Optimal Pricing Mechanisms with Unknown Demand

By Ilya Segal*

The standard profit-maximizing multiunit auction intersects the submitted demand curve with a preset reservation supply curve, which is determined using the distribution from which the buyers' valuations are drawn. However, when this distribution is unknown, a preset supply curve cannot maximize monopoly profits. The optimal pricing mechanism in this situation sets a price for each buyer on the basis of the demand distribution inferred statistically from other buyers' bids. The resulting profit converges to the optimal monopoly profit with known demand as the number of buyers goes to infinity, and convergence can be substantially faster than with sequential price experimentation. (*JEL* D42, D44, D82, D83)

Recent advances in information technology—most notably the Internet—have enabled the use of economic allocation mechanisms that had been impractical before. Many goods traditionally sold at posted prices are now sold using auction-like mechanisms, in which buyers express their preferences by making bids. Some Internet web sites, such as eBay.com, use traditional auction mechanisms, such as the English auction. Other web sites have developed new mechanisms. For example, so-called "demand aggregation" sites, such as Mercata.com, LetsBuyIt.com, and eWinWin.com, obtain the price by intersecting the demand curve formed by the buyers' bids with a downward-sloping "price curve."

What is the profit-maximizing pricing mechanism, and does it improve upon posted pricing?[1] The present paper examines this question in the context of selling multiple homogeneous units to buyers with unit demands. First the paper adopts the standard assumption of auction theory that the seller knows the distribution from which the buyers' valuations are drawn. Under this and other standard assumptions, the optimal auction can be represented by intersecting a supply curve submitted by the seller with the demand curve revealed by the buyers' bids, and selling to those buyers whose bids are above the intersection. The seller's profit-maximizing supply curve depends on her cost function as well as on the distribution of buyers' valuations. Furthermore, in two important special cases the seller cannot improve upon a posted price. One such case is when the seller's marginal cost is constant, and so her optimal supply curve is perfectly elastic. The other case is when the number of buyers is large, and by the Law of Large Numbers the seller can predict the aggregate demand curve and the price at which it intersects the optimal supply curve.

The problem ignored by this standard analysis is that in reality, the seller may not know the distribution from which buyers' valuations are drawn, and thus may be unable to calculate the optimal reservation supply curve. A typical example is the sale of tickets or subscriptions to a one-of-a-kind concert or sporting event. Even though there are many identical units for sale, such units have not been sold before and so the seller does not know the potential demand. As emphasized in microeconomic textbooks, in this

* Department of Economics, Stanford University, Landau Economics Building, 579 Serra Mall, Stanford, CA 94305 (e-mail: ilya.segal@stanford.edu). I am especially indebted to Paul Milgrom, who raised the question motivating this research and offered several invaluable suggestions. I am also grateful to Susan Athey, Guido Imbens, Jonathan Levin, Eric Maskin, Steven Tadelis, Hal Varian, Edward Vytlacil, Michael Whinston, and the participants of seminars at Duke University, New York University, the University of Maryland, the University of Pennsylvania, Princeton University, Stanford University, Yale University, and the 2002 Decentralization Conference and 2002 Stanford Institute for Theoretical Economics Summer Workshop for their invaluable comments. Vinicius Carrasco and David Miller provided excellent research assistance.

[1] While this paper focuses on the profit objective, the analysis is also applicable to a socially minded seller who needs to cover a fixed cost of production.

situation "a monopolistic market does not have a supply curve" (Robert S. Pindyck and Daniel L. Rubinfeld, 1995), because the profit-maximizing price depends on the overall shape, and in particular the elasticity, of the demand curve.[2]

This paper proposes a new pricing mechanism that maximizes the seller's profit without requiring prior knowledge of demand. The mechanism is based on the idea that buyers' bids reveal information about the distribution of their valuations. While standard auctions ignore this information, the optimal mechanism uses it for pricing. When the number of buyers is large, the seller learns the distribution precisely, and can price optimally given the revealed distribution.

To ensure that a buyer cannot obtain a better price by misreporting his valuation, he should face a price that depends only on other buyers' bids, and not on his own. Formally, such mechanisms are the only ones satisfying dominant-strategy incentive compatibility and *ex post* individual rationality. This paper characterizes the expected profit-maximizing mechanism satisfying these requirements. In the simple case in which the seller's marginal cost is constant, the optimal mechanism offers each buyer the optimal monopoly price against the demand curve inferred from other buyers' bids.

The proposed mechanism improves substantially upon posted pricing, but is qualitatively different from standard auctions. The key difference is that each buyer's bid has an *informational effect:* it affects other buyers' allocations directly, rather than through his own allocation. In particular, such a mechanism cannot be represented with a supply curve.[3]

With a small number of buyers, the seller's Bayesian prior affects her posterior beliefs about the distribution of valuations, and thereby optimal pricing. The optimal mechanism is thus still not completely "detail-free" in the sense of Robert B. Wilson (1987)—the dependence on the seller's prior is simply pushed to a higher level. However, as the number n of buyers grows, the information revealed by buyers' bids overwhelms the seller's prior. The paper shows that for any consistent estimation of demand and its elasticity, as $n \to \infty$, the seller's expected profit converges to the maximum profit achievable with the knowledge of the true demand distribution. In particular, this holds for Bayesian estimation provided that the prior's support includes the true distribution. This also holds for classical statistical estimation, both parametric and nonparametric. For example, the seller can use the reported empirical distribution of the valuations of all buyers other than i as an estimate of the distribution of buyer i's valuation, and offer buyer i the optimal monopoly price against this distribution.[4]

With a large number of buyers, there are many alternative ways to learn demand and attain the optimal monopoly profit asymptotically. For example, the seller can survey a small proportion of buyers and use their reported valuations to set the optimal price to the remaining buyers. Alternatively, the seller can experiment by pricing to different buyers sequentially and updating the price using purchase history (see, e.g., Philippe Aghion et al., 1991; Leonard J. Mirman et al., 1993; Yongmin Chen and Ruqu Wang, 1999; and Godfrey Keller and Sven Rady, 1999). However, both these strategies set a price to each buyer utilizing less information than the optimal mechanism derived in this paper. In particular, the price offered to a buyer depends only on the information received from the preceding buyers, but not from the subsequent buyers. This "informational inefficiency" may slow down convergence to the optimal monopoly profit, sometimes quite dramatically.

It should be noted that relaxing the "*ex post*" constraints of dominant-strategy incentive compatibility and *ex post* individual rationality to the corresponding "*interim*" constraints of

[2] A similar motivation underlies the analysis of Andrew V. Goldberg et al. (2001). A key difference is that they assume complete ignorance of the buyers' valuations, while in this paper these valuations are drawn from the same, although unknown, distribution. Also, they maximize the *worst-case* revenue (relative to that from the optimal posted price), for which purpose randomized mechanisms strictly dominate deterministic ones.

[3] Multiunit auctions that cannot be represented with a supply curve have also been considered by Yvan Lengwiler (1999) and David McAdams (2002), though with a different motivation.

[4] This mechanism is also suggested by Sandeep Baliga and Rakesh Vohra (2002), in independent and contemporaneous work.

Bayesian incentive compatibility and *interim* individual rationality would allow the seller to extract buyer surplus using mechanisms suggested by Jacques Cremer and Richard P. McLean (1985, 1988). However, such mechanisms are not "detail-free"—they are sensitive to the buyers' knowledge about the distribution and each other's valuations, and a seller who is ignorant of the extent of such knowledge may not want to use them.

The paper is organized as follows: Section I describes the model and the class of mechanisms being considered. Section II characterizes the optimal auction with a known demand distribution, and shows that it can normally be represented as the Vickrey-Groves-Clarke mechanism in which the seller manipulates her supply curve in a way that depends on the demand distribution. The section also examines circumstances under which the seller can do just as well with a posted price. Section III derives the optimal pricing mechanism when the seller does not know the distribution of demand but has a Bayesian prior over it, and so the buyers' valuations are correlated from her viewpoint. Section IV shows that the seller's expected profit converges to the maximum monopoly profit achievable with known demand as the number of buyers goes to infinity. Section V examines the rate of convergence and compares it to that achieved by sequential experimentation mechanisms. Section VI motivates the restriction to *ex post* mechanisms. Section VII concludes and discusses several potential extensions.

I. Setup

A monopolistic seller faces n buyers, each of whom has unit demand.[5] Each buyer $i = 1, \ldots, n$ privately observes his own valuation v_i; this valuation is not observed by the seller or by the other buyers. Buyers' valuations are independently drawn from a distribution F on $[0, \bar{v})$ (where $\bar{v} = \infty$ is allowed), with a positive continuous density function $f(v) = F'(v)$ and a finite expectation $E[v]$. Section II will consider the standard case in which the distribution F is common knowledge, while subsequent sections will suppose that F is not known. Section III, in particular, will suppose that the seller has a Bayesian prior over possible distributions, which makes the buyers' valuations correlated from her viewpoint.

An outcome is described by the allocation of the good and the buyers' payments to the seller. An allocation of the good is a vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$, where $\mathcal{X} = \{0, 1\}$ is the set of a buyer's possible purchases from the seller. The buyers' payments to the seller constitute a vector $\mathbf{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n$. All buyers' utilities as well as the seller's profit are quasilinear in the payments. The seller's cost of producing X units is $C(X)$.

By the Revelation Principle, the seller can restrict attention to direct revelation mechanisms, which ask each buyer to bid his valuation, and which ensure that all buyers participate and bid truthfully in equilibrium. To simplify exposition, we focus on deterministic mechanisms, which specify an allocation rule $\mathbf{x}$: $[0, \bar{v})^n \to \mathcal{X}^n$ and a payment rule $\mathbf{t}$: $[0, \bar{v})^n \to \mathbb{R}^n$. While the seller may sometimes gain by conditioning the outcome on a public randomization, this will prove not to be the case in the settings considered in this paper.

It is customary to impose the Bayesian Incentive Compatibility (BIC) and *Interim* Individual Rationality (IIR) constraints on the mechanism. Formally, the constraints say that for any buyer i and any $v_i, \hat{v}_i \in [0, \bar{v})$,

$$\text{(BIC)} \quad E_{\mathbf{v}_{-i}|v_i}[v_i x_i(\mathbf{v}) - t_i(\mathbf{v})] \geq E_{\mathbf{v}_{-i}|v_i}[v_i x_i(\hat{v}_i, \mathbf{v}_{-i}) - t_i(\hat{v}_i, \mathbf{v}_{-i})],$$

$$\text{(IIR)} \quad E_{\mathbf{v}_{-i}|v_i}[v_i x_i(\mathbf{v}) - t_i(\mathbf{v})] \geq 0.$$

Mechanisms satisfying these constraints will be called *interim mechanisms.*

This paper, however, will focus on mechanisms satisfying the stronger requirements of *Dominant-strategy* Incentive Compatibility (DIC) and *Ex post* Individual Rationality (EIR). Formally, for any buyer i, any valuation profile $\mathbf{v} \in [0, \bar{v})^n$, and any $\hat{v}_i \in [0, \bar{v})$,

[5] With obvious alterations the analysis could be applied to the problem of procuring from n sellers, each of whom has unit supply.

$$\text{(DIC)} \quad v_i x_i(\mathbf{v}) - t_i(\mathbf{v}) \geq v_i x_i(\hat{v}_i, \mathbf{v}_{-i}) - t_i(\hat{v}_i, \mathbf{v}_{-i}),$$

$$\text{(EIR)} \quad v_i x_i(\mathbf{v}) - t_i(\mathbf{v}) \geq 0.$$

These constraints require that each buyer's incentives to participate and bid truthfully are satisfied *ex post* (for any possible realization of other buyers' valuations), rather than just *interim* (on expectation over these valuations). Mechanisms satisfying (DIC) and (EIR) will be called *ex post mechanisms. Ex post* mechanisms are also studied by Kim-Sau Chung and Jeffrey C. Ely (2001), in the more general case of interdependent valuations.

