Correlated Information and Mecanism Design

R. Preston McAfee; Philip J. Reny

Econometrica, Volume 60, Issue 2 (Mar., 1992), 395-421.

Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28199203%2960%3A2%3C395%3ACIAMD%3E2.0.CO%3B2-4

Econometrica, Vol. 60, No. 2 (March, 1992), 395–421

CORRELATED INFORMATION AND MECHANISM DESIGN

BY R. PRESTON MCAFEE AND PHILIP J. RENY[1]

In most models of asymmetric information, possession of private information leads to rents for the possessors. This tends to induce mechanism designers to distort away from efficiency. We show that this is an artifact of the presumption that information is independently distributed. Rent extraction in a large class of mechanism design games is analyzed, and a necessary and sufficient condition for arbitrarily small rents to private information is provided. In addition, the two person bargaining game is shown to have an efficient solution under first order stochastic dominance and a hazard rate condition. Similar conditions lead to full rent extraction in Milgrom–Weber auctions.

KEYWORDS: Surplus extraction, efficiency, mechanism, correlated information, auctions, private information.

1. INTRODUCTION

IN MOST MODELS OF PRIVATE OR ASYMMETRIC INFORMATION, possessors of private information receive rents or profits. For example, in the independent private values auction, the winning buyer pays less for the item for sale than it is worth to him, even when the auction is designed to maximize the price paid to the seller.[2]

Milgrom and Weber (1982) show by example that these rents result from the privacy of the information rather than its accuracy. Basically, if the information is held by two players, it has no value to either player. One can think of a Bertrand competition set up by a third player (the mechanism designer) to extract the information. More generally, when players' private information is jointly distributed in a perfectly correlated manner, it is easily rendered public and hence provides its possessors no rents.

Many applications of the mechanism design paradigm include the assumption that the information held by the players is jointly independently distributed. This has the implication that the information is purely private, in the sense that one learns nothing about one player's information from another player's information. Hence the kind of Bertrand competition which reveals the private information when it is perfectly correlated fails to do so here. As a result, the independence assumption often leads to positive rents accruing to the possessors of private information. Under the assumption that agents are risk neutral, we find that introducing arbitrarily small amounts of correlation into the joint distribution of private information among the players is enough to render private information valueless, in the sense that its possessors earn no rents. It is worth noting that while this result does depend upon the agents' attitudes

[1] We thank Andreu Mas-Colell for helpful discussions, and an anonymous referee who both suggested that we pursue the characterization result (Theorem 2), and dramatically improved the proof of Corollary 3.

[2] Myerson (1981), Riley and Samuelson (1981). See McAfee and McMillan (1987) for a survey of the voluminous auction and mechanism design literature.

toward risk (risk neutrality is heavily used), it applies to virtually all mechanism design environments of interest. Furthermore, we provide a condition on the joint distribution of agents' private information which is both necessary and sufficient for reducing the value of this information to zero.[3] This condition essentially asks that each agent's private information not be entirely uninformative about other agents' private information, and thus rules out the case of independence.

Crémer and McLean (CM) (1985) and (1988) motivated the present study. The (1985) paper provides a condition on the joint (conditional) distribution over consumers' (uncertain) characteristics sufficient for a price-discriminating monopolist to extract from them the full surplus. Their (1988) paper focuses on the private values auction environment (each agent's value of the item for sale is known to, and only to, that agent), where the agents' values are correlated. A condition is provided on the joint (conditional) distribution over agents' values that is both necessary and sufficient for allowing the auctioneer to extract all of the surplus from the bidders.

One of our main objectives is to demonstrate that these results go well beyond the auction environment. The results we obtain apply to all mechanism design environments with risk neutral bidders in which an agent's type affects his payoff in a continuous manner. In particular, one can apply our results to problems involving the allocation of public goods, optimal taxation schemes, and a variety of agency and regulatory environments.

Both of CM's papers make heavy use of not only agents' risk neutrality, but also the assumed finite state space. In CM (1985), each consumer's utility function is characterized by a parameter that can take on at most finitely many distinct values. In CM (1988), each bidder's value of the item is one of finitely many fixed possible values.

Another of our objectives is to extend CM's results to the case where the agents (bidders, consumers, bargainers, etc.) may have infinitely many possible types. This is not merely an exercise in mathematical completeness. Note that CM's result implies that we cannot explain the predominant use of the standard auction forms based on a revenue maximizing seller, since these auctions leave bidders with positive rents. However, the finite values model they employ to obtain their result is only appropriate if the bidders and auctioneers we wish to model, explicitly take into account the fixed and finite number of possible values when making decisions. If, on the other hand, the bidders and auctioneers we wish to model are always willing to admit that "one more" value distinct from the (finitely many) others currently deemed possible, is also a possibility, then the appropriate model is not one with finitely many values, but one with infinitely many. Consequently, if CM's result does not hold in the infinite values model and the standard auctions are revenue maximizing there for "many" joint distributions (allowing correlation) over bidders' values, then not only do we regain a revenue maximization based explanation of the emergence of the

[3] The condition we present is the continuum analog of that given in Crémer and McLean (1988).

standard auction forms, we also (because of its inability to explain the facts) have some basis for rejecting the finite values model outright. Thus, it is a matter of some importance to pursue this line of research through to the infinite types case.

To understand CM's (1988) result, it suffices to consider the case of two bidders. The bidders (knowing only their own value) must simultaneously choose whether or not to participate in a Vickrey auction (a sealed bid auction in which the item goes to the highest bidder at a price equal to the second highest bid). If they choose not to participate, they get zero. If they choose to participate, then they must agree to pay a participation fee. The participation fee is allowed to be random. In particular, bidder 1's participation fee may depend upon bidder 2's bid in the Vickrey auction to follow and vice versa. Since honesty is a dominant strategy in a Vickrey auction, and since each bidder's fee is independent of his own reported value in the Vickrey auction, reporting honestly remains a dominant strategy. Hence, in equilibrium (if both are willing to participate), each bidder's ultimate participation fee is a random variable, the outcome of which depends upon the other bidder's value. From now on, we shall refer to this random variable as a participation fee schedule. In fact, the mechanism is just slightly more complicated than this. Instead of presenting each bidder with a single participation fee schedule, the auctioneer presents each with a finite set (a different set for each bidder perhaps) of participation fee schedules. After learning their own value, the bidders decide whether or not they wish to participate in the upcoming Vickrey auction. If so, they must choose one of the available participation fee schedules. They are then committed to paying the fee associated with the outcome of that schedule.

If both bidders are willing to participate, then as before honesty is a dominant strategy in the Vickrey auction to follow. Since a Vickrey auction is ex post efficient (the item goes to the bidder with the highest value), the mechanism (auction plus participation fees) will be optimal from a revenue maximizing point of view if the participation fee schedules can be constructed to recover any bidder's expected profits from the Vickrey auction (leaving a bidder with no surplus, regardless of his value; bidders are therefore willing to participate). We now show how CM (1988), with an appropriate restriction on the joint distribution of bidders' values, were able to construct the required sets of participation fee schedules.

Let $v_1, \dots, v_n$ be each bidder's set of possible values, and let P be the matrix of bidder 1's conditional probabilities. Thus, the ijth entry, p_{ij}, of P denotes the probability that bidder 2 has value v_j given that bidder 1 has value v_i. Denote by $p_{i\cdot}$ the ith row of P. Finally, let π_i be bidder 1's expected profit from the Vickrey auction (excluding any participation fees) when his value is v_i.

Consider now CM's restriction on P: for all $i = 1, \dots, n$, $p_{i\cdot} \notin \operatorname{co}\{p_{k\cdot}\}_{k \neq i}$. That is, the vector of conditional probabilities corresponding to any possible value of bidder 1 is not in the convex hull of the vectors of conditional probabilities corresponding to his other possible types. With the conditional distribution satisfying this condition, the auctioneer can extract all of bidder 1's surplus as follows:[4] For each $i = 1, \dots, n$, there is a hyperplane $x_i \in \mathbb{R}^n$ separating $p_{i.}$ and $\text{co}\{p_{k.}\}$ so that $x_i \cdot p_{i.} = 0$ and $x_i \cdot p_{k.} > 0$ for all $k \neq i$. Now, for each $m = 1, \dots, n$ construct the participation fee schedule (for bidder 1) $z_m(j) = \pi_m + \alpha \cdot x_{mj}$, where $\alpha > 0$ will be specified below. Thus, if bidder 1 wishes to participate in the Vickrey auction, he must first (knowing his own value) choose a participation fee schedule $z_m(\cdot)$ say, thereby agreeing to pay $z_m(j)$ if player 2 announces a value of v_j. Since 1's payoff in the auction itself is independent of the participation fee schedule he chooses, he will choose that schedule yielding the lowest expected fee. That is, bidder 1, given that his value is i, will choose $m = 1, \dots, n$ to minimize $p_{i.} \cdot z_m = \pi_m + \alpha p_{i.} \cdot x_m$. Now, since $p_{i.} \cdot x_m > 0$ whenever $m \neq i$ and $p_{i.} \cdot x_i = 0$, we may choose $\alpha > 0$ so that for every $i, p_{i.} \cdot z_m$ is minimized when $m = i$. Hence for every $i = 1, \dots, n$, if bidder 1 has value v_i he will optimally choose fee schedule $z_i(\cdot)$ and earn an expected surplus of zero. Using a similarly constructed set of fee schedules for bidder 2, the auctioneer can in this way extract the full surplus.

