V_a(\tilde{c}_a, \gamma^a) \quad \text{since } \tilde{c}_A \text{ satisfies SS},\\ &= \hat{V}_a(F, \gamma^a, \gamma^a) \quad \text{since } F \text{ satisfies (A3).}\end{aligned}$$

Thus F does satisfy (A4).

The mechanism $D_A = [1, \Theta^A, F]$ is the required mechanism since under D_A identity maps will be equilibrium strategies (since F satisfies (A4)). Therefore the equilibrium signals will be θ^A if the true value of the parameter vector is θ_A. Since all agents will report the same (true) value of jointly observed parameters, $\theta^A \in Z$, so that F evaluated at these equilibrium signals will simply be $\tilde{c}_A(\theta_A)$, i.e. $c_A(D_A, \theta_A) = \tilde{c}_A(\theta_A)$ as required. *Q.E.D.*

THEOREM 3: *If $\tilde{c}_A$ is a core allocation, then there is a direct core mechanism M_A such that the equilibrium allocation of M_A coincides with $\tilde{c}_A$, i.e. $\tilde{c}_A(\theta_A) \equiv c_A(M_A, \theta_A)$ for all θ_A. Conversely, if M_A is a core mechanism, then $c_A(M_A, \cdot)$ is a core allocation. The theorem also holds if "core" is replaced by "Pareto" or "individually rational Pareto."*

PROOF: We prove the theorem only for core mechanism and allocation. The proofs for Pareto and individually rational Pareto mechanism and allocation are exactly analogous.

Suppose $\tilde{c}_A$ is a core allocation. By definition $\tilde{c}_A$ satisfies SS. Using Theorem 2, let M_A be a mechanism for A such that

$$\tilde{c}_A(\theta_A) = c_A[M_A, \theta_A].$$

We claim that M_A is a core mechanism for E_A^0, for suppose not. Then there is a coalition $B \subseteq A$ and a mechanism M'_B for B such that

$$W_b[M'_B, \theta^b] \geq W_b[M_A, \theta^b]$$

for all agents $b \in B$ and $\theta^b \in \Theta^b$ with at least one inequality strict. Let $\tilde{x}_B = c_B[M'_B, \cdot]$. Then

$$V_b(\tilde{c}_b, \theta^b) = W_b[M_A, \theta^b]$$

and

$$V_b(\tilde{x}_b, \theta^b) = W_b[M'_B, \theta^b].$$

But $\tilde{x}_B$ satisfies SS (by Theorem 1), so B can block $\tilde{c}_A$ (using $\tilde{x}_B$) contrary to the assumption.

Now suppose M_A is a core mechanism for E_A^0. Let $\tilde{c}_A = c_A(M_A, \cdot)$. Thus $\tilde{c}_A$ satisfies SS by Theorem 1 and the fact that $V_a(\tilde{c}_a, \cdot) = W_a[M_A, \cdot]$. Now suppose that B can block $\tilde{c}_A$ for some coalition $B \subseteq A$. Then there is a p.c. allocation $\tilde{x}_B$ (feasible for B) such that $\tilde{x}_B$ satisfies SS and

$$V_b(\tilde{x}_b, \theta^b) \geq V_b(\tilde{c}_b, \theta^b)$$

for all b and θ^b with at least one inequality strict. Using Theorem 2, let M'_B be a mechanism for B such that $\tilde{x}_B = c_B(M'_B, \cdot)$. Then $V_b(\tilde{x}_b, \theta^b) \equiv W_b(M'_B, \theta^b)$, and so B can block M_A (using M'_B) in contradiction of the assumption that M_A is a core mechanism. *Q.E.D.*

THEOREM 4: *If $T = \{a, b\}$, $N_a = N$, $N_b \neq N$, and $\tilde{c}_T^*$ is a FI optimal p.c. allocation which satisfies SS, then $\tilde{c}_T^*$ is a Pareto p.c. allocation (in the sense defined at the beginning of Section 8).*