In the standard auction theory setup in which buyers' valuations are independently drawn from a known distribution F, the restriction to *ex post* mechanisms typically does not reduce the seller's expected profit, as explained in Section II below. However, when the distribution F is unknown, and so the valuations are correlated from the seller's viewpoint, the restriction does reduce the expected profit. Yet, *ex post* mechanisms have the important advantage of being robust to the buyers' beliefs about each other's valuations. A motivation for *ex post* mechanisms along these lines is offered in Section VI.

Ex post mechanisms have a very simple characterization:

LEMMA 1: *A deterministic mechanism* $\langle \mathbf{x}(\cdot), \mathbf{t}(\cdot) \rangle$ *is an ex post mechanism if and only if for each buyer i there exist functions* $p_i, s_i : [0, \bar{v})^{n-1} \rightarrow \mathbb{R}_+$ *such that for every valuation profile*[6] $\mathbf{v} \in [0, \bar{v})^n$,

$$x_i(\mathbf{v}) = \begin{cases} 1 & \text{if } v_i > p_i(\mathbf{v}_{-i}) \\ 0 & \text{if } v_i < p_i(\mathbf{v}_{-i}), \end{cases} \quad \text{and}$$

$$t_i(\mathbf{v}) = p_i(\mathbf{v}_{-i}) x_i(\mathbf{v}) - s_i(\mathbf{v}_{-i}).$$

PROOF:

The "if" part is easy to verify. The "only if" part follows from the Taxation Principle (see, e.g., Bernard Salanie, 1997), which under (DIC) allows us to represent the mechanism faced by buyer i for any given profile $\mathbf{v}_{-i}$ of other buyers' bids as a nondecreasing tariff $T_i(\cdot, \mathbf{v}_{-i}) : \mathcal{X} \rightarrow \mathbb{R}$. Let $s_i(\mathbf{v}_{-i}) = -T_i(0, \mathbf{v}_{-i})$ and $p_i(\mathbf{v}_{-i}) = T_i(1, \mathbf{v}_{-i}) - T_i(0, \mathbf{v}_{-i})$. (EIR) for $v_i = 0$ requires that $s_i(\mathbf{v}_{-i}) \geq 0$.

The mechanism described in Lemma 1 offers each buyer i a lump-sum subsidy $s_i(\mathbf{v}_{-i}) \geq 0$ and a price $p_i(\mathbf{v}_{-i}) \geq 0$ that depend on other buyers' reports. Buyer i receives a unit at this price if and only if the price is below his reported valuation. Such mechanisms will be called *pricing mechanisms,* and the functions $p_i(\cdot)$ and $s_i(\cdot)$ will be called the *pricing* and *subsidy rule,* respectively.

Lemma 1 implies that any allocation rule that is implemented in a deterministic *ex post* mechanism satisfies the following monotonicity condition:

$$\text{(M)} \quad x_i(v_i, \mathbf{v}_{-i}) \text{ is nondecreasing in } v_i \text{ for all } i, \text{ all } \mathbf{v}_{-i} \in [0, \bar{v})^{n-1}.$$

In the unique pricing rule implementing such an allocation rule, the price to each buyer equals the minimum bid procuring him a unit:

$$(1) \quad p_i(\mathbf{v}_{-i}) = \inf\{v_i \in [0, \bar{v}) : x_i(v_i, \mathbf{v}_{-i}) = 1\}.$$

As for the subsidies, a profit-maximizing seller will set them identically to zero.

One example of a pricing mechanism is the Vickrey-Groves-Clarke mechanism, in which $p_i(\mathbf{v}_{-i})$ equals the externality that giving buyer i a unit imposes on the others. Another example is a *posted-price mechanism,* in which $p_i(\mathbf{v}_{-i}) \equiv p^*$ for all i, i.e., buyers face a single price that does not depend on any bids.

Observe that any pricing mechanism could in principle be implemented with a two-stage procedure, in which (1) buyers report their valuations $(v_1, \ldots, v_n)$, and (2) each buyer i decides whether to purchase at price $p_i(\mathbf{v}_{-i})$. Since a buyer's stage 1 report has no effect on the price he faces in stage 2, truthtelling is a weak equilibrium at stage 1. However, there are several

[6] The consumption $x_i(\mathbf{v})$ for $v_i = p_i(\mathbf{v}_{-i})$ is left indeterminate, which is not important because the probability of this occurrence is zero.

concerns with such implementation. One concern is that a buyer who has an (arbitrarily small) cost of learning his valuation would not expend the cost at stage 1, expecting to avoid it when offered a very high or very low price at stage 2. Another concern is that arbitrarily small bribes could induce collusion at stage 1. For these (unmodeled) reasons, it is preferable to eliminate the buyers' discretion at stage 2, instead determining their purchases on the basis of their reported valuations. This makes truth-telling *uniquely* optimal for a buyer who faces sufficient uncertainty about other buyers' reports.

II. The Optimal Mechanism with a Known Distribution

This section describes the optimal mechanism when the distribution F is known by the seller. This problem was first analyzed by Roger B. Myerson (1981) for a single-unit setting, and the analysis was extended to the multiunit case by Jeremy I. Bulow and John Roberts (1989). Here we offer new characterizations of the optimal mechanism for important special cases, and provide a useful benchmark for the subsequent analysis of the case in which F is unknown.

A. *The Virtual Surplus Representation*

By the Revenue Equivalence Theorem, the allocation rule $\mathbf{x}(\cdot)$ fully determines the information rents of buyers in any Bayesian incentive-compatible mechanism in which the participation constraints of zero-valuation buyers bind. The seller's expected profit can be expressed as the difference between the expected social surplus and the sum of buyers' expected information rents. Upon integration by parts, this difference can be written as the expectation of the *virtual surplus*:

$$J(\mathbf{x}, \mathbf{v}) = \sum_i m(v_i)x_i - C\left(\sum_i x_i\right),$$

$$\text{where } m(v_i) = v_i - \frac{1 - F(v_i)}{f(v_i)}.$$

The difference between $m(v_i)$, called buyer i's *virtual valuation*, and his true valuation v_i accounts for the buyer's information rent that is not captured by the seller. The function $m(\cdot)$ is called the *marginal revenue function.*[7]

If the seller can use an *interim* mechanism, she chooses an allocation rule to maximize the expected virtual surplus $E_{\mathbf{v}}J(\mathbf{x}(\mathbf{v}), \mathbf{v})$ subject to $E_{\mathbf{v}_{-i}}x_i(\mathbf{v})$ being nondecreasing in v_i for all i, which is the monotonicity constraint stemming from Bayesian incentive compatibility. If the solution turns out to satisfy the stronger monotonicity condition (M), then, according to Lemma 1, it can be implemented with a pricing mechanism, which yields the same expected revenue by the Revenue Equivalence Theorem. Therefore, the restriction to *ex post* mechanisms can hurt the seller only by strengthening the monotonicity constraint to (M).[8] In the cases considered below, this restriction does not reduce the seller's expected profit.

B. *When a Reservation Supply Curve Is Optimal*

We begin by considering the case in which the marginal revenue function $m(\cdot)$ is increasing[9] and the cost function $C(\cdot)$ is convex. For this case, we describe the allocation rule that maximizes the virtual surplus in each state $\mathbf{v}$. Since this allocation rule turns out to satisfy (M), it solves the seller's problem.[10] Since the cost function $C(\cdot)$ is convex, the virtual surplus is maximized by allocating units to buyers in descending order of their virtual valuations

[7] This name, suggested by Bulow and Roberts (1989), comes from the following parallel to the monopoly pricing problem. Thinking of $D(p) = 1 - F(p)$ as the expected demand curve for a given buyer, $R(X) = D^{-1}(X) \cdot X$ is the revenue function, and $m(p) = R'(D(p))$—i.e., the marginal revenue expressed as a function of price.

[8] A more general version of this argument is given by Dilip Mookherjee and Stefan Reichelstein (1992).

[9] A sufficient condition for this is the monotonicity of the hazard rate $f(v)/[1 - F(v)]$, which holds for many important distributions (see Mark Bagnoli and Ted Bergstrom, 1989).

[10] This also implies that a randomized mechanism would not provide an improvement, since the seller's expected profit in such a mechanism can still be written as the expected virtual surplus, while the proposed deterministic mechanism maximizes the virtual surplus for any given valuation profile $\mathbf{v}$.

$m(v_i)$, proceeding while these virtual valuations exceed the incremental cost $C(X) - C(X - 1)$. Since the marginal revenue function $m(\cdot)$ is increasing, the buyers receive units in descending order of their true valuations while the valuations exceed $m^{-1}(C(X) - C(X - 1)) \equiv S(X)$. The function $S(X)$, obtained by transforming the seller's incremental cost curve $C(X) - C(X - 1)$ upward with the inverse marginal revenue function $m^{-1}(\cdot)$, can be interpreted as the seller's *inverse reservation supply curve.* Intersecting this curve with the inverse demand curve reported by the buyers yields the optimal quantity X^*. Formally, X^* is described by

$$(2) \qquad v^{(X^*)} \geq S(X^*) \qquad \text{if } X^* > 0, \text{ and}$$

$$v^{(X^*+1)} \leq S(X^* + 1) \text{ if } X^* < n,$$

where $v^{(X)}$ denotes the Xth-highest order statistic of vector $\mathbf{v}$. Note that a buyer is more likely to receive a unit when he has a higher valuation, therefore the described allocation rule indeed satisfies (M).

By Lemma 1, the *ex post* mechanism implementing the described allocation rule is a pricing mechanism, whose pricing rule is uniquely determined by formula (1).[11] This implies that each buyer receiving a unit pays the price equal to his highest bid that would entail either not producing his unit or giving it to the first runner-up, buyer $X^* + 1$:

$$(3) \qquad p = \max\{S(X^*), v^{(X^*+1)}\}.$$

These conclusions are summarized as follows:

PROPOSITION 1: *Suppose that the buyers' valuations are independently drawn from a known distribution F, the marginal revenue function $m(\cdot)$ is increasing, and the cost function $C(\cdot)$ is convex. Then any optimal mechanism* (*under either ex post or interim constraints*) *allocates units to buyers in descending order of their valuations while the valuations exceed* $S(X) \equiv m^{-1}(C(X) - C(X - 1))$. *The optimal quantity X^* is thus described by* (2). *In the optimal ex post mechanism, losers do not pay, and all winners pay price* (3).