Since P satisfies the condition described above for almost every distribution of values (in Lebesgue measure), full rent extraction is "usually" possible. Note that for such distributions, the precise manner in which P determines the bidder's rents in the original Vickery auction need not be considered to conclude that an optimal auction in this environment must extract all of these rents. It is this observation of Crémer and McLean's that we wish to exploit in Section 2.

Now consider the continuum analogue to CM's result. Let $f(s|t)$ be the density of s conditional on an agent's type $t \in [0,1]$, and suppose this agent anticipates profits $\pi(t)$ on average from participation in the Vickery auction. The analogous full rent extraction problem for the seller is: Construct finitely many participation fee schedules $z_1(\cdot), \dots, z_N(\cdot)$ so that for all $t \in [0,1]$

$$(1.1) \qquad \pi(t) = \min_{1 \leqslant n \leqslant N} \int_0^1 z_n(s) f(s|t)\, ds.$$

If such schedules exist, and the agent is risk neutral, then the agent's rents can be extracted.

There are several simple observations to be made. First, (1.1) is not generally solvable. If f does not depend on t (i.e. s and t are independent), and π is not a constant function, then (1.1) has no solution. Second, if the support of $f(\cdot|t)$ is monotonic in t, then (1.1) reduces to a Volterra equation and is always solvable.[5] We shall assume only that the support of f is contained in $[0,1]^2$. Third, solutions to (1.1) never exist for all continuous π, if f is continuous. That is, when $f \in C[0,1]^2$ one can always find a $\pi \in C[0,1]$ such that (1.1) has no solution.

[4] We are grateful to an anonymous referee for providing this argument.

[5] See Hochstadt (1973). The support of f is monotonic if it is of the form $[a(t), b(t)]$ and $a' \leqslant 0$ and $b' > 0$, $f(b(t)) > 0$. See Demougin (1987) for an economic application.

Thus, unlike the finite dimensional case it appears as though we can never conclude that an optimal auction in an environment yielding a continuous conditional density $f(s|t)$ must extract all of bidder 1's (receiver of signal t) rents, without investigating the precise manner in which f determines π, since the π we are faced with may be one of those for which (1.1) is not solvable. Suppose however that given f, the following were true:

$$(1.2) \qquad \forall \varepsilon > 0, \forall \pi \in C[0,1], \qquad \exists z_1, \ldots, z_N \in C[0,1] \text{ such that } \forall t \in [0,1]$$

$$0 \leqslant \pi(t) - \min_{1 \leqslant n \leqslant N} \int_0^1 z_n(s) f(s|t)\, ds < \varepsilon.$$

Then, regardless of the π determined by f as a result of the Vickery auction there is a participation charge which does not induce bidder 1 to refuse to participate given his type (the first inequality in (1.2)) and extracts all but ε of his rents where ε is arbitrarily small. Hence, if an optimal auction exists, it must extract all of bidder 1's rents regardless of his type. Thus, it is enough for f to satisfy (1.2) in order that we may extend CM's result to the continuum case.

In addition to showing that the continuum analogue of the condition on the conditional density provided by CM is both necessary and sufficient for full rent extraction, we shall provide remarkably simple sufficient conditions on f under which the mechanism designer can extract almost all of the rents (up to an arbitrary $\varepsilon > 0$) for all type realizations t.

The techniques developed in Section 2 apply not only to auction environments but to a large class of mechanism design problems, and do not impose much structure on the environment (only properties of the density f). As previously mentioned, these techniques apply to Groves mechanisms for the allocation of a public good, taxation schemes, agency and regulatory environments, and generally to environments where the presence of information correlated to private information is reasonable. However, when a specific environment is given, somewhat sharper results obtain, because properties of the relationship between the rent function π and f can be utilized. This is illustrated in Sections 3 and 4, where specific environments are described. In particular, conditions leading to complete rent extraction, rather than almost complete rent extraction, are given.

Myerson and Satterthwaite (1983) consider the following bargaining problem with two sided asymmetric information: Suppose that a seller's cost s and a buyer's value t are known only to the respective agents, are independently distributed, and that the supports of the densities overlap, so that the decision of whether to trade is nontrivial. It then turns out that any ex-post efficient trading mechanism requires subsidies from outside, that is, there is no efficient trading mechanism that "breaks even."

We shall show in Section 3 (Theorem 3) that the combination of first order stochastic dominance (increases in the buyer's value tend to increase the seller's opportunity cost of sale) and a "hazard rate" assumption on the cumulative distribution function of the seller's reported value given the buyer's type imply the existence of efficient solutions to the bargaining problem. Moreover, one such solution has the property that the buyer pays a positive participation charge and in addition pays a price less than his value for the good, provided that trade is efficient. The mechanism is constructed to be incentive compatible and for the buyer to exactly break even. Moreover, the seller obtains all of the rents (he gets to sell the good at a price equal to the buyer's value, making honest reporting a dominant strategy), and only honest reporting survives iterative elimination of dominated strategies.

Finally, we consider rent extraction in the Milgrom–Weber (1982) auction environment, and provide an alternative to the Crémer–McLean results, outside environments with finitely many private values. A condition analogous to that used in the bargaining environment leads to full rent extraction in the auction environment.

The conclusion explores the implications we wish to draw from this analysis. Although the paper develops tools for solving mechanism design problems with correlated information, the results (full rent extraction) cast doubt on the value of the current mechanism design paradigm as a model of institutional design.

2. SURPLUS EXTRACTION

We shall focus on extracting rents from a particular agent, who has type t known only to himself. We take the view that the agent is participating in a game which leaves the agent with rents equal to $\pi(t)$ on average. This game might be an auction, a bargaining game, or any other game involving private information. We assume that the agent's type falls in $[0,1]$ and that π is continuous, the latter being a feature of any mechanism design game in which type enters the payoff functions continuously.[6] Finally, we assume that the agent can achieve a payoff of zero by not participating in the game in question by defining $\pi(t)$ as the surplus in excess of the value of nonparticipation. By the revelation principle, we restrict attention (without loss of generality) to incentive compatible mechanisms. Thus, in the game under consideration, we focus attention on equilibria in which all participants report their types truthfully.

We also assume that there is a mechanism designer who may charge the agent a participation fee for the right to play the game. Thus, for example, an auctioneer may charge for the right to bid, or an arbitrator to a bargaining problem might charge both agents some amount, etc. In addition, the participation fee may be a function z of some random variable s, the realization of which the agent does not know at the time he makes the decision to participate,

[6] For most of the results of this section, s and t could be members of convex, compact subsets of Euclidean space. In particular, this holds for Theorems 1 and 2, Corollaries 2 and 3, and Lemma 1, using minor variations of the proofs.

although the agent does know t, his own type. Call such a $z(\cdot)$ a *participation fee schedule*. The random variable s, which determines his *participation fee* (i.e. a particular value $z(s)$ of $z(\cdot)$), might be another bidder's reported value in an auction, the other party's reported value in a bargaining environment, etc. What is important is that the realization of s is not influenced by the agent's reported value of his type. The agent is assumed to be risk neutral.

Let f: $[0,1]^2 \to R$ be the continuous conditional density of s, given t. Given the participation fee schedule $z(\cdot)$, $y(t) = \int_0^1 z(s)f(s|t)\,ds$ is the agent's (expected) *participation charge* given that his type is t. Supplementing the original mechanism by adding this kind of participation charge, and *assuming* (for the moment) that the agent chooses to participate for every realization of his type, again renders truth-telling as an equilibrium since the agent's participation charge is independent of his report. Let $R(f)$ denote the set of all such participation charges. Hence,

$$R(f) = \left\{ y : (\exists z \in C[0,1])(\forall t \in [0,1])\, y(t) = \int_0^1 z(s)f(s|t)\,ds \right\}$$
$$\subseteq C[0,1].$$

Note that $R(f)$ is a linear subspace of $C[0,1]$. Although we restrict attention to continuous participation fee schedules z, none of the results change if we allow for instance $z \in L^1[0,1]$.

As suggested by (1.2), we employ the supnorm $\|y\| = \max_{0 \leqslant t \leqslant 1}|y(t)|$. For any $A \subseteq C[0,1]$, we shall denote the closure of A under this norm by $\bar{A}$. Hence, $y \in \bar{A}$ if there exist arbitrarily good uniform approximations $x \in A$:

$$(\forall \varepsilon > 0)(\exists x \in A)(\forall t \in [0,1]) \quad |y(t) - x(t)| < \varepsilon.$$

As noted in the introduction (and by Crémer and McLean (1988)), the mechanism designer also has available participation charges that are independent of the agent's report and are not contained in $R(f)$. These charges are constructed as follows: Let N be a finite set of indices, and let z_n be a member of $C[0,1]$ for every $n \in N$. Present the agent with a choice of participation charges from $R(f)$. That is, the agent selects $n \in N$, and is then charged $z_n(s)$ when s is realized. The agent of type t will select n minimizing the participation charge:

$$\int_0^1 z_n(s)f(s|t)\,ds.$$

If the agent's choice of n is not used in the game to follow, the participation charge given by

$$y(t) = \min_n \int_0^1 z_n(s)f(s|t)\,ds$$

is independent of his reported value in the game to follow. We denote the set of such participation charges by $r(f) \supseteq R(f)$. Thus,

$$(2.1)\qquad r(f)=\left\{y:(\exists N)(\forall t\in[0,1])\,y(t)=\min_{1\leqslant n\leqslant N}\int_0^1 z_n(s)f(s|t)\,ds\right\}\subset C[0,1].$$

The following facts are easily established:

$$(2.2)\qquad y_1,y_2\in r(f)\Rightarrow y_1+y_2\in r(f),$$

$$(2.3)\qquad y\in r(f),\ \alpha\geqslant 0\Rightarrow \alpha y\in r(f),$$

$$(2.4)\qquad y_1,\dots,y_k\in r(f)\Rightarrow \min_{1\leqslant n\leqslant k} y_n\in r(f),$$

$$(2.5)\qquad 1,-1\in r(f),$$

$$(2.6)\qquad y_1\in r(f),\ y_2\in R(f)\Rightarrow y_1-y_2\in r(f).$$

Now, as outlined in the introduction for the special case of an auction environment, our goal is to establish conditions under which (1.2) is satisfied. Even in our more general environment, if this is the case, then regardless of the π determined by f and the equilibrium of the given mechanism being played, for any $\varepsilon>0$ there is a participation charge in $r(f)$ which induces the agent to participate in the original game (playing the original equilibrium there), and which extracts all but ε of the agent's rents. Since (1.2) is equivalent to the condition that $r(f)$ is dense in $C[0,1]$ (with respect to $\|\ \|$), the problem of full rent extraction is equivalent to finding conditions upon the conditional density $f(\cdot\,|\,\cdot)$ so that $\bar{r}(f)=C[0,1]$. The rest of this section is devoted to precisely this issue.