PROOF: First note that since $N_a = N$, for any $\theta \in \Theta$

$$V_a(\tilde{c}_a, \theta^a) = U_a[\tilde{c}_a(\theta), \theta]. \tag{A5}$$

Now suppose $\tilde{c}_T^*$ is FI optimal and satisfies SS. We need only show that no other $\tilde{c}'_T$ can strictly dominate $\tilde{c}_T^*$ in the sense of Section 8. Suppose, to the contrary, that $\tilde{c}'_T$ is a p.c. allocation which satisfies SS and

$$V_t(\tilde{c}'_t, \theta^t) \geq V_t(\tilde{c}_t^*, \theta^t)$$

for each $t \in T$ and $\theta^t \in \Theta^t$ with at least one inequality strict.

CASE 1: $V_b(\tilde{c}'_b, \theta^b) > V_b(\tilde{c}_b^*, \theta^b)$ for some $\theta^b \in \Theta^b$. Then for some $\xi \in \Theta$ such that $\xi^b = \theta^b$,

$$U_b[\tilde{c}'_b(\xi), \xi] > U_b[\tilde{c}_b^*(\xi), \xi].$$

Since $\tilde{c}_T^*$ is FI optimal, this implies

$$U_a[\tilde{c}'_a(\xi), \xi] < U_a[\tilde{c}_a^*(\xi), \xi]$$

or using (A5),

$$V_a[\tilde{c}'_a, \xi] < V_a[\tilde{c}_a^*, \xi].$$

This is the desired contradiction.

CASE 2: $V_a(\tilde{c}'_a, \theta^a) > V_a(\tilde{c}^*_a, \theta^a)$ for some $\theta \in \Theta$. Using (1), this implies that

$$U_a[\tilde{c}'_a(\theta), \theta] > U_a[\tilde{c}^*_a(\theta), \theta],$$

and since $\tilde{c}^*_T$ is FI optimal, we must have

(A6) $$U_b[\tilde{c}'_b(\theta), \theta] < U_b[\tilde{c}^*_b(\theta), \theta].$$

Furthermore, since $V_a(\tilde{c}'_a, \xi^a) \geq V_a(\tilde{c}^*_a, \xi^a)$ for every $\xi^a \in \Theta^a$, using (A5),

$$U_a[\tilde{c}'_a(\xi), \xi] \geq U_a[\tilde{c}^*_a(\xi), \xi] \quad \text{for every } \xi \in \Theta.$$

Consequently, since $\tilde{c}^*_T$ is FI optimal,

(A7) $$U_b[\tilde{c}'_b(\xi), \xi] \leq U_b[\tilde{c}^*_b(\xi), \xi] \quad \text{for every } \xi \in \Theta.$$

But (A6), (A7) and the definition of V_b imply that

$$V_b(\tilde{c}'_b, \theta^b) < V_b(\tilde{c}^*_b, \theta^b),$$

the desired contradiction. *Q.E.D.*

REFERENCES

[1] ARROW, K. J.: "The Property Rights Doctrine and Demand Revelation Under Incomplete Information," Technical Report No. 243, IMSSS, Stanford University, August, 1977.

[2] BROCK, W. A.: "The Design of Mechanisms for Efficient Allocation of Public Goods," Report No. 7655, Center for Mathematical Studies in Business and Economics, University of Chicago, December, 1976.

[3] CHAMPSAUR, P., D. J. ROBERTS, AND R. W. ROSENTHAL: "On Cores in Economies with Public Goods," *International Economic Review*, 16 (1975), 751–764.

[4] DASGUPTA, P., P. HAMMOND, AND E. MASKIN: "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility," *Review of Economic Studies*, 46 (1979), 185–216.

[5] D'ASPREMONT, C., AND L.-A. GERARD-VARET: "Incentives and Incomplete Information," *Journal of Public Economics*, 11 (1979), 25–45.

[6] GREEN, E.: "Non-Cooperative Games and Equilibrium Market Strategies," unpublished Ph.D. dissertation, GSIA, Carnegie-Mellon University, October, 1977.