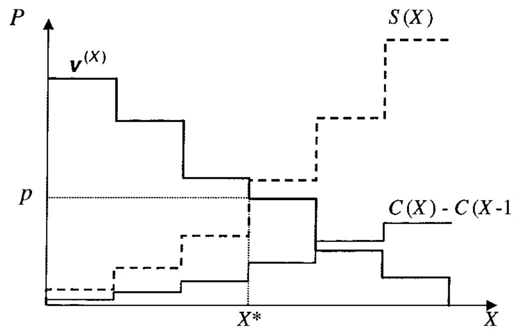

FIGURE 1. STANDARD OPTIMAL AUCTION

The optimal mechanism is thus equivalent to the Vickrey-Groves-Clarke mechanism in which each buyer pays the externality he imposes on others, except that the seller misrepresents his incremental cost to be $S(X) > C(X) - C(X - 1)$. The mechanism is depicted in Figure 1.

Many features of this characterization extend to the case in which the cost function $C(\cdot)$ is not convex. Namely, the virtual surplus-maximizing allocation rule still satisfies (M),[12] and it still allocates units to buyers in descending order of their valuations. The optimal quantity X^* must still satisfy the seller's "discrete first-order conditions" (2), for otherwise she would prefer to sell either one more or one less unit. Thus, the optimal quantity still lies at an intersection of the reservation supply curve $S(\cdot)$ and the demand curve revealed by the buyers. Observe that when the seller's cost function is not convex, her optimal supply curve is not upward sloping. Auctions with downward-sloping supply curves have been implemented by "demand aggregation" web sites such as Mercata.com, LetsBuyIt.com, and eWinWin.com, presumably reflecting sellers' economies of scale.[13]

[11] The same allocation rule can be implemented by many interim mechanisms.

[12] This can be seen using the Monotone Selection Theorem of Paul R. Milgrom and Chris Shannon (1994).

[13] Complications arise when the revealed demand curve intersects the reservation supply curve more than once, in which case the reservation supply curve alone cannot de-

C. *When a Posted Price Is Optimal*

When the seller's marginal cost is a constant c, her inverse reservation supply curve $S(X) = m^{-1}(c)$ is horizontal. Then the optimal mechanism derived in Proposition 1 reduces to a posted price $p^* = m^{-1}(c)$, which maximizes the expected per capita profit:

$$(4)\quad p^* \in \arg\max_{p \in [0,\bar{v})} \pi(p),\ \pi^* = \pi(p^*),$$

$$\text{where } \pi(p) = (p - c)(1 - F(p)) = \int_p^{\bar{v}} (m(v) - c)f(v)\,dv.$$

While the above argument relies on the assumption of increasing marginal revenue, the optimality of posted pricing is more general:

PROPOSITION 2: *If the buyers' valuations are independently drawn from a known distribution F, and $C(X) = cX$, the posted price p^* is an optimal mechanism among all interim mechanisms.*

PROOF:

Since the seller's program is additively separable across buyers, she can restrict attention to mechanisms in which each buyer i's allocation $x_i(\mathbf{v})$ depends only on his own valuation v_i. In particular, in this case the interim constraints are equivalent to the *ex post* constraints. Randomized mechanisms are not useful because the seller's problem of choosing a nondecreasing randomized allocation rule $x_i : [0, \bar{v}) \to [0, 1]$ to maximize the expected virtual surplus on buyer i is a linear program, whose solution is attained at an extreme point, all of which are deterministic allocation rules. By Lemma 1, the seller can then use a pricing mechanism, with a price p_i to each buyer independent of others' announcements. Finally, by (4), p^* is an optimal price offer to each buyer.

Even if the marginal cost is not constant, a posted price becomes optimal asymptotically as the number n of buyers goes to infinity. For normalization across n, consider the asymptotic setting in which the set of each buyer's possible purchases is $\mathcal{X} = \{0, 1/n\}$. This ensures that the expected demand at any posted price p is $1 - F(p)$ for any n. As $n \to \infty$, by the Strong Law of Large Numbers the empirical demand at price p converges to $1 - F(p)$ almost surely, and the resulting profit converges almost surely to $p(1 - F(p)) - C(1 - F(p))$ [provided that the cost function $C(\cdot)$ is continuous]. The asymptotically optimal posted price can then be defined as

$$(5)\quad p^* \in \arg\max_{p \in [0,\bar{v})} \pi(p),\ \pi^* = \pi(p^*),$$

$$\text{where } \pi(p) = p(1 - F(p)) - C(1 - F(p)).$$

The asymptotic optimality of posted pricing is again the easiest to see in the case of increasing marginal revenue and nondecreasing marginal cost, the optimal mechanism for which is described in Proposition 1. Intuitively, as $n \to \infty$, the reported demand curve converges to $1 - F(p)$ and the reservation supply curve converges to $m^{-1}(C'(X))$; therefore the price at which they intersect converges to p^* (see Figure 2). Thus, the optimal mechanism asymptotically reduces to posting price p^*. This conclusion carries over to a more general setting:

PROPOSITION 3: *Suppose that the buyers' valuations are independently drawn from a known distribution F, and let $D(p) = 1 - F(p)$ and $R(X) = D^{-1}(X) \cdot X$ (the revenue function). Suppose that the cost function $C : [0, 1] \to \mathbb{R}$ is continuous, and that for $X^* = D(p^*)$, there exists $\gamma \in \mathbb{R}_+$ such that*[14]

termine the optimal quantity. (This does not happen at demand aggregation web sites, because their dynamic bidding procedures stop once the first intersection is achieved.) Furthermore, in this case some buyers may be "pivotal," meaning that without them it would be optimal to drop some other buyers so as to switch to a lower intersection point (perhaps even to shut down to save a fixed cost). In the pricing mechanism implementing the optimal allocation rule, pivotal buyers face prices that are different from (3). See Francesca Cornelli (1996) for a characterization of the optimal mechanism in the setting with a fixed cost and a constant marginal cost.

[14] Note that when $X^* \in (0, 1)$, condition (6) implies that $\gamma = R'(X^*) = C'(X^*)$ (provided that the latter derivative exists).

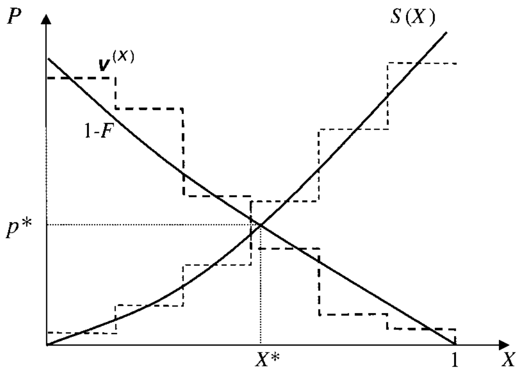

FIGURE 2. ASYMPTOTICS WITH KNOWN DISTRIBUTION

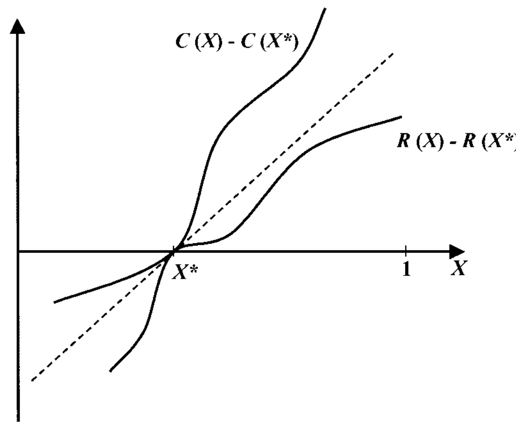

FIGURE 3. THE SEPARATION CONDITION

$$X^* \in \arg\max_{X \in [0,1]}[R(X) - \gamma X], \tag{6}$$

$$X^* \in \arg\max_{X \in [0,1]}[\gamma X - C(X)].$$

Then the seller's expected profit in any interim mechanism with n buyers and $\mathfrak{X} = \{0, 1/n\}$ *cannot exceed* π^**, while her profit from posting price* p^* *converges to* π^* *almost surely as* $n \to \infty$.

PROOF:

Divide the seller's expected profit into two terms, one being as though her marginal cost were constant and equal to γ, and the other being $E[\gamma X - C(X)]$ (where X is the quantity sold by the mechanism). By Proposition 2, the first term is maximized by a posted-price mechanism, and the first line in (6) implies that p^* is an optimal posted price, yielding the maximum value $(p^* - \gamma)X^*$. As for the second term, by the second line in (6) it cannot exceed $\gamma X^* - C(X^*)$. Adding up, we see that the seller's expected profits cannot exceed $p^*X^* - C(X^*) = \pi^*$. On the other hand, as noted above, by the Strong Law of Large Numbers the profit from posting price p^* converges to $\pi(p^*) = \pi^*$ almost surely as $n \to \infty$.

Condition (6) says that the graphs of $R(X) - R(X^*)$ and $C(X) - C(X^*)$ are separated with a straight line passing through the point $(X^*, 0)$ (see Figure 3).[15] By the Separating Hyperplane Theorem, this condition is weaker than the concavity of the revenue function $R(\cdot)$ and the convexity of the cost function $C(\cdot)$, which are assumed in Proposition 1 [as noted in footnote 7, $R'(X) = m(D^{-1}(X))$].

The asymptotic setting considered in Proposition 3, in which the aggregate expected demand is held fixed, should be distinguished from the setting in Dov Monderer and Moshe Tennenholtz (2001) and Zvika Neeman (2001), in which demand grows proportionally to n (e.g., $\mathfrak{X} = \{0, 1\}$ for any n). In the latter setting, the expected demand curve in the limit becomes perfectly elastic at price $\bar{v}$. By posting a price just slightly below $\bar{v}$ and optimally rationing demand at the price, the seller can extract nearly all buyer surplus, while realizing almost all available total surplus as the number of buyers goes to infinity.[16] Monderer and Tennenholtz (2001) and Neeman (2001) instead propose the Vickrey-Groves-Clarke mechanism, which the seller can use to achieve the same profit asymptotically even without knowing $\bar{v}$. The present

[15] The first line in (6) can also be interpreted as saying that the "ironed-out" marginal revenue curve coincides with $R'(X)$ at X^* [since ironing corresponds to the convexification of $R(\cdot)$]. If this does not hold, then profit maximization requires convexification, as discussed in Bulow and Roberts (1989). With a large n, this convexification can be achieved by posting two different prices to different groups of buyers, and so the seller again need not resort to bidding mechanisms.

[16] If $\bar{v} = \infty$, the seller's profits would be unbounded.

model offers a better approximation of real-life situations with many buyers in which the aggregate demand is well known and is downward sloping, while the individual buyers' valuations are not observed by the seller. In such situations, the Vickrey-Groves-Clarke mechanism asymptotically reduces to posting the competitive equilibrium price, which is clearly suboptimal when demand is not perfectly elastic.

D. *Extension to Asymmetric Buyers*

While we have assumed that all buyers are a priori identical, much of the analysis extends to the case where the valuations of different buyers are independently drawn from different (and known) distributions. In this case, the optimal mechanism may no longer be representable with a supply curve, because a low-valuation buyer can have a higher virtual valuation than another buyer with a higher valuation, and so the units may not be allocated in the order of the buyers' true valuations. Thus, Proposition 1 no longer holds. We abstain from describing the optimal mechanism for the asymmetric case, because an even more general setting with correlated valuations is examined in Section III, subsection A, below.