Our first result provides conditions sufficient for subsets of $C[0,1]$ to be dense in $C[0,1]$. It is therefore analogous to the classical Stone–Weierstrauss approximation theorem (see Friedman (1970, p. 116). Our proof follows similar lines, even though they assume that the closure of their class of functions is closed under the taking of both minima and maxima whereas we assume that it is closed only under minima. This accounts for our additional hypothesis (2.11).

Before stating the theorem we define for any $\varepsilon>0$, $\delta>0$, and $t_0\in[0,1]$, the set $U(\varepsilon,\delta,t_0)$ of (ε,δ) u-shaped functions at t_0 as follows: $u\in C[0,1]$ is in $U(\varepsilon,\delta,t_0)$ if and only if

(i) $\quad u(t)\geqslant 0\quad$ for all $\quad t\in[0,1]$,

(ii) $\quad u(t_0)\leqslant\varepsilon,\quad$ and

(iii) $\quad u(t)\geqslant 1\quad$ whenever $\quad |t-t_0|>\delta$.

Note that if $\varepsilon\leqslant\varepsilon_0$ and $\delta\leqslant\delta_0$, then $U(\varepsilon,\delta,t_0)\subseteq U(\varepsilon_0,\delta_0,t_0)$. Also note that $U(\varepsilon,\delta,t_0)$ is convex with a nonempty interior.[7]

[7] To see that $U(\varepsilon,\delta,t_0)$ has nonempty interior, note that it contains the $\varepsilon/2$ ball centered at $\varepsilon/2+|t-t_0|/\delta$.

THEOREM 1: *Suppose* $A \subseteq C[0,1]$ *satisfies*

(2.7) $x, y \in A \Rightarrow x + y \in A,$

(2.8) $x \in A,\ \alpha > 0 \Rightarrow \alpha x \in A,$

(2.9) $x_1, \ldots, x_n \in \bar{A},\ y(t) = \min\{x_1(t), \ldots, x_n(t)\} \Rightarrow y \in \bar{A},$

(2.10) $1, -1 \in A,$

(2.11) *for all* $\varepsilon, \delta > 0$ *and every* $t \in [0,1]$, $U(\epsilon, \delta, t) \cap \bar{A} \neq \varnothing.$

Then $\bar{A} = C[0,1]$.

PROOF: Fix $\pi \in C[0,1]$, and $\epsilon > 0$. Let α be a positive constant satisfying $\alpha > 2\|\pi\|$. Choose now any $t_0 \in [0,1]$, and corresponding to it a $\delta_{t_0} > 0$ such that $|\pi(t) - \pi(t')| < \varepsilon/2$ for all $t, t' \in [t_0 - \delta_{t_0}, t_0 + \delta_{t_0}]$. Next, choose $u_{t_0} \in U(\varepsilon/(2\alpha), \delta_{t_0}, t_0) \cap \bar{A}$, and let $x_{t_0}(t) = \alpha(u_{t_0}(t) - u_{t_0}(t_0)) + \pi(t_0)$. Finally, choose $\eta_{t_0} > 0$ such that $|x_{t_0}(t) - \pi(t)| < \varepsilon/2$ for all $t \in (t_0 - \eta_{t_0}, t_0 + \eta_{t_0})$. This is possible because $x_{t_0}(t_0) = \pi(t_0)$ and both x_{t_0} and π are continuous functions.

Varying $t_0 \in [0,1]$, the intervals $(t_0 - \eta_{t_0}, t_0 + \eta_{t_0})$ form an open cover of the compact set $[0,1]$. So, there is a finite subcover, and we denote the center of the k intervals comprising the subcover by $t_1, t_2, \ldots, t_k$. As in the construction above, we have corresponding to each t_i, a $\delta_{t_i} > 0$, $x_{t_i} \in \bar{A}$, and $\eta_{t_i} > 0$ with the properties endowed them by the construction.

Now, $x_{t_i}(t) < \pi(t) + \varepsilon/2$ for all $t \in (t_i - \eta_{t_i}, t_i + \eta_{t_i})$. Hence, letting $y(t) = \min_{1 \leqslant i \leqslant k}\{x_{t_i}(t)\}$ for every $t \in [0,1]$, we have that $y(t) < \pi(t) + \varepsilon/2$ for every $t \in [0,1]$. Also, for each i, and every $t \in [0,1]$,

$$\pi(t) - x_{t_i}(t) = (\pi(t) - \pi(t_i)) + \alpha(u_{t_i}(t_i) - u_{t_i}(t))$$
$$\leqslant \pi(t) - \pi(t_i) + \varepsilon/2,$$

since $u_{t_i} \in U(\varepsilon/(2\alpha), \delta_{t_i}, t_i)$. Hence, if $t \in [t_i - \delta_{t_i},\ t_i + \delta_{t_i}]$ we have $\pi(t) - x_{t_i}(t) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. On the other hand, if $t \notin [t_i - \delta_i,\ t_i + \delta_i]$, $x_{t_i}(t) \geqslant \alpha - \varepsilon/2 + \pi(t_i)$ (since $u_{t_i} \in U(\varepsilon/2\alpha, \delta_{t_i}, t_i)$).

Putting these together yields:

$$x_{t_i}(t) > \pi(t) - \varepsilon, \quad \text{if} \quad t \in [t_i - \delta_i, t_i + \delta_i]$$

and

$$x_{t_i}(t) \geqslant \alpha - \varepsilon/2 + \pi(t_i), \quad \text{if} \quad t \in [0,1] \setminus [t_i - \delta_{t_i}, t_i + \delta_{t_i}].$$

But since α was chosen so that $\alpha > 2\|\pi\|$, we have $\alpha - \varepsilon/2 + \pi(t_i) > \pi(t) - \varepsilon$ for all $t \in [0,1]$. Hence, for every $i = 1, \ldots, k$, $x_{t_i}(t) > \pi(t) - \varepsilon$ for all $t \in [0,1]$. This implies then that $y(t) > \pi(t) - \varepsilon$ for all $t \in [0,1]$, so that $\|y - \pi\| < \varepsilon$. Since $\varepsilon > 0$ was arbitrary and $y \in \bar{A}$, we conclude that $\pi \in \bar{A}$. *Q.E.D.*

REMARK 1: For the set $A = r(f)$, (2.7)–(2.10) are satisfied (see (2.2)–(2.5)). Thus, we need only ensure that (2.11) holds in order to produce $\bar{r}(f) = C[0,1]$.

Also, Theorem 1 continues to hold if s and t vary over a compact metric space. In particular it covers the finite types case as well.

We now present our main result which provides a necessary and sufficient condition (the continuum analogue of that in Crémer and McLean (1988)) for (almost) full rent extraction.

Let $\Delta[0,1]$ denote the set of probability measures on the Borel subsets of $[0,1]$. We have the following theorem.

THEOREM 2: $\bar{r}(f) = C[0,1]$ *if and only if the following condition holds*:

(*) *For every* $t_0 \in [0,1]$ *and every* $\mu \in \Delta[0,1]$,

$$\mu(\{t_0\}) \neq 1 \quad \textit{implies that} \quad f(\cdot\,|t_0) \neq \int_0^1 f(\cdot\,|t)\mu(dt).$$

PROOF: We first prove the necessity of (*) for $\bar{r}(f) = C[0,1]$. So, suppose that $\bar{r}(f) = C[0,1]$, and that $t_0 \in [0,1]$ and $\mu \in \Delta[0,1]$ satisfy $f(\cdot\,|t_0) = \int_0^1 f(\cdot\,|t)\mu(dt)$. We will show that $\mu(\{t_0\}) = 1$.