[7] GREEN, J. R.: "Information, Efficiency, and Equilibrium," Discussion Paper #284, Harvard Institute of Economic Research, March, 1973.

[8] GROSSMAN, S.: "The Existence of Futures Markets, Noisy Rational Expectations and Informational Externalities," *Review of Economic Studies*, 44 (1977), 431–450.

[9] GROVES, T., AND J. LEDYARD: "Optimal Allocation of Public Goods: A Solution to the 'Free Rider Problem'," *Econometrica*, 45 (1977), 783–810.

[10] HARRIS, M., AND A. RAVIV: "Allocation Mechanisms and the Design of Auctions," GSIA Working Paper No. 5-78-79, Carnegie-Mellon University, July, 1979.

[11] ———: "A Theory of Monopoly Pricing Schemes with Uncertain Demand," GSIA Working Paper, Carnegie-Mellon University, August, 1979.

[12] HARRIS, M., AND R. M. TOWNSEND: "Allocation Mechanisms for Asymmetrically Informed Agents," GSIA Working Paper No. 35-76-77, Carnegie-Mellon University, September, 1978.

[13] HARSANYI, J. C.: "Games with Incomplete Information Played by 'Bayesian' Players (Parts I–III)," *Management Science*, 14 (1967–68), 159–182, 320–334, 481–502.

[14] HOLT, C.: "Competitive Bidding for Contracts Under Alternative Auction Procedures," *Journal of Political Economy*, 88 (1980), 433–445.

[15] HURWICZ, L.: "On Informationally Decentralized Systems," Chapter 14 of *Decision and Organization*, ed. by C. B. McGuire and R. Radner. Amsterdam: North-Holland, 1972.

[16] HURWICZ, L., AND L. SHAPIRO: "Incentive Structures Maximizing Residual Gain Under Incomplete Information," *Bell Journal of Economics*, 9 (1978), 180–191.

[17] KYDLAND, F., AND E. PRESCOTT: "Rules Rather Than Discretion: The Inconsistency of the Optimal Plan," *Journal of Political Economy*, 85 (1977), 473–492.

[18] MYERSON, R. B.: "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47 (1979), 61–74.

[19] PRESCOTT, E. C., AND R. M. TOWNSEND: "Equilibrium Under Uncertainty: Multiagent Statistical Decision Theory," in *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*, ed. by Arnold Zellner. Amsterdam: North Holland, 1980, pp. 169–194.

[20] PRESCOTT, E. C., AND M. VISSCHER: "Sequential Location Among Firms with Foresight," *Bell Journal of Economics*, 8 (1977), 378–393.

[21] REITER, S.: "Information and Performance in the $(\text{New})^2$ Welfare Economics," *American Economic Review*, 67 (1977), 226–234.

[22] RILEY, J. G.: "Informational Equilibrium," *Econometrica*, 47 (1979), 331–360.

[23] ROSENTHAL, R. W.: "External Economies and Cores," *Journal of Economic Theory*, 3 (1971), 182–188.

[24] ———: "Cooperative Games in Effectiveness Form," *Journal of Economic Theory*, 5 (1972), 88–101.

[25] ROTHSCHILD, M., AND J. STIGLITZ: "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90 (1976), 629–649.

[26] SALOP, J., AND S. SALOP: "Self-Selection and Turnover in Labor Markets," *Quartlerly Journal of Economics*, 90 (1976), 619–627.

[27] SELTEN, R.: "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4 (1975), 25–55.

[28] SHAVELL, S.: "Risk Sharing and Incentives in the Principal and Agent Relationship," *Bell Journal of Economics*, 10 (1979), 55–74.

[29] SPENCE, A. M.: *Market Signalling*. Cambridge: Harvard University Press, 1974.

[30] SPENCE, A. M., AND R. ZECKHAUSER: "Insurance, Information, and Individual Action," *American Economic Review*, 61 (1971), 380–387.

[31] WILSON, C.: "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 16 (1977), 176–207.

[32] WILSON, R. B.: "A Bidding Model of Perfect Competition," *Review of Economic Studies*, 44 (1977), 511–518.