On the other hand, Propositions 2 and 3 can be generalized to the asymmetric case. With a constant marginal cost, the optimal mechanism may post different prices to different buyers—i.e., engage in third-degree price discrimination—but still does not use buyers' bids. The same is true with a general cost function, provided that there is only a finite number of observable buyer types, and there are sufficiently many buyers of each type with independently and identically distributed (i.i.d.) valuations so that their total demand is predictable by the Law of Large Numbers. If the demand of each type satisfies the first separability condition in (6) and the cost function satisfies the second condition in (6), then asymptotically it is again optimal to use (discriminatory) posted pricing rather than a bidding mechanism.

III. The Bayes Optimal Mechanism with Unknown Distribution

Now we turn to the mechanism design problem in which the distribution F from which the buyers' valuations are drawn is unknown. This section considers the case in which the seller is endowed with a Bayesian prior over possible distributions. For example, the seller may know that the distribution belongs to a parametric family $\{F(\cdot|\theta)\}_{\theta \in \Theta}$, and have a prior over the parameter θ. Note that the buyers' valuations $(v_1, \ldots, v_n)$, while independent conditional on F, are correlated from the seller's viewpoint.

A. *The Case of a General Joint Distribution of Valuations*

We begin by deriving the optimal mechanism for the case where the buyers' valuations have an arbitrary joint distribution, without regard to the source of their correlation. We again restrict attention to *ex post* mechanisms, though this is no longer without loss (the restriction is motivated in Section VI below). Just as in the independent value case, dominant-strategy incentive compatibility and the binding *ex post* participation constraints of zero-valuation buyers pin down the information rents of each buyer i. Upon integration by parts, the seller's expected profit can be expressed as the expectation of the virtual surplus

$$(7)\qquad J(\mathbf{x}, \mathbf{v}) = \sum_i m_i(\mathbf{v})x_i - C\left(\sum_i x_i\right),$$

$$\text{where } m_i(\mathbf{v}) = v_i - \frac{1 - \hat{F}_i(v_i|\mathbf{v}_{-i})}{\hat{f}_i(v_i|\mathbf{v}_{-i})}.$$

The only difference from the independent value case is that buyer i's virtual valuation $m_i(v_i, \mathbf{v}_{-i})$ is calculated using the conditional distribution and density functions $\hat{F}_i(\cdot|\mathbf{v}_{-i})$ and $\hat{f}_i(\cdot|\mathbf{v}_{-i})$ respectively, and so it depends on other buyers' valuations as well as his own. The seller's problem again takes the form of maximizing the expected virtual surplus subject to the monotonicity constraint (M).

When the marginal cost is constant, the seller's program is additively separable across buyers. With an appeal to Proposition 2, the optimal mechanism reduces to setting the optimal price to each buyer using the information gleaned from other buyers' bids:

PROPOSITION 4: *If $C(X) = cX$, then the Bayes optimal ex post mechanism is a pricing mechanism with the pricing rule*

(8)
$$p_i(\mathbf{v}_{-i}) \in \arg\max_{p \in [0,\bar{v})} (p - c) \cdot [1 - \hat{F}_i(p|\mathbf{v}_{-i})].$$

For more general cost functions, we identify conditions under which the monotonicity constraint (M) does not bind and so the optimal allocation rule is obtained by maximizing the virtual surplus in each state:[17]

PROPOSITION 5: *Suppose that (i) $m_i(\mathbf{v})$ is increasing in v_i, (ii) $m_i(\mathbf{v}) \geq m_j(\mathbf{v})$ implies that $m_i(v_i', \mathbf{v}_{-i}) > m_j(v_i', \mathbf{v}_{-i})$ for all $v_i' > v_i$, and (iii) $C(\cdot)$ is convex. Then any Bayes optimal ex post mechanism allocates units to buyers in descending order of their virtual valuations while they exceed the incremental cost. Thus, the optimal quantity X^* is described by*

$$m^{(X^*)}(\mathbf{v}) \geq C(X^*) - C(X^* - 1) \quad \text{if } X^* > 0,$$

$$m^{(X^*+1)}(\mathbf{v}) \leq C(X^* + 1) - C(X^*) \quad \text{if } X^* < n.$$

The losers in the mechanism do not pay, and the price p_i paid by a winner i satisfies

(9)
$$m_i(p_i, \mathbf{v}_{-i}) = \max\{C(\hat{X}_i) - C(\hat{X}_i - 1),\ m^{(\hat{X}_i+1)}(p_i, \mathbf{v}_{-i})\},$$

where $\hat{X}_i$ is the largest optimal quantity for the valuation profile $(p_i, \mathbf{v}_{-i})$.

PROOF:

Due to (iii), the proposed allocation rule maximizes the virtual surplus in every state. The allocation rule satisfies (M), because raising a buyer's valuation increases both his virtual valuation by (i) and its rank among all virtual valuations by (ii), thus making him more likely to receive a unit. (This implies that a randomized mechanism would not be useful, by the same argument as in footnote 10.) According to (1), each winner i in the mechanism pays the price p_i equal to his lowest bid that would procure him a unit. When buyer i bids exactly p_i, the seller is indifferent between serving him and either giving his unit to the first runner-up or not producing it at all. This is described in (9), with $\hat{X}_i$ representing the optimal quantity when buyer i is still served in this situation.

Proposition 5 generalizes Proposition 1 to the case of a general joint distribution of valuations. The new condition (ii) ensures that buyer i's marginal revenue function crosses buyer j's at most once, and from below, as buyer i's valuation increases. With independent valuations, a buyer's virtual valuation depends only on his own valuation, and so condition (ii) is implied by condition (i).[18]

The calculation of the optimal quantity is straightforward, and in the special case of symmetrically and independently distributed valuations studied in Proposition 1 it agrees with (2). However, the calculation of prices is more complicated than in Proposition 1. In the case of independent valuations, a reduction in buyer i's bid from v_i to p_i does not affect the allocations of other buyers as long as buyer i is still served. Thus we could take $\hat{X}_i = X^*$ in the pricing formula (9), which yields formula (3) for the symmetric case. In the general correlated case, however, a reduction in buyer i's bid affects other buyers' virtual valuations and thus the quantity sold, even if buyer i still receives a unit. For this reason, identifying the price to buyer i now requires solving a system of two equations with two unknowns, $\hat{X}_i$ and p_i.

Note that in the optimal mechanism, a buyer's bid v_i has an *informational effect:* it affects other buyers' allocations $\mathbf{x}_{-i}(\mathbf{v})$ even when it does not affect his own allocation $x_i(\mathbf{v})$. This

[17] Giuseppe Lopomo (2001) offers a related characterization of the profit-maximizing *ex post* mechanism for selling a single object, the buyers' valuations for which may be interdependent.

[18] Observe also that when the joint distribution of valuations is symmetric with respect to the buyers (while not necessarily a product distribution), $m_i(\mathbf{v}) = m_j(\mathbf{v})$ whenever $v_i = v_j$, and therefore condition (ii) simply says that the buyers' virtual valuations are ordered in the same way as their true valuations.

informational effect cannot arise in auctions that are optimal for the standard case of independent valuations. In such auctions, as well as in most mechanisms observed in real life (such as those represented with a supply curve), a buyer's bid can affect other buyers' allocations only through his own allocation.

The informational effect identified here also arises in *efficient* mechanisms in the case where the buyers' valuations are interdependent, studied by Lawrence M. Ausubel (1999), Partha Dasgupta and Eric Maskin (2000), and Motty Perry and Philip J. Reny (2002). For example, with interdependent valuations, the efficient allocation of a unit between buyers i and j may depend on the information of a losing buyer k, and this dependence cannot be implemented with a standard auction. Though the present model has purely private values, a buyer's valuation does convey information about other buyers' valuations *to the seller,* which gives rise to the interdependence of the buyers' *virtual* valuations, thus creating an informational effect of messages in the profit-maximizing mechanism.[19]

B. *The Case of Affiliated Valuations*

The analysis is simplified when the buyers' valuations are *affiliated,* as defined by Milgrom and Robert J. Weber (1982). In this case, condition (ii) of Proposition 5 can be dispensed with. Indeed, with affiliated valuations, an increase in $\mathbf{v}_{-i}$ (weakly) increases the conditional distribution $\hat{F}_i(v_i|\mathbf{v}_{-i})$ in the monotone likelihood ratio order, and therefore (weakly) reduces the distribution's hazard rate $\hat{f}_i(v_i|\mathbf{v}_{-i})/(1 - \hat{F}_i(v_i|\mathbf{v}_{-i}))$ (see Louis Eeckhoudt and Christian Gollier, 1995, Lemmas 1, 2). This implies that buyer i's virtual valuation $m_i(\mathbf{v})$ is nonincreasing in $\mathbf{v}_{-i}$, which, together with the proposition's condition (i), implies its condition (ii).

With affiliated valuations, we can also say more about the pricing rule in the optimal mechanism described in Proposition 5. Since an increase in $\mathbf{v}_{-i}$ reduces buyer i's virtual valuation, while raising the other buyers' virtual valuations by the proposition's condition (i), buyer i becomes less likely to receive a unit for any given v_i. By (1), this implies that the pricing rule $p_i(\mathbf{v}_{-i})$ is nonincreasing in $\mathbf{v}_{-i}$. Intuitively, an increase in $\mathbf{v}_{-i}$ raises in a stochastic sense the posterior distribution of buyer i's valuation, which, coupled with the fact that the other buyers are now more deserving of the good, makes it optimal to raise the price to buyer i.

When the buyers' valuations are drawn independently from an unknown distribution, they are affiliated if the family $\{F(\cdot|\theta)\}_{\theta\in\Theta}$ of possible distributions is ordered in the monotone likelihood ratio order. By symmetry, in this case the pricing rule $p_i(\cdot)$ is now the same for all buyers. This implies that for any two buyers with valuations $v_i > v_j$ in a given valuation profile $\mathbf{v}$, buyer i faces a lower price than buyer j because $\mathbf{v}_{-i} = (v_j, v_{-i-j})$ is lower than $\mathbf{v}_{-j} = (v_i, v_{-i-j})$. When the valuations are *strictly* affiliated, a buyer's virtual valuation is *strictly* decreasing in the others' valuations, and the same chain of arguments implies that higher-valuation buyers pay *strictly* lower prices. Contrast this with the case where valuations are drawn independently from a *known* distribution F, in which, by Proposition 1, all winners pay the same price.

C. *A Parametric Example*

Let the buyers' valuations be drawn independently from an exponential distribution:[20] $F(v|\theta) = 1 - e^{-\theta v}$, with the hazard parameter $\theta > 0$ not known by the seller. To simplify analysis, suppose that the seller's prior over θ lies in the conjugate family to exponential distributions, which, according to Morris H. DeGroot (1970, p. 166), consists of gamma

[19] To make the analogy precise, imagine that, with v_i representing buyer i's private signal, his valuation is given by $m_i(\mathbf{v})$. Then the function $J(\mathbf{x}, \mathbf{v})$ in (7) describes the social surplus, and so Proposition 5 gives the *surplus*-maximizing allocation rule. Under the proposition's assumptions, this allocation rule satisfies (M), and it is implementable in an *ex post* mechanism (as defined by Chung and Ely, 2001, for the interdependent value case). However, this mechanism differs from that described in Proposition 5: since buyer i's valuation is now $m_i(\mathbf{v})$, his truthful reporting of v_i is induced by charging him price $m_i(p_i, \mathbf{v}_{-i})$, where p_i is the minimum report procuring him a unit (characterized in Proposition 5).