Let $y(t) = (t-t_0)^2$ for every $t \in [0,1]$. Hence, $y \in C[0,1] = \bar{r}(f)$. There must therefore be a sequence $\{y_n\}_{n=1}^{\infty}$ of functions in $r(f)$ so that $y_n \to y$. Since each $y_n \in r(f)$ we have

$$y_n(t) = \min_{1 \leqslant i \leqslant m_n} \{w_1^n(t), \dots, w_{m_n}^n(t)\}$$

for every n, and every $t \in [0,1]$, where

$$w_i^n(t) = \int_0^1 z_i^n(s) f(s|t)\,ds \quad \text{for some} \quad z_i^n \in C[0,1].$$

Thus, for each n and every $t \in [0,1]$, $y_n(t) = w_{i(n,t)}^n(t)$ for some $i(n,t) \leqslant m_n$. In particular,

$$\begin{aligned} y_n(t_0) &= w_{i(n,t_0)}^n(t_0) \\ &= \int_0^1 z_{i(n,t_0)}^n(s) f(s|t_0)\,ds \\ &= \int_0^1 \int_0^1 z_{i(n,t_0)}^n(s) f(s|t)\mu(dt)\,ds \\ &= \int_0^1 w_{i(n,t_0)}^n(t)\mu(dt). \end{aligned}$$

Since $y_n(t_0) \to y(t_0) = 0$, this implies that the last integral converges to zero. Now, by definition, $y_n(t) \leqslant w_{i(n,t_0)}^n(t)$ so that (since $\mu \in \Delta[0,1]$)

$$\int_0^1 y_n(t)\mu(dt) \leqslant \int_0^1 w_{i(n,t_0)}^n(t)\mu(dt) \to 0.$$

Hence, $0 \geqslant \int_0^1 y(t)\mu(dt) = \int_0^1 (t-t_0)^2 \mu(dt)$, so that $\mu(\{t_0\}) = 1$.

We turn now to sufficiency, and proceed by way of contradiction. Suppose that $\bar{r}(f) \neq C[0,1]$ and that (*) holds. Since hypotheses (2.7)–(2.10) of Theorem 1 are satisfied when $A = r(f)$, it must be the case (by Theorem 1) that (2.11) fails when $\bar{A}$ is replaced by $\bar{r}(f)$. Thus, there exist ε_0, $\delta_0 > 0$, and $t_0 \in [0,1]$ such that $U(\varepsilon_0, \delta_0, t_0) \cap \bar{r}(f) = \varnothing$. Since $\bar{R}(f) \subseteq \bar{r}(f)$ we have *a fortiori* that $U(\varepsilon_0, \delta_0, t_0) \cap \bar{R}(f) = \varnothing$.

Now, $\bar{R}(f)$ is convex (being a linear subspace) and as previously noted, $U(\varepsilon_0, \delta_0, t_0)$ is convex and has a nonempty interior. So, by the separating hyperplane theorem (Dunford and Schwartz (1958, 1988; Theorem 8, p. 417), there is a continuous linear functional on $C[0,1]$ separating $\bar{R}(f)$ and $U(\varepsilon_0, \delta_0, t_0)$. Equivalently, by the Riesz representation theorem (Dunford and Schwartz (1958, 1988; Theorem 3, p. 265)), there is a regular, countably additive, signed measure $\mu \neq 0$ on the Borel subsets of $[0,1]$ and a constant $c \in \mathbb{R}$ such that

$$\int_0^1 x(t)\mu(dt) \leqslant c \qquad \text{for all} \quad x \in \bar{R}(f), \quad \text{and} \tag{2.12}$$

$$\int_0^1 x(t)\mu(dt) \geqslant c \qquad \text{for all} \quad x \in U(\varepsilon_0, \delta_0, t_0). \tag{2.13}$$

Since $\bar{R}(f)$ is a linear subspace, we must therefore have $\int_0^1 x(t)\mu(dt) = 0$ for every $x \in \bar{R}(f)$. (Otherwise there is an $x_0 \in \bar{R}(f)$ with $\int_0^1 x_0(t)\mu(dt) \neq 0$, and a suitable choice of $\alpha \in \mathbb{R}$ yields $\int_0^1 \alpha x_0(t)\mu(dt) > c$, violating (2.12) since $\alpha x_0 \in \bar{R}(f)$.) Hence, c can be taken to be zero without loss of generality.

Combining (2.12) and the definition of $R(f)$ we then have

$$\int_0^1 \left\{ \int_0^1 z(s) f(s|t)\, ds \right\} \mu(dt) = 0 \quad \text{for every} \quad z \in C[0,1].$$

By Fubini's theorem, this is equivalent to

$$\int_0^1 z(s) \left[\int_0^1 f(s|t)\mu(dt) \right] ds = 0 \quad \text{for every} \quad z \in C[0,1].$$

Hence, the continuous function of s in square brackets is identically zero. That is

$$\int_0^1 f(\cdot\,|t)\mu(dt) = 0. \tag{2.14}$$

By the Jordan decomposition theorem (Cohn (1980, Corollary 4.1.5, p. 125)), we may write μ as the difference between two positive measures μ^+ and μ^- at least one of which is finite. Furthermore, there are disjoint Borel subsets of $[0,1]$, A^+ and A^-, such that $\mu^+(A^-) = \mu^-(A^+) = 0$, and $A^+ \cup A^- = [0,1]$. Thus (2.14) becomes

$$\int_{A^+} f(\cdot\,|t)\mu^+(dt) = \int_{A^-} f(\cdot\,|t)\mu^-(dt). \tag{2.15}$$

Regarding both sides of (2.15) as functions of $s \in [0,1]$, integrating over s (with respect to Lebesgue measure) and using Fubini's theorem yields $\int_{A^+} d\mu^+ = \int_{A^-} d\mu^- = 1$, where the second equality is without loss of generality. Hence, both μ^+ and μ^- are in $\Delta[0,1]$.

Combining (*), (2.15), and the fact that $\mu \neq 0$, yields that neither μ^+ nor μ^- is a point mass on t_0. In particular, since $\mu^- \in \Delta[0,1]$ is regular (see Billingsley (1968, Theorem 1.1)), there is a closed subset B of A^-, and a $\delta \in (0, \delta_0]$ such that $B \cap (t_0 - \delta, t_0 + \delta) = \varnothing$ and $\mu^-(B) > 0$. Choose $K > 1/\mu^-(B) \geqslant 1$, and define the step function x on $[0,1]$ as follows:

$$x(t) = \begin{cases} 0, & \text{if } t \in (t_0 - \delta, t_0 + \delta), \\ K, & \text{if } t \in B, \\ 1, & \text{otherwise}. \end{cases}$$

Hence, $\int_0^1 x(t)\mu(dt) \leqslant 1 - K\mu^-(B) < 0$.

Now, using Theorem 1.2 of Billingsley, it is straightforward to construct a sequence of continuous functions $\{x_n\}_{n=1}^{\infty}$ on $[0,1]$ such that for every n,

(i) $x_n(t) \geqslant 1$, for every $t \notin (t_0 - \delta, t_0 + \delta)$,

(ii) $x_n(t) \geqslant 0$, for every $t \in [0,1]$,

(iii) $x_n(t_0) = 0$,

(iv) for every $t \in [0,1]$, $x_n(t) \to x(t)$,

(v) for every $t \in [0,1]$, $x_n(t) \leqslant K$.

By (i)–(iii) $x_n \in U(\varepsilon_0, \delta, t_0) \subseteq U(\varepsilon_0, \delta_0, t_0)$ (since $\delta \leqslant \delta_0$), for every n. And by (ii), (iv), (v), and Lebesgue's dominated convergence theorem, $\int_0^1 x_n(t)\mu(dt) \to \int_0^1 x(t)\mu(dt) < 0$. Thus for n large enough, $\int_0^1 x_n(t)\mu(dt) < 0$, contradicting (2.13).

Q.E.D.

To better understand condition (*), consider μ to be a prior on the agent's type t. The induced distribution on the signal s is then given by $\int_0^1 f(\cdot\,|t)\mu(dt)$. Hence, if $f(\cdot\,|t_0) = \int_0^1 f(\cdot\,|t)\mu(dt)$, then learning that the agent's type is t_0 provides no new information about the signal s. Condition (*) asks that unless one's prior is already concentrated on an agent's type t_0 say, learning the agent's type is always informative about the signal s. In particular, when s is the reported value of another agent (which in equilibrium is a truthful report), (*) asks that each agent's private information not be entirely uninformative about other agents' private information.

REMARK 2: Note that if for every $t_0 \in [0,1]$, there is an $x_{t_0} \in \bar{r}(f)$ taking a minimum uniquely at t_0, then setting $y(t) = x_{t_0}(t) - x_{t_0}(t_0)$ in the proof of necessity above is enough to show that (*) holds and hence (by sufficiency) that $\bar{r}(f) = C[0,1]$. This observation is at the heart of the three corollaries which follow.[8] Like Theorem 1, Theorem 2 also holds if $[0,1]$ is replaced by any

[8] Equivalently, if for every $t_0 \in [0,1]$, such an $x_{t_0} \in \bar{r}(f)$ exists, then $\bar{r}(f)$ satisfies (2.11) and hence all the hypotheses of Theorem 1, by Remark 1. Again this yields $\bar{r}(f) = C[0,1]$.

compact metric space. In particular, the theorem and proof given here cover the finite types case.

COROLLARY 1: *Suppose* $x \in R(f)$, $y \in r(f)$ *satisfy*

$$(\forall t)\, x'(t) > 0,$$

$$(\forall t)\, y'(t)/x'(t) \text{ is strictly increasing in } t.$$

Then $\bar{r}(f) = C[0,1]$.