[20] This demand model has been considered by Jeffrey M. Perloff and Steven C. Salop (1985).

distributions. A gamma distribution of θ is defined by a density function of the form[21]

$$\mu(\theta|\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta},$$

with parameters $\alpha, \beta > 0$.

More precisely, if the prior distribution of θ is a gamma distribution with parameters (α_0, β_0), then its posterior conditional on a vector $\mathbf{v}_{-i}$ of $n-1$ independent draws from $F(\cdot|\theta)$ is also a gamma distribution, with parameters $(\alpha, \beta) = (\alpha_0 + n - 1, \beta_0 + \Sigma_{j\neq i} v_j)$. The posterior distribution of $v_i|\mathbf{v}_{-i}$ can then be calculated as

$$\hat{F}_i(v_i|\mathbf{v}_{-i})=\int_0^{\infty}F(v_i|\theta)\mu(\theta|\mathbf{v}_{-i})\,d\theta$$

$$=1-\frac{\beta^{\alpha}}{(\beta+v_i)^{\alpha}}.$$

It is easy to verify that the marginal revenue of this distribution is increasing in v_i, thus condition (i) of Proposition 5 is satisfied. Its condition (ii) is also satisfied by the argument in subsection B, because the family of exponential distributions is ordered in the monotone likelihood ratio order. Thus, the optimal *ex post* mechanism for this setting with a convex cost function is described in Proposition 5.

When the marginal cost is a constant c, the optimal price to each buyer i solves (8), which yields

$$p_i(\mathbf{v}_{-i})=\frac{\alpha c+\beta}{\alpha-1}$$

$$=\left(\frac{\alpha_0+n-1}{\alpha_0+n-2}\right)c+\frac{\beta_0+\Sigma_{j\neq i}v_j}{\alpha_0+n-2}.$$

In particular, if the seller lacks any prior information about the demand parameter θ, she could use the improper uniform prior on $\mathbb{R}_+$ given by parameters $(\alpha_0, \beta_0) = (1, 0)$, which yields the pricing rule

$$p_i(\mathbf{v}_{-i})=\frac{n}{n-1}c+\frac{1}{n-1}\Sigma_{j\neq i}v_j.$$

Observe that a similar pricing rule obtains if, instead of updating a Bayes prior, the seller uses maximum likelihood estimation of parameter θ. The log-likelihood of a vector $\mathbf{v}_{-i}$ is

$$(n-1)\log\theta-\theta\Sigma_{j\neq i}v_j,$$

which is maximized by

$$\hat{\theta}(\mathbf{v}_{-i})=\left(\frac{1}{n-1}\Sigma_{j\neq i}v_j\right)^{-1}.$$

If the seller assumes that buyer i's valuation is distributed according to the estimated parameter, i.e., takes $\hat{F}_i(\cdot|\mathbf{v}_{-i}) = F(\cdot|\hat{\theta}(\mathbf{v}_{-i}))$, then program (8) yields the pricing rule

$$p_i(\mathbf{v}_{-i})=c+1/\hat{\theta}=c+\frac{1}{n-1}\Sigma_{j\neq i}v_j.$$

That is, each buyer is offered a price equal to the marginal cost plus the average of other buyers' bids.

Note that as $n \to \infty$, under both Bayesian and maximum likelihood estimation the prices conditional on a "true" parameter value θ_0 converge to the optimal monopoly price for this parameter value. Indeed, by the Strong Law of Large Numbers, $\frac{1}{n-1}\Sigma_{j\neq i} v_j$ converges almost surely to $E[v|\theta_0] = 1/\theta_0$, and therefore $p_i(\mathbf{v}_{-i})$ converges almost surely to $c + 1/\theta_0$, which is the price solving the profit-maximization program (4) for the true distribution $F(\cdot|\theta_0)$. This implies that, as $n \to \infty$, the seller's expected profit converges to the maximum profit from monopoly pricing with known demand.

IV. Convergence

The optimal mechanism derived in Section III depends on the seller's prior. However, as the example in subsection C illustrates, for a large n the prior is overwhelmed by the information obtained from the buyers' bids. As $n \to \infty$, the seller learns the distribution F from which the buyer's valuations are drawn, the prices converge to the optimal posted price for F, and the resulting profit converges to the optimal monopoly profit given F. General for-

[21] Where $\Gamma(\alpha) = \int_0^{\infty} z^{\alpha-1}e^{-z}\,dz$.

mulations of the convergence result are given in this section.

We adopt the "frequentist" approach of classical statistics, assuming the existence of a "true" distribution to be estimated, and examining convergence conditional on this distribution.[22] This approach allows us to dispense with the prior altogether, letting $\hat{F}_i(\cdot|\mathbf{v}_{-i})$ stand for *any* consistent estimator of the true distribution F, and not necessarily the Bayes posterior distribution. For simplicity, from now on we restrict attention to the symmetric case in which the same estimating function $\hat{F}(\cdot|\cdot)$ is used for all buyers. The simplest convergence result obtains for the case of constant marginal cost:[23]

PROPOSITION 6: *Suppose that* $C(X) = cX$; *that for each* $v_i \in [0, \bar{v})$, $\hat{F}(v_i|\mathbf{v}_{-i}) \xrightarrow{p} F(v_i)$ *as* $n \to \infty$; *and that* $v_i(1 - \hat{F}(v_i|\mathbf{v}_{-i})) \xrightarrow{p} 0$ *as* v_i, $n \to \infty$.[24] *Then as* $n \to \infty$, *the expected per capita profit in the pricing mechanism solving program* (8) *converges to the maximum expected per capita profit* π^* *achievable with F known* [*given by* (4)].

For more general cost functions, a similar convergence result can be established for the asymptotic setting in which each unit contains quantity $1/n$, under the assumptions of Proposition 5 ensuring that the optimal mechanism maximizes the virtual surplus state-by-state:

PROPOSITION 7: *Suppose that as* $n \to \infty$, *for each* $v_i \in [0, \bar{v})$, $\hat{F}(v_i|\mathbf{v}_{-i}) \xrightarrow{p} F(v_i)$, $\hat{f}(v_i|\mathbf{v}_{-i}) \xrightarrow{p} f(v_i)$, *and* $m_i(v_i, \mathbf{v}_{-i}) = v_i - (1 - \hat{F}(v_i|\mathbf{v}_{-i}))/\hat{f}(v_i|\mathbf{v}_{-i})$ *is asymptotically uniformly integrable.*[25] *Suppose also that conditions* (*i*)–(*iii*) *of Proposition* 5 *hold and* C: $[0, 1] \to \mathbb{R}$ *is continuous. Then as* $n \to \infty$, *the expected profit in the mechanism described in Proposition* 5 *with n buyers and* $\mathcal{X} = \{0, 1/n\}$ *converges to the maximum expected profit* π^* *achievable asymptotically with F known* [*given by* (5)].

When the distribution estimator $\hat{F}(\cdot|\mathbf{v}_{-i})$ is a posterior distribution obtained by Bayes updating of a parameter θ whose prior distribution is μ, the consistency assumptions of Propositions 6 and 7 are verified for μ-almost all parameter values using Doob's Consistency Theorem. The Theorem states that the Bayes posterior distribution $\theta|\mathbf{v}_{-i}$ converges to the true parameter value θ_0 weakly, in probability, as $n \to \infty$. This in turn implies that the posterior distribution and density functions, $\hat{F}(v_i|\mathbf{v}_{-i}) = E_{\theta|\mathbf{v}_{-i}}F(v_i|\theta)$ and $\hat{f}(v_i|\mathbf{v}_{-i}) = E_{\theta|\mathbf{v}_{-i}}f(v_i|\theta)$, are consistent estimators of the true distribution and density functions, respectively.

Propositions 6 and 7 are also applicable to non-Bayesian estimation. For example, the maximum likelihood estimator of the parameter,

$$(10) \qquad \hat{\theta}(\mathbf{v}_{-i}) \in \arg\max_{\theta \in \Theta} \Pi_{j \neq i} f(v_j|\theta),$$

is consistent under standard assumptions, leading to the consistent distribution and density estimators $\hat{F}(\cdot|\mathbf{v}_{-i}) = F(\cdot|\hat{\theta}(\mathbf{v}_{-i}))$ and $\hat{f}(\cdot|\mathbf{v}_{-i}) = f(\cdot|\hat{\theta}(\mathbf{v}_{-i}))$, respectively. Alternatively, the seller can use nonparametric estimation, the simplest example of which is given by the empirical distribution of $\mathbf{v}_{-i}$:

$$(11) \quad \hat{F}(v|\mathbf{v}_{-i}) = \frac{1}{n-1}\left|\{j \neq i : v_j < v\}\right|.$$

Consistency of this estimator is established by the Glivenko-Cantelli Theorem.

Application of Proposition 7 to nonparametric estimation may be problematic because the estimation would typically yield virtual valuation estimates $m_i(\cdot)$ that fail assumptions (i) and (ii) of Proposition 5, in which case the proposed allocation rule may fail (M). For example, an increase in v_j can raise the hazard rate of the distribution estimate $\hat{F}(v|\mathbf{v}_{-i})$, thus raising buyer i's virtual valuation $m_i(\mathbf{v})$ to such an

[22] If convergence is uniform across possible distributions, then it also implies the convergence of the unconditional expectation of profit given any Bayesian prior over possible distributions.

[23] The statistical concepts and results used below can be found in A. W. van der Vaart (1998). The proofs of this section's results are given in the Appendix.

[24] The last assumption is vacuous when $\bar{v} < \infty$.

[25] The last assumption holds, in particular, when for each v_i, $E_{\mathbf{v}_{-i}|F}(m_i(\mathbf{v}))^2$ is uniformly bounded across n. This can be easily verified in the parametric example in Section III, subsection C.

extent that it becomes optimal to reallocate buyer j's unit to buyer i, violating (M) for buyer j. The problem can be avoided using instead the following mechanism, inspired by Goldberg et al. (2001) and Baliga and Vohra (2002): partition buyers into two equal-sized subsets S_1, S_2, and offer each subset $S = S_1, S_2$ an optimal price against the distribution estimate using the bids from the other subset:

$$p_S(\mathbf{v}_{N\setminus S}) \in \arg\max_{p \in [0,\bar{v})} [p(1 - \hat{F}(p|\mathbf{v}_{N\setminus S})) - C(1 - \hat{F}(p|\mathbf{v}_{N\setminus S}))].$$

Provided that $\hat{F}(\cdot|\mathbf{v}_{N\setminus S})$ is a consistent estimator of the true distribution F, and the profit-maximizing price p^* defined in (5) is unique, by the Theorem of the Maximum the prices $p_S(\mathbf{v}_{N\setminus S})$ to both groups $S = S_1, S_2$ converge to p^* in probability, and therefore the expected profit converges to π^*. However, this pricing mechanism is not informationally efficient, for in setting each price it ignores the information received from half of the buyers.