PROOF: As noted in Remark 2, we need only find for each $t_0 \in [0,1]$ a function in $\bar{r}(f)$ taking a minimum uniquely at t_0. Let

$$q(t) = y(t) - \frac{y'(t_0)}{x'(t_0)} x(t).$$

By (2.6), $q \in r(f)$. Moreover

$$q'(t) = y'(t) - \frac{y'(t_0)}{x'(t_0)} x'(t) \gtreqless 0 \quad \text{as} \quad \frac{y'(t)}{x'(t)} \gtreqless \frac{y'(t_0)}{x'(t_0)} \quad \text{as} \quad t \gtreqless t_0.$$

Thus $q(t)$ achieves a minimum uniquely at $t = t_0$. *Q.E.D.*

REMARK 3: Consider a monotonic transformation of F, the conditional c.d.f. of s given t:

$$G(s|t) = F(\varphi(s)|\psi(t))$$

where $\varphi', \psi' > 0$, $\varphi(0) = \psi(0) = 0$, $\varphi(1) = \psi(1) = 1$. Then $R(g) = \{x\colon\ x(t) = y(\psi^{-1}(t)),\ y \in R(f)\}$. To see this, note

$$y(t) = \int_0^1 z(s) g(s|t)\, ds = \int_0^1 z(u) f(\varphi(u)|\psi(t)) \varphi'(u)\, du$$

$$= \int_0^1 z(\varphi^{-1}(s)) f(s|\psi(t))\, ds.$$

Thus, if $y \in R(f)$, $y(\psi^{-1}(t)) \in R(g)$ and vice versa. Now suppose, $x, y \in R(f)$ satisfy the hypotheses of Corollary 1. Then $x(\psi^{-1}(t))$, $y(\psi^{-1}(t)) \in R(g)$, and

$$\frac{\partial}{\partial t} \frac{\frac{\partial}{\partial t} y(\psi^{-1}(t))}{\frac{\partial}{\partial t} x(\psi^{-1}(t))} = \frac{\partial}{\partial t} \frac{y'(\psi^{-1}(t))}{x'(\psi^{-1}(t))} = \left[\frac{\partial}{\partial u} \frac{y'(u)}{x'(u)}\right] \psi^{-1'}(u) \Big|_{u = \psi^{-1}(t)} > 0.$$

Thus, if Corollary 1 applies to f, it applies to a rescaling of f.

By Corollary 1, it is straightforward to show that combined with first order stochastic dominance, a sufficient condition for $\bar{r}(f) = C[0,1]$ is that $E[s^2|E(s|t) = \mu]$ be a strictly convex function of μ. The following example illustrates this.

EXAMPLE 1: $f(s|t) = ts^{t-1}$. $\mu = E(s|t) = t/t+1$ implies that $t = \mu/(1-\mu)$. Also, $E(s^2|t) = t/t+2$, so that $E\{s^2|Es = \mu\} = \mu/(2-\mu)$, a convex function of $\mu \in [0, 1/2]$. Thus, letting $x(t) = E(s|t) = t/t+1$ and $y(t) = E(s^2|t) = t/t+2$, we have that $y'(t)/x'(t) = 2(t+1/t+2)^2$ is increasing in t. Since by first order stochastic dominance $x'(t) > 0$, Corollary 1 can be directly applied to conclude that $\bar{r}(f) = C[0,1]$. In general, with first order stochastic dominance and $E[s^2|E(s|t) = \mu]$ a convex function of μ, $x(t) = E(s|t)$ and $y(t) = E(s^2|t)$ will satisfy the hypotheses of Corollary 1. Furthermore, in this case the participation fee schedules $z_n(s)$, can be chosen to be quadratic in s. These conditions are satisfied for many common distributions, in particular those with mean and variance increasing in t.

The lemma to follow establishes a useful equivalence for placing functions in $\bar{R}(f)$. $[A]$ denotes the linear span of A. The proof of Lemma 1 is in the Appendix.

LEMMA 1: $\bar{R}(f) = \overline{[\{f(s|\cdot): 0 \leqslant s \leqslant 1\}]}$.

COROLLARY 2: *Suppose that*

$$(2.16) \qquad (\forall t)(\exists s)(\forall t' \neq t) \quad f(s|t) > f(s|t').$$

Then $\bar{r}(f) = C[0,1]$.

PROOF: By Lemma 1, $-f(s|\cdot) \in \bar{r}(f)$. By (2.16), for each t, there exists an s with $-f(s|\cdot)$ taking a minimum uniquely at t. In light of Remark 2, $\bar{r}(f) = C[0,1]$. *Q.E.D.*

Figure 1 presents a density satisfying the hypotheses of Corollary 2. The condition on f expressed in Corollary 2 has a number of interpretations. The first is the direct one, namely that for each of a player's types t, there exists a value of the signal s, so that t maximizes the likelihood of s. Alternatively, one can relate the hypotheses of Corollary 2 to a strengthening of first order stochastic dominance. For instance, suppose that in addition to f satisfying first order stochastic dominance, for every $t_0, t_1 \in [0,1]$ there exists a unique $s \in [0,1]$ such that $f(s|t_0) = f(s|t_1)$ and that fixing t_0, this s is strictly monotonic in t_1. (Figure 2 below illustrates this.) It is not hard to show that in such a case the hypotheses of Corollary 2 must be satisfied.

The final result of this section provides further conditions for rent extraction which, in some instances, are simple to verify.

COROLLARY 3: *Suppose there exists a set* $S \subseteq [0,1]$ *such that*

$$(2.17) \qquad (\forall s \in S) \quad f(s|t) \text{ is strictly concave in } t, \quad \text{and}$$

$$(2.18) \qquad (\forall t_0, t_1 \in [0,1])(\exists s \in S) \quad f(s|t_0) \geqslant f(s|t_1).$$

Then $\bar{r}(f) = C[0,1]$.

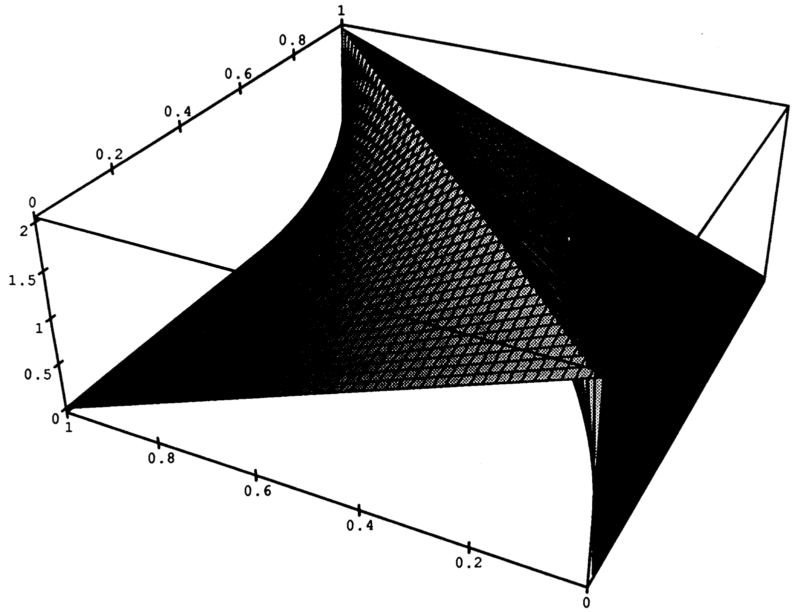

FIGURE 1.—$f(s|t) = 2\min\{s/t, (1-s)/(1-t)\}$.

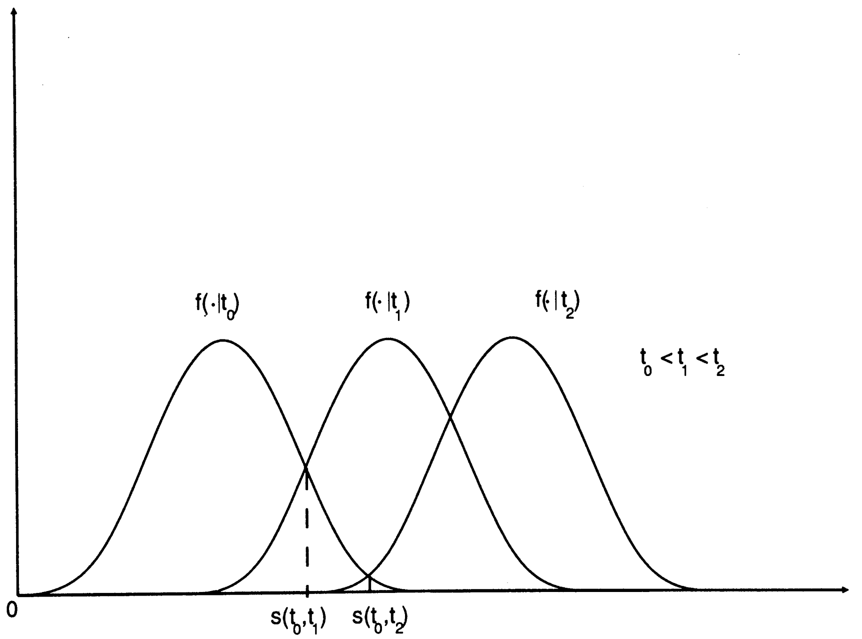

FIGURE 2.

PROOF: Suppose (2.17) and (2.18) hold, but (*) fails. Then there exists a $t_0 \in [0,1]$ and $\mu \in \Delta[0,1]$ not a point mass on t_0, with $f(\cdot|t_0) = \int f(\cdot|t)\mu(dt)$. Define $t_1 = \int t\mu(dt)$. Since $f(s|t)$ is strictly concave in t, for all $s \in S$, we have, by Jensen's inequality,

$$(\forall s \in S) \quad f(s|t_0) = \int f(s|t)\mu(dt) < f(s|t_1),$$

which contradicts (2.18). *Q.E.D.*

We gratefully acknowledge that a referee provided us with this dramatically improved proof.

REMARK 4: If the set S in the hypotheses of Corollary 3 is compact and convex, it follows that for all $t_0 \in [0,1]$, there exists an $s \in S$ so that $f(s|t)$ is maximized at $t = t_0$, in which case the hypothesis of Corollary 2 is satisfied. However, there are examples where S is not convex and the hypothesis of Corollary 2 fails to hold, even though the hypotheses of Corollary 3 hold.