V. Rates of Convergence

Convergence to the optimal per capita profit π^* is not the only useful asymptotic criterion. In fact, approximating π^* with a large number n of buyers is not at all hard. For example, the seller could experiment on some buyers by offering them different prices, as in Aghion et al. (1991) and Keller and Rady (1999). Alternatively, she could ask some buyers to report their valuations, refraining from selling to them to ensure truthful reporting. Either experimentation on or surveying of a sufficiently large "test group" of buyers would reveal the demand curve and enable the seller to set an approximately optimal price to the remaining buyers. At the same time, when n is large, the size of the "test group" can be small relative to n, ensuring that the per capita profit approaches π^*. This section compares the asymptotic performance of mechanisms such as surveying and experimentation to that of the optimal mechanism, using as the criterion the rate of convergence to the optimal monopoly profit π^* as $n \to \infty$.[26]

It is clear that no mechanism can achieve uniformly faster profit convergence on a set of distributions F than the Bayes optimal mechanism for a prior concentrated on this set (for otherwise it would achieve a higher expected profit than the supposedly optimal mechanism for n large enough). In fact, surveying and experimentation mechanisms are likely to have a slower convergence rate than the optimal mechanism because they both ignore useful information in setting prices. For example, both mechanisms are *sequential pricing mechanisms*, which set the price to a given buyer using only information obtained from the preceding buyers, rather than from all the other buyers. In addition, while the optimal sequential pricing mechanism would offer each buyer i the optimal price $p_i\ (v_1, \ldots, v_{i-1})$ given the preceding buyers' reported valuations,[27] both experimentation and surveying sacrifice profits on the first buyers (by setting a suboptimal price to them or not selling to them at all) in order to acquire information about their valuations.

We examine this intuition in the simple case in which the marginal cost is a constant c, and so the seller's maximum expected profit π^* is given by (4). The expected loss on a given buyer i when his price $p(\mathbf{v}_{-i})$ is determined from $n - 1$ other buyers' bids is

$$L_n = \pi^* - E_{\mathbf{v}_{-i}|F}[\pi(p(\mathbf{v}_{-i}))].$$

By Proposition 6, $L_n \to 0$ as $n \to \infty$. We examine the rate of this convergence using the

[26] The same asymptotic criterion for mechanism design with many agents is adopted by Thomas A. Gresik and Mark A. Satterthwaite (1989) and Tymon Tatur (2001), but their objective is efficiency rather than the designer's profits.

[27] For example, each buyer i could be asked to report his valuation after deciding whether to buy at the quoted price $p_i(v_1, \ldots, v_{i-1})$. However, recall from the discussion at the end of Section I that for buyer i to have a *strict* incentive to report truthfully, the price $p_i(v_1, \ldots, v_{i-1})$ should not be revealed to him until after his report, and he should receive the good at the revealed price if and only if his reported valuation exceeds the price.

following terminology: Two sequences $\{\alpha_n\}_{n=1}^{\infty}$ and $\{\beta_n\}_{n=1}^{\infty}$ *have the same convergence rate* if there exist two positive numbers $\underline{a}$, $\bar{a}$ such that $\alpha_n/\beta_n \in [\underline{a}, \bar{a}] < \infty$ for n large enough. The two sequences satisfy the stronger property of being *asymptotically proportional*, written as $\alpha_n \propto \beta_n$, if $\alpha_n/\beta_n \to a \in (0, +\infty)$ as $n \to \infty$.

The convergence rate of L_n will depend on how large the family $\{F(\cdot|\theta)\}_{\theta \in \Theta}$ of possible demand distributions is. We consider three cases in turn: (1) *hypothesis testing*, in which Θ is a finite set of parameters ("simple hypotheses"), (2) *parametric estimation*, in which Θ is a Euclidean (finite-dimensional) parameter space, and (3) *nonparametric estimation*, in which Θ is an infinite-dimensional space (for example, including all distribution functions of sufficient smoothness). Suppose without loss of generality that all distributions in the family are distinct, and let θ_0 denote the true parameter value, so that the true distribution is $F(\cdot|\theta_0) = F(\cdot)$.

A. *Hypothesis Testing*

In this case, the optimal mechanism achieves exponential convergence to the optimal monopoly profit as $n \to \infty$. For example, the maximum likelihood estimator given by (10) selects a false hypothesis $\hat{\theta}(\mathbf{v}_{-i}) \neq \theta_0$ with an exponentially small probability (this follows from Chernoff's Theorem—see, e.g., Robert J. Serfling, 1980, Sec. 10.3). Therefore, offering buyer i the optimal price $p(\mathbf{v}_{-i}) = p^*(\hat{\theta}(\mathbf{v}_{-i}))$ against this estimator, where

$$(12) \quad p^*(\theta) \in \arg\max_{p \in [0, \bar{v})} (p - c)(1 - F(p|\theta)),$$

yields exponentially small expected loss. Since this pricing mechanism is also available to a Bayesian decision maker, the expected loss in the Bayes optimal mechanism must converge to zero at least as fast conditional on each positive-probability parameter value θ. In fact, the expected loss in the Bayes optimal mechanism is exponentially small because for any full-support prior, the expected posterior probabilities of false hypotheses shrink exponentially (see, e.g., Erik N. Torgersen, 1991, Sec. 1.4).[28] Thus, under both Bayesian and maximum likelihood estimation, the expected per capita loss L_n satisfies

$$(13) \quad \log L_n \propto -n.$$

On the other hand, the expected per capita loss in any sequential pricing mechanism is at least of the order $1/n$, because the mechanism sets a price to buyer 1 without the benefit of any information. Thus, sequential pricing mechanisms converge exponentially slower than the optimal mechanism. The optimal sequential pricing mechanism in fact *achieves* convergence rate n^{-1}, because setting price $p(v_1, \dots, v_{i-1}) = p^*(\hat{\theta}(v_1, \dots, v_{i-1}))$ to each buyer i yields expected loss L_i on this buyer, hence the *total* expected loss is bounded above by $\sum_{i=1}^{\infty} L_i < \infty$, due to (13).

Experimentation can only perform worse than the optimal sequential pricing mechanism because it uses only past buyers' purchases rather than their reported valuations. Here, however, optimal experimentation achieves the same convergence rate as the optimal sequential pricing mechanism, under the generic condition $F(p^*(\theta')|\theta_0) \neq F(p^*(\theta')|\theta)$ for all $\theta, \theta' \in \Theta$. To see this, note that the seller can offer each buyer the optimal price $p^*(\theta)$ given the maximum likelihood estimate of θ that uses only the past purchase observations at the most frequently set price. Since there are only $|\Theta|$ possible prices, buyer i's price will be based on at least $\lfloor (i-1)/|\Theta| \rfloor$ i.i.d. purchase observations. A Chernoff's Theorem argument then again implies that the expected loss L_i on buyer i is exponentially small in i, and therefore $\sum_{i=1}^{\infty} L_i < \infty$, hence the per capita expected loss $\frac{1}{n}\sum_{i=1}^{n} L_i$ converges to zero at rate n^{-1}.

B. *Parametric Estimation*

Assume that $p^* = p^*(\theta_0) > 0$ is a unique solution to the expected profit-maximization

[28] Under the second-order Taylor expansion (14) below, the minimization of Bayesian expected loss yields a price error $p(\mathbf{v}_{-i}) - p^*(\theta_0)$ that is asymptotically proportional to the seller's posteriors on false hypotheses, hence the expected loss is proportional to the square of these posteriors.

program (4). The first-order condition for the program can be written as $m(p^*) = c$, and the second-order condition as $m'(p^*) \geq 0$. Assume that in fact $m'(p^*) > 0$, in which case the second-order Taylor expansion of π around p^* yields

(14)

$$\pi^* - \pi(p) = A(p - p^*)^2 + o((p - p^*)^2),$$

$$\text{where } A = -\tfrac{1}{2}\pi''(p^*) = m'(p^*)f(p^*) > 0.$$

This implies that the expected loss on buyer i is asymptotically proportional to the squared price error, $E(p(\mathbf{v}_{-i}) - p^*)^2$.

Suppose that the seller offers buyer i the optimal price (12) against the maximum likelihood estimator (10) of the parameter: $p(\mathbf{v}_{-i}) = p^*(\hat{\theta}(\mathbf{v}_{-i}))$. Suppose also that the function $p^*(\theta)$ is uniquely defined in a neighborhood of $\theta = \theta_0$, with the gradient $p^*_\theta(\theta_0) \neq 0$.[29] It is well known that under standard regularity conditions, $\sqrt{n}(\hat{\theta}(\mathbf{v}_{-i}) - \theta_0)$ is asymptotically normal with zero mean and a nondegenerate covariance matrix (see van der Vaart, 1998, Sec. 5.5). By the "delta method" based on the first-order Taylor expansion of $p^*(\theta)$ around θ_0 (see van der Vaart, 1998, Theorem 3.1), $\sqrt{n}(p(\mathbf{v}_{-i}) - p^*)$ is also asymptotically normal with zero mean and a positive variance. By the Bernstein-von Mises Theorem, the same asymptotic normality holds also for the Bayes optimal price $p(\mathbf{v}_{-i})$, which can be viewed as the Bayes point estimate of the optimal price p^* with the loss function (14). Therefore, in both cases, $E(p(\mathbf{v}_{-i}) - p^*)^2$ is asymptotically proportional to $1/n$, hence by (14), the per capita expected loss is

$$L_n \propto n^{-1}.$$

The optimal sequential pricing mechanism has a slower convergence rate. Indeed, since the expected loss on buyer i in this mechanism is L_i, the per capita expected loss is

$$\frac{1}{n}\sum_{i=1}^{n} L_i \propto \frac{1}{n}\sum_{i=1}^{n}\frac{1}{i} \propto \frac{1}{n}\int_1^n \frac{di}{i} \propto \frac{\log n}{n}.$$

(The first proportionality is by Cesàro's Theorem and the second by the Integral Test for series—see Thomas John I'Anson Bromwich, 1931.) Thus, here sequentiality slows down convergence by the factor $\log n$. Optimal experimentation may in fact achieve this convergence rate: Intuitively, even if the seller sets the myopically optimal price to each buyer on the basis of past purchase observations, the price will eventually arrive in a neighborhood of the optimal price p^* in which the partial derivative $F_\theta(p|\theta)$ is bounded away from zero, and so the amount of information about θ received from a purchase observation is bounded below. Then the expected loss on buyer i is asymptotically proportional to $1/i$, yielding again the expected per capita loss of the order of $n^{-1}\log n$.

C. *Nonparametric Estimation*

The simplest nonparametric distribution estimator $\hat{F}(v|\mathbf{v}_{-i})$ is the empirical distribution of the other buyers' valuations, given by (11). The price $p(\mathbf{v}_{-i})$ solving program (8) against this distribution is an "M-estimator" of the correct price p^* (see van der Vaart, 1998). Kislaya Prasad (2001) shows that the distribution of $n^{1/3}(p(\mathbf{v}_{-i}) - p^*)$ converges to a distribution with a finite positive variance. Under the assumptions of the previous subsection, (14) implies that the expected per capita loss $L_n \propto n^{-2/3}$.