REMARK 5: The conclusion of Corollary 3 continues to obtain if, in the hypotheses, "concave" is replaced with "convex."

REMARK 6: One case of interest occurs when $f(s_0|t)$ is strictly concave and decreasing in t, while $f(s_1|t)$ is strictly concave and increasing in t. In this case, $S = \{s_0, s_1\}$ suffices, as (2.18) follows immediately. Thus, we can establish the desired property without any information about "most" of f.

REMARK 7: Corollaries 2 and 3 continue to obtain when s and t are members of convex, compact subsets of Euclidean space, which includes the case of many players. Thus, there is nothing special about the one dimensional case, at least for these results.

We now show by example that the combination of first order stochastic dominance and affiliation is not sufficient to guarantee $\bar{r}(f) = C[0,1]$. As the example illustrates, the combination of these properties admits an f with $R(f)$ comprised of only linear functions and no u-shaped functions.

EXAMPLE 2: $f(s|t) = 1 + (2s-1)t$. Note $\int_0^1 f(s|t)\,ds = 1 + (s^2 - s)t|_0^1 = 1$, and $f(s|t) \geqslant 1 - t \geqslant 0$, so f is an admissible conditional density.

$$F(s|t) = \int_0^s f(u|t)\,du = s + t(s^2 - s),$$

$$F_t(s|t) = s^2 - s < 0 \quad \text{for} \quad s \in (0,1),$$

so F satisfies strict first order stochastic dominance. Also,

$$\frac{\partial^2}{\partial s\,\partial t}\log f(s|t)=\frac{\partial}{\partial s}\frac{2s-1}{1+(2s-1)t}>0,$$

so f is affiliated (see Milgrom and Weber (1982)). Equivalently, f has the monotone likelihood ratio property.

Finally, $R(f)=[\{1,t\}]$, the set of linear functions. (1 indicates the constant function, t the identity.) It is easily seen that $\bar{r}(f)$ is then the set of concave functions, and is thus a strict subset of $C[0,1]$.

We mention briefly that the results of this section can be extended in a straightforward manner to the unbounded support case. This may require allowing players to choose from among countably (rather than finitely) many participation charges, so in what immediately follows $r(f)$ is:

$$r(f)=\left\{y(t)\in C(\mathbb{R})|y(t)=\min_{n\geqslant 1}\int z_n(s)f(s|t)\,ds\right.$$
$$\left.\text{for some countable subset}\quad \{z_n\}_{n=1}^{\infty}\text{ of }C(\mathbb{R})\right\},$$

where $C(\mathbb{R})$ denotes the set of bounded continuous functions on $\mathbb{R}$. Corollary 1 above goes through verbatim.

COROLLARY 1′: *If* $x\in R(f)$, $y\in r(f)$, $x'>0$, *and* y'/x' *is strictly increasing, then* $\bar{r}(f)=C(\mathbb{R})$.

Corollary 2 above also admits a natural counterpart namely, Corollary 2′:

COROLLARY 2′: *Suppose*

(i) $\forall t\in\mathbb{R},\exists s\in\mathbb{R}$ *such that*:

(*a*) $f(s|t)>f(s|t')\quad \forall t'\neq t$,

(*b*) $f(s|t)>\limsup_{|t'|\to\infty}f(s|t')$,

and

(ii) $f(\cdot\,|\,\cdot)$ *is uniformly continuous on* $\mathbb{R}^2$.

Then $\bar{r}(f)=C(\mathbb{R})$.

Both Corollaries 1′ and 2′ are of particular interest since their hypotheses are satisfied when s and t are jointly normally distributed, with nonzero covariance. The proofs of Corollaries 1′ and 2′ follow from suitable modifications of Theorem 1 and Lemma 1. The hypotheses of Corollary 1′ are also satisfied when the agent's type and the signal are additively related (i.e. $f(s|t)=h(s-t)$), and to cases in which $s=x_0+x_1$ and $t=x_0+x_2$ where the x_i's are independent draws from gamma distributions with parameters (α_i, β_i), $i = 0, 1, 2$. In each of these cases, participation fee schedules $z_n(s)$ that are quadratic in s suffice.

We now return to the case of bounded support and end this section with a brief application. (A more detailed application is provided in the next section.) Consider a principal designing a contract for a risk neutral agent possessing private information $t \in [0, 1]$. The principal knows he can receive signal s correlated to t, sometime in the future. What is the value of s? Consider the full information gains from trade G, and the solution to the informationally constrained contract design problem, which gives the principal profits of G'. We have shown that if an efficient mechanism exists, then, for many densities, the value of the correlated information is $G - G'$. This follows since the principal can set up a mechanism which is full-information efficient, producing rents G, and then extract those rents via a participation charge $z_n(s)$. That is, the principal "sells the agency" to the agent for $z_n(s)$. We believe that, in many economic problems, the presence of correlated information is natural, and destroys the "inefficiencies resulting from private information" so often cited in the literature (see McAfee and McMillan (1987a) for references). The third and fourth sections provide two such examples.

3. BARGAINING MECHANISMS

Consider a buyer with value t, known only to himself, of an item and a potential seller, who privately observes his own opportunity cost of sale, s. It is common knowledge that s and t were drawn from a joint density $g(s, t)$ with support $[0, 1]^2$. Both buyer and seller are risk neutral.

As Myerson and Satterthwaite (1983) showed, if s and t are independent, then there is no efficient mechanism for arranging trades in this environment that does not lose money on average. As we shall show, however, under alternative conditions, there is an efficient mechanism. We think it is plausible that the determinants of the buyer's value may also influence the seller's opportunity cost. Thus when the buyer's value is high (e.g. due to an increased estimate of resale value) the seller's opportunity cost will typically be higher than usual. Thus, independence of values is likely the exception rather than the rule. We now show how the results obtained in the last section (Corollaries 1, 2, and 3) can be used to demonstrate the existence of an efficient mechanism when the values are correlated.

Consider first the following "pre"-mechanism which includes a risk neutral third party who acts solely as a budget balancer when necessary. Let r_σ, r_β denote the seller's, buyer's reported signal respectively. If $r_\sigma \leqslant r_\beta$, then the seller receives r_β for the good and the buyer pays r_σ. The difference $r_\beta - r_\sigma$ is made up by the budget balancer. If $r_\sigma > r_\beta$, then the good is unsold and no payments are made. Honesty is a dominant strategy for both buyer and seller here and in equilibrium the budget balancer is expected to lose

$$G = \int_0^1 \int_0^t (t - s) g(s, t) \, ds \, dt,$$

the expected gain from trade, which we assume is positive. This pre-mechanism is clearly ex-post efficient although it requires "outside" money. Myerson and Satterthwaite showed that any ex-post efficient mechanism requires outside money (in the sense that a budget balancer must be included and would expect to lose money) when the valuations of the buyer and seller are independently distributed and their supports intersect in an interval.

However, let $h(s|t)$ be the conditional density of the seller's value given that the buyer's value is t and let $k(t|s)$ be the conditional density of the buyer's value given the seller's value is s, and suppose that both h and k satisfy (*) (ruling out independence, in particular). Then, letting π^σ, π^β denote the seller's, buyer's rent function (a function of their respective value of the good) respectively obtained from participation in the game defined by the pre-mechanism, we have, for $(s,t) \in [0,1]^2$,

$$\pi^\sigma(s) \equiv \int_0^1 (t-s)k(t|s)\,dt, \tag{3.1}$$

$$\pi^\beta(t) \equiv \int_0^1 (t-s)h(s|t)\,ds. \tag{3.2}$$

Now, by assumption $\bar{r}(k) = \bar{r}(h) = C[0,1]$. Hence, given any $\varepsilon > 0$ there exist finite sets of participation fee schedules $\{z_n^\beta\}_{n \in N_\beta}$, $\{z_n^\sigma\}_{n \in N_\sigma}$, one for the buyer and one for the seller, such that for all $(s,t) \in [0,1]^2$,

$$0 \leqslant \pi^\sigma(s) - \min_{n \in N_\sigma} \int_0^1 z_n^\sigma(t)k(t|s)\,dt < \varepsilon \tag{3.3}$$

and

$$0 \leqslant \pi^\beta(t) - \min_{n \in N_\beta} \int_0^1 z_n^\beta(s)h(s|t)\,ds < \varepsilon. \tag{3.4}$$

Hence, if we supplement the pre-mechanism above by offering the seller a choice among participation fee schedules from $\{z_n^\sigma\}_{n \in N_\sigma}$ and his value is s, he will expect to be charged

$$c^\sigma(s) = \min_{n \in N_\sigma} \int_0^1 z_n^\sigma(t)k(t|s)\,dt,$$

since by choosing $n \in N_\sigma$ he will be charged $z_n^\sigma(t)$ if the buyer reports value t. His rents therefore become $\pi^\sigma(s) - c^\sigma(s)$ and lie between 0 and ε by (3.3). The buyer must similarly choose among $\{z_n^\beta\}_{n \in N_\beta}$ and his rents become $\pi^\beta(t) - c^\beta(t) \in [0, \varepsilon]$ when his value is t and where $c^\beta(t)$ is defined analogously to $c^\sigma(s)$. All revenue generated by these participation fees is given to the budget balancer.