Faster convergence rates can be achieved using kernel estimation of the density function f, provided that f is smooth. For example, Charles J. Stone (1983) shows that if f is known to be r times continuously differentiable, then the optimal uniform probabilistic convergence rate of the kernel density estimator $\hat{f}(\cdot|\mathbf{v}_{-i})$ to the true density f is $(n^{-1}\log n)^{r/(2r+1)}$. This implies that the optimal price against the estimated distribution converges in probability to p^* at least as fast, and therefore by (14) the expected per capita loss satisfies

$$L_n = O\left(\left(\frac{n}{\log n}\right)^{-\alpha}\right), \text{ where } \alpha = \frac{2r}{2r+1} < 1.$$

[29] By the Implicit Function Theorem, both assumptions hold when $m_\theta(p^*|\theta_0) \neq 0$, where $m(v|\theta) = v - (1 - F(v|\theta))/f(v|\theta)$.

Optimal sequential pricing mechanisms may in fact achieve the same convergence rate. For example, suppose that $L_n \propto (n/\log n)^{-\alpha}$ or $L_n \propto n^{-\alpha}$, with $\alpha \in (0, 1)$ (recall that empirical distribution estimation yields the latter with $\alpha = 2/3$). In both cases, Cesàro's Theorem implies that

$$\lim_{n\to\infty} \frac{nL_n}{\sum_{i=1}^{n} L_i} = \lim_{n\to\infty} \frac{(n+1)L_{n+1} - nL_n}{L_{n+1}}$$
$$= 1 - \alpha > 0.$$

Thus, $\frac{1}{n} \sum_{i=1}^{n} L_i \propto L_n$; i.e., the expected per capita loss in the optimal sequential mechanism converges at the same rate in as in the fully optimal mechanism.

The optimal experimentation mechanism would be very difficult to characterize in this setting. Intuitively, it appears that its convergence rate may be slower, because the early purchases at prices that are far from p^* will prove useless for fine-tuning the price around p^*.

VI. Justifying *Ex Post* Mechanisms

If the *ex post* constraints (DIC) and (EIR) are relaxed to the corresponding *interim* constraints (BIC) and (IIR), the seller is typically able to extract all buyer surplus, while implementing the surplus-maximizing allocation. Cremer and McLean (1988) show how this can be done, even while satisfying (DIC) [but not (EIR)]. Namely, the seller can employ the Vickrey-Groves-Clarke mechanism, but in addition charge each buyer i a participation fee $\phi_i(\mathbf{v}_{-i})$ that depends on other buyers' reports. For a generic joint distribution of valuations, the fee function $\phi_i(\cdot)$ can be chosen so that the expected *interim* payoff of buyer i is zero no matter what valuation v_i he has.[30]

Neeman (2002) notes that the surplus extraction mechanisms of Cremer and McLean (1988) exploit a one-to-one correspondence between a buyer's own valuation and his belief about the others' types. In a more general information structure, two different types of buyer i with different valuations may share the same beliefs about the others' types, in which case it is impossible to fully extract the information rents of both types of buyer i. In the extreme case in which a buyer's valuation does not constrain his beliefs about others, any mechanism that is robust to the buyers' beliefs (as Wilson, 1987, calls it, *detail-free*) must be an *ex post* mechanism, which is formally shown by John O. Ledyard (1978) and Dirk Bergemann and Stephen Morris (2001).

To be sure, if buyers' beliefs stem from their information about each other's valuations, the "second-best" optimal mechanism, rather than being detail-free, will elicit these beliefs. For example, if buyer i knows the distribution F from which other buyers' valuations are drawn, the mechanism can ask this buyer to set the optimal monopoly price to the other buyers. However, the seller may be wary of using this mechanism if she is not sure how well-informed buyer i is about F. For the same reason, she might also be wary of using the Cremer-McLean mechanism described above. More generally, a seller who is "ignorant" about the buyers' knowledge of each other's valuations (while being confident that they are drawn independently from an unknown distribution) may be concerned with the mechanism's worst-case performance over all information structures. I conjecture that such worst-case performance is maximized by an *ex post* mechanism that elicits only the buyers' valuations and not their beliefs.

VII. Conclusion

This paper has examined the profitability of bidding mechanisms relative to posted pricing. The advantage of bidding mechanisms is that they create *interdependence* among buyers, whereby one buyer's bid v_i affects other buyers' allocations $\mathbf{x}_{-i}$. In the standard auction theory setting in which the buyers' valuations are independently drawn from known distributions, interdependence is desirable to the extent that the seller's cost is nonseparable across buyers (in the extreme case, the seller has a capacity

[30] For example, consider the parametric setting of Section III, subsection C, with $C(X) \equiv 0$ (so that the Vickrey-Groves-Clarke mechanism gives the good for free to all buyers). It can be verified that in this setting, charging buyer i the fee $\phi_i(v_j) = \alpha_0 v_j - \beta_0$ (which depends on the report of another buyer $j \neq i$) ensures that his expected surplus in the mechanism is $v_i - E[\phi_i(v_j)|v_i] = 0$ for all v_i.

constraint). Indeed, a buyer's bid v_i affects his allocation x_i, which due to interactions in the seller's cost function affects the other buyers' optimal allocations $\mathbf{x}_{-i}$. However, this reasoning does not apply when the seller's marginal cost is either constant or little affected by a single buyer (e.g., when there are many small buyers). In these practically important cases, interdependence is not useful, hence optimal auctions do not improve upon posted pricing in the standard setting.

Interdependence becomes useful, however, when the buyers' valuations are correlated from the seller's viewpoint, and in particular when they are drawn independently from an *unknown* distribution. In this case, one buyer's bid v_i conveys information to the seller about other buyers' valuations $\mathbf{v}_{-i}$, and therefore affects their optimal allocations $\mathbf{x}_{-i}$ even when it does not affect the buyer's own allocation x_i. The optimal mechanism is thus qualitatively different from standard auctions (in particular, it cannot be represented with a supply curve). Rather, it resembles (but differs from) the efficient mechanisms suggested by Ausubel (1999), Dasgupta and Maskin (2000), and Perry and Reny (2002) for the case of interdependent values.

The mechanisms suggested in the present paper satisfy Wilson's (1987) desideratum of being "detail-free," i.e., robust to buyers' beliefs about each other's valuations. This is ensured by imposing the "*ex post*" constraints of dominant-strategy incentive compatibility and *ex post* individual rationality, which rule out the surplus extraction schemes proposed by Cremer and McLean (1988).

Another dimension of detail-freeness is robustness to the *seller's* beliefs about buyers' valuations. The rationale for this kind of robustness is not as strong: if the seller has some prior information about the distribution of buyers' valuations (for example, from a history of selling similar products), there is no reason not to utilize it in designing the mechanism. At the same time, it is useful to have mechanisms that can be used even when the seller has "no idea" of the distribution from which the buyers' valuations are drawn. Both kinds of mechanisms are suggested in the present paper. While the Bayes optimal mechanism utilizes the seller's prior, this prior become irrelevant with a large number of buyers, and asymptotically the seller does just as well using classical statistical estimation of demand. Non-Bayesian knowledge about the distribution, such as that of its smoothness or functional form, can also be used to accelerate convergence to optimal monopoly profits. Thus, the paper provides a flexible framework allowing the utilization of different kinds of prior knowledge, Bayesian or non-Bayesian, in designing the optimal mechanism.

An important concern for the practical implementation of any novel economic mechanism is whether its participants can understand how the mechanism works. However, a buyer participating in a mechanism proposed in this paper does not need to understand the intricacies of the pricing formula: as long as he believes that his own bid will not affect the price he ends up facing, he will find it optimal to bid his valuation. Of course, it is crucial that the seller's commitment not to base the price offered to a buyer on his own bid be credible. The same commitment issue arises in second-price auctions, and usually it is successfully resolved in real life. For example, eBay.com allows buyers to submit proxy bids, effectively converting the English auction into a second-price auction. The web site explains to buyers that since a buyer's proxy bid will not be used to raise the price above the minimum necessary for him to win the auction, he should bid his true valuation for the object. There is no reason for the same explanation not to be effective for the mechanisms proposed in this paper.

The proposed mechanisms could be made more transparent by allowing buyers to submit and raise their bids over time, while observing the current price they face. Many Internet pricing mechanisms are realized in this dynamic fashion.[31] An interesting distinguishing feature of our mechanism is that in its dynamic realiza-

[31] A downside of such transparency is that it facilitates tacit collusion among buyers (the same is true of other dynamic mechanisms, such as the English auction). For example, at a bid profile at which each buyer receives a unit, no buyer has a strict incentive to raise his bid, even if it is below his valuation. He may even strictly prefer not to raise his bid to avoid retaliation by other bidders. The rules may have to be modified to reward buyers for breaking such collusive equilibria (see, e.g., the suggestions in McAdams, 2002). Note that collusive equilibria are unlikely in the one-shot mechanism, since a buyer with sufficient uncertainty about others' behavior will find it strictly optimal to bid truthfully.

tion, the price facing a buyer can either rise or fall as demand grows. This contrasts with standard auctions, in which price can only rise as demand grows, and with "demand aggregation" mechanisms, in which it can only fall.

The present analysis can be extended in several directions. One extension is to allow buyers to demand more than one unit. The optimal mechanism would in general involve second-degree price discrimination, charging each buyer different prices for different units, as in the model of Maskin and John Riley (1984). When the seller's marginal cost is constant, her problem is again additively separable across buyers, and she should offer each buyer the optimal nonlinear tariff using the information inferred from other buyers' reported preferences. It should be kept in mind, however, that unless the buyers' preferences are seriously restricted a priori (say, to a one-dimensional domain with a single-crossing property), computing the optimal tariff may be quite difficult.

Another possible extension is the addition of value interdependence (i.e., a common-value component) among buyers, which can be analyzed using Chung and Ely's (2001) concept of *ex post* implementation. The buyers in this setting may need to submit more complex bids. For example, with unit demands, they could report functions describing how their valuations depend on those of others, as in the mechanism of Dasgupta and Maskin (2000).

Finally, note that the mechanisms proposed in this paper typically charge different buyers different prices for identical units. This occurs because the price to each buyer is calculated upon excluding this buyer's bid. However, I conjecture that a uniform-pricing mechanism in which the price is calculated using *all* buyers' bids would work just as well when the number of buyers is large. Indeed, each buyer will then realize that his bid's effect on the price is very small, and therefore will bid close to his true valuation.