Hence, the budget balancer's net expected revenue becomes:

$$\begin{aligned} R &\equiv \int_0^1 \int_0^1 c^{\sigma}(s) g(s,t)\, ds\, dt + \int_0^1 \int_0^1 c^{\beta}(t) g(s,t)\, dt\, ds - G \\ &\geqslant (G-\varepsilon) + (G-\varepsilon) - G \\ &= G - 2\varepsilon \end{aligned}$$

where the inequality arises from the definitions of G, $c^{\sigma}(\cdot)$ and $c^{\beta}(\cdot)$, and (3.1)–(3.4). So, for ε small enough, $R > 0$. Finally, choose $\alpha \in [0,1]$ and require the budget balancer to give the seller αR and the buyer $(1-\alpha)R$ (independently of reports), so that the budget balancer now expects to break even. Since none of the charges we have introduced affect the buyer's or seller's incentive to reveal his true value, this mechanism (i.e. pre-mechanism plus net participation charge) is ex-post efficient and allows the budget balancer to break even on average. Note that since $R \to G$ as $\varepsilon \to 0$, by choosing $\alpha \in [0,1]$ appropriately we may give the seller δG of the gains from trade for any $\delta \in (0,1)$. On the other hand, giving all of the gains from trade to either the seller or the buyer may not be possible with participation charges of this sort.

As mentioned in the introduction, the analysis of the rent extraction problem is greatly simplified when one need not take into account the precise relationship between the conditional density f and a player's rent function π. On the other hand in some environments this more detailed analysis is tractable and provides stronger results; namely, all rents rather than almost all rents can be extracted and more importantly the mechanism which extracts the rents can actually be constructed rather than simply shown to exist. As we now show, the Myerson and Satterthwaite bargaining environment is amenable to this more detailed approach.

As before, $g(s,t)$ is the joint density between the buyer's and seller's valuation of the good. Let

$$f(s \mid t) = g(s,t) \Big/ \int_0^1 g(u,t)\, du$$

so that f is the conditional density, and let

$$F(s \mid t) = \int_0^s f(u \mid t)\, du$$

be the distribution function of s, conditional on t. Let $F_2(s \mid t) = \partial/\partial t\, F(s \mid t)$. By taking advantage of our explicit description of this mechanism design environment, we get the following result:

THEOREM 3: *Suppose* $\forall (s,t) \in (0,1)^2$,

(3.5) $\quad F_2(s \mid t) < 0,$

(3.6) $\quad \dfrac{\partial}{\partial_t}\left[t + \dfrac{F(s \mid t)}{F_2(s \mid t)} \right] \geqslant 0.$

Then there exists an efficient trading mechanism giving all of the rents to the seller.

PROOF: We display the mechanism. The buyer and seller make reports (r and $\hat{s}$, respectively) of their values t and s, respectively. The seller receives r for the item if $r > \hat{s}$, and otherwise no trade occurs and the seller gets nothing. Honesty is a dominant strategy for the seller, and should the buyer choose to be honest (as he will in equilibrium), all the rents go to the seller. The buyer is awarded the good if $r \geqslant \hat{s}$, and is required to pay

$$\begin{cases} -\dfrac{F(r|r)^2}{F_2(r|r)} & \text{if} \quad r < \hat{s}, \\ r + \dfrac{F(r|r)}{F_2(r|r)} - \dfrac{F(r|r)^2}{F_2(r|r)} & \text{if} \quad r \geqslant \hat{s}. \end{cases}$$

Since the seller is honest, this provides a buyer with value t who reports r rents equal to

$$u(r,t) = \frac{F(r|r)^2}{F_2(r|r)} + \left[t - r - \frac{F(r|r)}{F_2(r|r)}\right] F(r|t)$$

as $\hat{s} < r$ with probability $F(r|t)$.

Since $u(t,t) = 0$, the buyer is willing to participate, as he can obtain at least 0. We need only verify that he can do no better than 0 to complete the proof:

$$\begin{aligned} r \gtreqless t \quad & \text{as} \quad t + \frac{F(r|t)}{F_2(r|t)} \lesseqgtr r + \frac{F(r|r)}{F_2(r|r)} \quad \text{(by (3.6))}, \\ & \text{as} \quad \left[t - r - \frac{F(r|r)}{F_2(r|r)}\right] + \frac{F(r|t)}{F_2(r|t)} \lesseqgtr 0, \\ & \text{as} \quad \left[t - r - \frac{F(r|r)}{F_2(r|r)}\right] F_2(r|t) + F(r|t) \gtreqless 0 \quad \text{(by (3.5))}, \\ & \text{as} \quad u_2(r,t) \gtreqless 0. \end{aligned}$$

Thus $u_2(r,t) \gtreqless 0$ as $r \gtreqless t$. We conclude that $u(r,t) \geqslant 0$, since $u(r,t)$ increases in t for $t < r$ up to $u(r,r) = 0$, and then decreases in $t > r$. Consequently, incentive compatibility is satisfied. *Q.E.D.*

REMARK 8: The mechanism requires a budget balancer, since payments by the buyer equal payments to the seller on average, but not for all realizations.

We wish to argue that the hypotheses of Theorem 3 are plausible. The first hypothesis, first order stochastic dominance, merely requires that the aspects of the environment that increase the buyer's value also tend to increase the seller's opportunity cost. Thus there is a "common value" aspect to sale: increases in the buyer's value also increase the seller's use value of the item. The second condition (3.2), looks like a hazard rate condition (see McAfee and McMillan

(1987a) for an intuitive explanation), and is equivalent to

$$\frac{\partial^2}{(\partial t)^2}\frac{1}{F(s|t)} \geqslant 0.$$

This condition may be related to the "usual" hazard rate condition,

$$\partial/\partial s(s + F(s)/f(s)) \geqslant 0,$$

by observing that the latter is equivalent to

$$\frac{\partial^2}{(\partial s)^2}\frac{1}{F(s)} \geqslant 0.$$

Thus, (3.6) requires the standard hazard property to hold for t instead of s. This condition is satisfied for many examples. Example 1 in the preceding section satisfies it, so that, although $r(f)$ contains only concave functions, and the profits net of participation charges to the buyer in efficient mechanisms are strictly convex, the rents may still be extracted. This paradox is resolved by noting that the mechanism described here is not of the Section 2 form (which consist of participation charges alone) but links the participation charge to the sale of the item. Therefore, by exploiting aspects of the game, a mechanism designer can attain allocations which cannot be attained using unconnected participation charges and gambles.

To summarize, under reasonable assumptions on the distribution of values, which imply a certain amount of correlation, there exists an efficient solution to the bilateral bargaining problem with asymmetric information.

Next we show that Milgrom and Weber (1982) auction environments are also amenable to the more detailed analysis just applied to the bargaining problem. This supplements the implications of our analysis in Section 2 that under the hypotheses of either Theorems 2, 4, or 5 an optimal auction extracts all bidders' rents by providing conditions under which an optimal auction exists and a construction of such an auction.

4. OPTIMAL AUCTIONS

Milgrom and Weber (1982) present a general model of the auction environment, allowing for correlation among valuations, and valuations which are viewed by the bidders as random. For simplicity, we shall consider the case of two bidders. It is assumed that the seller values the item at zero. The bidders receive signals s and t respectively, privately observed, which are generated by a density $g(s,t)$ with support $[0,1]^2$. We assume g is symmetric and focus on bidder "1" who has signal t. Let

$$f(s|t) = \frac{g(s,t)}{\int_0^1 g(u,t)\,du}, \quad \text{and}$$

$$F(s|t) = \int_0^s f(u|t)\,du.$$

Even if bidder 1 knows s and t, his valuation may be random, so we let $u(t,s)$ be the expected value of the object to bidder 1 given s and t. Symmetrically, we assume that the other bidder values the item at $u(s,t)$. Finally, we assume that u is strictly increasing in its first argument, and that

$$t > s \Rightarrow u(t,s) > u(s,t)$$

so that the agent with the highest signal is the efficient consumer of the item. Define

$$v(r,t) = \int_0^r u(t,s) f(s|t)\, ds = E\{u(t,s)|s \leqslant r\} F(r|t),$$

which is the expected valuation to bidder 1 if his signal is t and he obtains the item whenever $s \leqslant r$. As before, subscripts denote partial derivatives.

THEOREM 4: *Suppose* $\forall (s,t) \in (0,1)^2$

$$F_2(s|t) < 0, \qquad \textit{and}$$

$$\frac{\partial}{\partial t} \frac{v_2(s,t)}{F_2(s|t)} \geqslant 0.$$

Then there exists an efficient mechanism which extracts exactly all of the bidders' surplus, and hence is optimal from the seller's point of view.

PROOF: Suppose bidder 2 reports his signal honestly, and bidder 1 has signal t and reports r. The mechanism awards bidder 1 the item if $r \geqslant s$, with a charge of

$$v(r,r) - \frac{v_2(r,r)F(r|r)}{F_2(r|r)}, \quad s > r,$$

$$v(r,r) - \frac{v_2(r,r)F(r|r)}{F_2(r|r)} + \frac{v_2(r,r)}{F_2(r|r)}, \quad s \leqslant r.$$

Bidder 1's profit is

$$\pi(r,t) = v(r,t) - v(r,r) + \frac{v_2(r,r)(F(r|r) - F(r|t))}{F_2(r|r)}.$$

Clearly $\pi(t,t) = 0$, so individual rationality is satisfied.

$$\pi_2(r,t) = v_2(r,t) - \frac{v_2(r,r)F_2(r|t)}{F_2(r|r)} \gtreqless 0 \quad \text{as}$$

$$\frac{v_2(r,t)}{F_2(r|t)} \lesseqgtr \frac{v_2(r,r)}{F_2(r|r)} \quad \text{as} \quad t \lesseqgtr r.$$

Thus, $\pi(r,t)$ is maximized, as a function of t, at $t=r$, that is,

$$\pi(r,t) \leqslant \pi(r,r) = 0.$$

Consequently, incentive compatibility is satisfied. The mechanism is then efficient, incentive compatible, and extracts all the rents. *Q.E.D.*

REMARK 9: If the auction is a private values one, $u(s,t)=t$ and $v(r,t)=tF(r|t)$. Thus

$$\frac{\partial}{\partial t}\frac{v_2(r,t)}{F_2(r|t)} = \frac{\partial}{\partial t}\left(t + \frac{F(r|t)}{F_2(r|t)}\right),$$

which is the hazard condition (31) from the previous section. Generally, this condition says that the expected value of receiving the item whenever $s \leqslant r$ is a convex function of $F(r|t)$, which is not (at least to us) an intuitive requirement.