Appendix: Proofs of Propositions 6 and 7

PROOF OF PROPOSITION 6:

Let $\pi_i(p|\mathbf{v}_{-i})$ denote the objective function in (8). If the seller uses the pricing rule $p(\mathbf{v}_{-i})$ solving (8), her expected loss relative to π^* on buyer i given $\mathbf{v}_{-i}$ is bounded above as follows:

$$\text{(A1)}\quad \pi^* - \pi(p(\mathbf{v}_{-i})) = [\pi(p^*) - \pi_i(p^*|\mathbf{v}_{-i})] + [\pi_i(p^*|\mathbf{v}_{-i}) - \pi_i(p(\mathbf{v}_{-i})|\mathbf{v}_{-i})] + [\pi_i(p(\mathbf{v}_{-i})|\mathbf{v}_{-i}) - \pi(p(\mathbf{v}_{-i}))] \le 2 \sup_{p\in[0,\bar{v})} |\pi(p) - \pi_i(p|\mathbf{v}_{-i})|.$$

In words, the loss is bounded above by twice the supremum absolute difference between the objective functions in (5) and (8). This supremum absolute difference can in turn be bounded above as follows:

$$\sup_{p\in[0,\bar{v})} |\pi(p) - \pi_i(p|\mathbf{v}_{-i})| \le M \sup_{p\in[0,\bar{v})} |\hat{F}(p|\mathbf{v}_{-i}) - F(p)| + \sup_{p\ge M}[p(1-\hat{F}(p|\mathbf{v}_{-i}))] + \sup_{p\ge M}[p(1-F(p))]$$

for any $M > 0$. A simple extension of Lemma 2.11 in van der Vaart (1998) shows that, by the consistency of $\hat{F}(v_i|\mathbf{v}_{-i})$, the first term above goes to zero in probability as $n \to \infty$ for any fixed M. The other assumption on $\hat{F}(v_i|\mathbf{v}_{-i})$ implies that the second term goes to zero in probability as $M, n \to \infty$. Finally, the third term goes to zero as $M \to \infty$ due to the assumption that $E[v|F] < \infty$. Putting together, we see that for all $\varepsilon, \delta > 0$ we can find $M = M(\varepsilon, \delta)$ and a corresponding $\hat{n}(\varepsilon, \delta)$ such that for all $n \ge \hat{n}(\varepsilon, \delta)$, each term is less than $\varepsilon/3$ with probability at least $1 - \delta/3$. This implies that $\Pr\{\sup_{p\in[0,\bar{v})}|\pi(p) - \pi_i(p|\mathbf{v}_{-i})| < \varepsilon\} \ge 1 - \delta$ for $n \ge \hat{n}(\varepsilon, \delta)$, which by (A1) implies that $\pi(p(\mathbf{v}_{-i})) \xrightarrow{p} \pi^*$ as $n \to \infty$. Since $\pi(p(\mathbf{v}_{-i}))$ is bounded, it follows that the expected profit $E_{\mathbf{v}_{-i}|F}[\pi(p(\mathbf{v}_{-i}))] \to \pi^*$ as $n \to \infty$.

PROOF OF PROPOSITION 7:

Note that the proposition's assumptions verify those of Proposition 3, which implies that the maximum expected profit with F known converges to π^* as $n \to \infty$. Therefore, it suffices to show that the expected loss from not knowing F goes to zero as $n \to \infty$.

The allocation rule described in Proposition 5 maximizes the virtual surplus (7) in each state, and therefore maximizes its expectation $E_{\mathbf{v}|F}J(\mathbf{x}(\mathbf{v}), \mathbf{v})$. The seller's expected profit under the true distribution F is instead

$$E_{\mathbf{v}|F}\left[\sum_i m_F(v_i)x_i(\mathbf{v}) - C\left(\sum_i x_i(\mathbf{v})\right)\right],$$

where $m_F(v_i) = v_i - (1 - F(v_i))/f(v_i)$. By an argument similar to that in the beginning of the proof of Proposition 6, the seller's expected loss from not knowing F is bounded above by twice the supremum absolute difference between the two expectations over all allocation rules $\mathbf{x}(\cdot)$. Using symmetry, this supremum absolute difference is in turn bounded above as follows:

$$\sup_{\mathbf{x}:[0,\bar{v})^n \to \{0,1/n\}^n}\left|E_{\mathbf{v}|F}\left[\sum_i [m_i(\mathbf{v}) - m_F(v_i)]x_i(\mathbf{v})\right]\right|$$

$$\leq E_{\mathbf{v}|F}|m_i(\mathbf{v}) - m_F(v_i)|.$$

The consistency of $\hat{F}(v_i|\mathbf{v}_{-i})$ and $\hat{f}(v_i|\mathbf{v}_{-i})$ implies that for each v_i, $m_i(\mathbf{v}) - m_F(v_i) \xrightarrow{p} 0$ as $n \to \infty$. Since the absolute value of the difference is asymptotically uniformly integrable by assumption, Theorem 2.20 in van der Vaart (1998) implies that $E_{\mathbf{v}_{-i}|F}|m_i(\mathbf{v}) - m_F(v_i)| \to 0$ as $n \to \infty$. Take expectation $E_{v_i|F}$ using Lebesgue's Dominated Convergence Theorem to see that the right-hand side of the above inequality goes to zero.

REFERENCES

Aghion, Philippe; Bolton, Patrick; Harris, Christopher and Jullien, Bruno. "Optimal Learning by Experimentation." *Review of Economic Studies*, June 1991, *58*(4), pp. 621–54.

Ausubel, Lawrence M. "A Generalized Vickrey Auction." Working paper, University of Maryland, 1999.

Bagnoli, Mark and Bergstrom, Ted. "Log-Concave Probability and its Applications." Center for Research on Economic and Social Theory Working Paper No. 89-23, University of Michigan, 1989.

Baliga, Sandeep and Vohra, Rakesh. "Market Research and Market Design." Working paper, Northwestern University, 2002.

Bergemann, Dirk and Morris, Stephen. "Robust Mechanism Design." Working paper, Yale University, 2001.

Bromwich, Thomas John I'Anson. *An introduction to the theory of infinite series*. London: Macmillan, 2nd Ed., 1931.

Bulow, Jeremy I. and Roberts, John. "The Simple Economics of Optimal Auctions." *Journal of Political Economy*, October 1989, *97*(5), pp. 1060–90.

Chen, Yongmin and Wang, Ruqu. "Learning Buyers' Valuation Distribution in Posted-Price Selling." *Economic Theory*, August 1999, *14*(2), pp. 417–28.

Chung, Kim-Sau and Ely, Jeffrey C. "*Ex Post* Incentive Compatible Mechanism Design." Working paper, Northwestern University, 2001.

Cornelli, Francesca. "Optimal Selling Procedures with Fixed Costs." *Journal of Economic Theory*, October 1996, *71*(1), pp. 1–30.

Cremer, Jacques and McLean, Richard P. "Optimal Selling Strategies under Uncertainty for a Discriminating Monopolist When Demands Are Interdependent." *Econometrica*, March 1985, *53*(2), pp. 345–62.

______. "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions." *Econometrica*, November 1988, *56*(6), pp. 1247–57.

Dasgupta, Partha and Maskin, Eric. "Efficient Auctions." *Quarterly Journal of Economics*, May 2000, *105*(2), pp. 341–88.

DeGroot, Morris H. *Optimal statistical decisions*. New York: McGraw-Hill, 1970.

Eeckhoudt, Louis and Gollier, Christian. "Demand for Risky Assets and the Probability Ratio Order." *Journal of Risk and Uncertainty*, September 1995, *11*(2), pp. 113–22.

Goldberg, Andrew V.; Hartline, Jason D. and Wright, Andrew. "Competitive Auctions and Digital Goods." 12th Annual ACM-SIAM Symposium on Discrete Algorithms, Washington, DC, 2001, pp. 735–44.

Gresik, Thomas A. and Satterthwaite, Mark A. "The Rate at Which a Simple Market Converges to Efficiency as the Number of Traders Increases: An Asymptotic Result for Optimal Trading Mechanisms." *Journal of Economic Theory*, June 1989, *48*(1), pp. 304–32.

Keller, Godfrey and Rady, Sven. "Optimal Experimentation in a Changing Environment." *Review of Economic Studies*, July 1999, *66*(3), pp. 475–507.

Ledyard, John O. "Incentive Compatibility and Incomplete Information." *Journal of Economic Theory*, June 1978, *18*(1), pp. 171–89.

Lengwiler, Yvan. "The Multiple Unit Auction with Variable Supply." *Economic Theory*, August 1999, *14*(2), pp. 373–92.

Lopomo, Giuseppe. "Optimality and Robustness of the English Auction." *Games and Economic Behavior*, August 2001, *36*(2), pp. 219–40.

Maskin, Eric and Riley, John. "Monopoly with Incomplete Information." *RAND Journal of Economics*, Summer 1984, *15*(2), pp. 171–96.

McAdams, David. "Modifying the Uniform-Price Auction to Eliminate 'Collusive-Seeming Equilibria.' " Working paper, Massachusetts Institute of Technology, 2002.

Milgrom, Paul R. and Shannon, Chris. "Monotone Comparative Statics." *Econometrica*, January 1994, *62*(1), pp. 157–80.

Milgrom, Paul R. and Weber, Robert J. "A Theory of Auctions and Competitive Bidding." *Econometrica*, September 1982, *50*(5), pp. 1089–122.

Mirman, Leonard J.; Samuelson, Larry and Urbano, Amparo. "Monopoly Experimentation." *International Economic Review*, August 1993, *34*(3), pp. 549–63.

Monderer, Dov and Tennenholtz, Moshe. "Asymptotically Optimal Multi-Object Auctions." Working paper, Technion-Israel Institute of Technology, 2001.

Mookherjee, Dilip and Reichelstein, Stefan. "Dominant Strategy Implementation of Bayesian Incentive Compatible Allocation Rules." *Journal of Economic Theory*, April 1992, *56*(2), pp. 378–99.

Myerson, Roger B. "Optimal Auction Design." *Mathematics of Operations Research*, February 1981, *6*(1), pp. 58–73.

Neeman, Zvika. "The Effectiveness of Simple Auctions." Working paper, Boston University, 2001.

______ . "The Relevance of Private Information in Mechanism Design." Working paper, Boston University, 2002.

Perloff, Jeffrey M. and Salop, Steven C. "Equilibrium with Product Differentiation." *Review of Economic Studies*, January 1985, *52*(1), pp. 107–20.

Perry, Motty and Reny, Philip J. "An Efficient Auction." *Econometrica*, May 2002, *70*(3), pp. 1199–212.

Pindyck, Robert S. and Rubinfeld, Daniel L. *Microeconomics*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

Prasad, Kislaya. "Price Asymptotics." Working paper, Florida State University, 2001.

Salanie, Bernard. *The economics of contracts: A primer*. Cambridge, MA: MIT Press, 1997.

Serfling, Robert J. *Approximation theorems of mathematical statistics*. New York: John Wiley & Sons, 1980.

Stone, Charles J. "Optimal Uniform Rate of Convergence for Nonparametric Estimators of a Density Function or its Derivatives," in M. Haseeb Rizvi, Jagdish S. Rustagi, and David Siegmund, eds., *Recent advances in statistics: Papers in honor of Herman Chernoff on his sixtieth birthday*. New York: Academic Press, 1983, pp. 393–406.

Tatur, Tymon. "Asymptotically Optimal Market Mechanisms." CMS-EMS Discussion Paper No. 1315, Northwestern University, 2001.

Torgersen, Erik N. *Comparison of statistical experiments*. Encyclopedia of Mathematics and its Applications, Volume 36. Cambridge: Cambridge University Press, 1991.

van der Vaart, A. W. *Asymptotic statistics*. Cambridge: Cambridge University Press, 1998.

Wilson, Robert B. "Game-Theoretic Approaches to Trading Processes," in T. Bewley, ed., *Advances in economic theory: Fifth world congress*. Cambridge: Cambridge University Press, 1987.