For the common value environment, where $u(t,s)=u(s,t)$, and where there is no issue of efficiency (awarding the item to an agent chosen at random is efficient), a stronger result holds. As long as s and t are independently distributed conditional on the true valuation (which is unknown), all but an arbitrarily small fraction of the rents may be extracted. This result may be found in McAfee, McMillan, and Reny (1989).

5. CONCLUSION

We have examined the robustness of mechanism design solutions when independence of information does not hold. We found that private information is often worthless; it does not lead to rents for its possessors in a variety of contexts.

A common reaction to this paper focuses attention on the heavy use made of the agents' risk neutrality, and argues that only together do the assumptions of risk neutral agents (RNA) and independently and identically distributed information (iid) combine to make a good proxy for the more difficult real world case of risk averse agents and correlated information. Consider the auction environment. With RNA and iid, any of the usual auction forms maximize the seller's revenue. This is taken as corroboration of the mechanism design paradigm. On the other hand, we know that when either RNA (Maskin and Riley (1984), Matthews (1983)) or identically distributed information (Myerson (1981)) fail, the usual auction forms do not maximize the seller's revenue. From this paper, if independence fails, the usual auction forms do not maximize the seller's revenue. Thus, if either RNA or iid fail to hold, we do not have a mechanism design explanation of the usual auctions. In light of this, it is difficult to believe that if both RNA and iid fail, the mechanism design solution will be similar to the case in which they both hold. We consider this a strong argument that mechanism design has no reasonable explanation of the usual auction forms, and that some other criteria must be invoked.

This begs a substantially more difficult question which we have not addressed. It is a remarkable fact that the English auction (with reserve price) is the solution to a mechanism design problem, maximizing the seller's expected profits, in the independent private values framework. However, minor perturbations of the environment destroy this result. To see this, note that if $f(s|t)$ is a conditional density of s given t satisfying (*) or the hypotheses of any one of Corollaries 1, 2, or 3, then for any $\varepsilon \in (0,1)$ $g_\varepsilon(s|t) \equiv (1-\varepsilon) + \varepsilon f(s|t)$ is also a conditional density of s given t and satisfies the hypotheses of one of our theorems. Furthermore, g_ε converges in sup norm to the independence case as ε goes to zero. This indicates (at least to us) that the prevalence of the English auction in selling items whose value is uncertain is almost certainly not due to the fact that sellers are maximizing expected revenue.

The English auction does possess some important features. Milgrom and Weber (1982) showed that because the English auction reveals a lot of information as bidders drop out of the bidding, prices are pushed higher on average. Moreover, we suspect that the English auction does well in a variety of circumstances precisely because it does not depend, as a selling mechanism, on information about the specific environment, such as densities of valuations, etc.

We are not surprised that the mechanisms described in this paper are not in common use, because these mechanisms (the z functions) will generally be sensitive to the environment's description (e.g. f).[9] Thus, this paper is really more about economists' models of asymmetric information than about asymmetric information itself, since generally the description of the environment, at the level of detail required by mechanism designers, is absurd.

Therefore, a reasonable question for the mechanism design literature is how to capture the importance of robustness. Specifically, we think the answer to questions like "under what circumstances are English auctions used?" has much to do with the need for an institution to perform "well" in a variety of circumstances. Indeed, one might well imagine that the circumstances are at least partially determined by the institution. That is, English auctions will attract buyers who prefer the English auction over another selling institution, and thus the choice of mechanism affects the distributions. One cannot hold the distributions fixed in the experiment of choosing the mechanism.[10]

These concerns have led several authors (Holmstrom and Milgrom (1987), McAfee and McMillan (1987b), Laffont and Tirole (1985)) to look for environments in which "simple" contracts or mechanisms are optimal in a wide variety of circumstances. These papers share a theme that a mechanism designer's desire to use complicated mechanisms which exploit aspects of the environment (utility functions, distributions) is reduced by enlarging an agent's action space.

[9] Indeed, it can be easily shown that the "optimal" mechanism, where optimal maximizes one agent's rents, is not continuous in the density of values.

[10] A glimmer of this idea may be found in McAfee and McMillan (1987c), which examines optimal auctions when participation is costly, and the participation decision is made after the mechanism is chosen. This destroys the seller's incentive to post a reserve price, common to the literature, and exchange is efficient.

Effectively, an agent can thwart the mechanism by exerting effort at a small cost. The lesson of this paper is that asymmetric information, when combined with risk neutrality, plays a small or nonexistent role in such a research program. Generally, in environments with correlated information, the importance of private information is near zero.

Department of Economics, University of Texas, Austin, TX 78712, U.S.A.
and
Department of Economics, University of Western Ontario, Social Science Center, London N6A5C2, Canada

Manuscript received August, 1988; final revision received July, 1991.

APPENDIX

PROOF OF LEMMA 1: ($\supseteq$): Fix $s_0 \in [0,1]$ and $\varepsilon > 0$. For $s \in [0,1]$, let

$$z(s) = \begin{cases} 1/2\alpha & \text{if} \quad s_0 - \alpha \leqslant s \leqslant s_0 + \alpha, \\ 0 & \text{otherwise,} \end{cases}$$

and choose α so that

$$|s - s_0| < \alpha \Rightarrow |f(s|t) - f(s_0|t)| < \varepsilon$$

(this is feasible since f is continuous on a compact set, and hence uniformly continuous). Then

$$\begin{aligned} &\left| \int_0^1 z(s) f(s|t)\, ds - f(s_0|t) \right| \\ &\quad = \left| \int_{s_0-\alpha}^{s_0+\alpha} \frac{1}{2\alpha} f(s|t)\, ds - f(s_0|t) \right| = \frac{1}{2\alpha} \left| \int_{s_0-\alpha}^{s_0+\alpha} (f(s|t) - f(s_0|t))\, ds \right| \\ &\quad \leqslant \frac{1}{2\alpha} \int_{s_0-\alpha}^{s_0+\alpha} |f(s|t) - f(s_0|t)|\, ds \leqslant \frac{1}{2\alpha} \int_{s_0-\alpha}^{s_0+\alpha} \varepsilon\, ds = \varepsilon. \end{aligned}$$

Thus, for $s_0 \in [0,1]$, $f(s_0|\cdot) \in \bar{R}(f)$. Since $\bar{R}(f)$ is closed under linear combinations, we have established one inclusion.

($\subseteq$): Since z, f are continuous, $\forall \varepsilon > 0 \exists s_1, \ldots, s_k$ such that for all $t \in [0,1]$,

$$\left| \int_0^1 z(s) f(s|t)\, ds - 1/k \sum_{i=1}^{k} z(s_i) f(s_i|t) \right| < \varepsilon.$$

Thus $\int_0^1 z(s) f(s,\cdot)\, ds \in \overline{[\{f(s|\cdot)|s \in [0,1]\}]}$ as desired. *Q.E.D.*

REFERENCES

BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. New York: John Wiley and Sons.

COHN, D. (1980): *Measure Theory*. Boston: Birkhäuser.

CRÉMER, J., AND R. MCLEAN (1985): "Optimal Selling Strategies Under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent," *Econometrica*, 53, 345–361.

——— (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1257.

DEMOUGIN, D. (1987): "A Principal-Agent Model with Ex Post Pareto Optimality," mimeo, University of Toronto.

DUNFORD, N., AND J. T. SCHWARTZ (1958; 1988): *Linear Operators, Part I: General Theory*. New York: John Wiley and Sons; Interscience.

FRIEDMAN, A. (1970): *Foundations of Modern Analysis*. New York: Dover.

HOLMSTROM, B., AND P. MILGROM (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives," *Econometrica*, 55, 303–328.

HOCHSTADT, H. (1973): *Integral Equations*. New York: John Wiley and Sons.

LAFFONT, J. J., AND J. TIROLE (1985): "Auctioning Incentive Contracts," mimeo, MIT.

MASKIN, E., AND J. RILEY (1984): "Optimal Auction with Risk Averse Buyers," *Econometrica*, 52, 1473–1518.

MATTHEWS, S. (1983): "Selling to Risk Averse Buyers with Unobservable Tastes," *Journal of Economic Theory*, 30, 370–400.

MCAFEE, R. P., AND J. MCMILLAN (1987a): "Auctions and Bidding," *Journal of Economic Literature*, 25, 699–738.

——— (1987b): "Competition for Agency Contracts," *Rand Journal of Economics*, 18, 296–307.

——— (1987c): "Auctions with Entry," *Economics Letters*, 23, 343–347.

MCAFEE, R. P., J. MCMILLAN, AND P. RENY (1989): "Extracting the Surplus in Common Value Auctions," *Econometrica*, 57, 1451–1459.

MILGROM, P., AND R. WEBER (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

MYERSON, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

MYERSON, R., AND M. SATTERTHWAITE (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265–282.

RILEY, J., AND W. SAMUELSON (1981): "Optimal Auctions," *American Economic Review*, 71, 381–392.