Monopoly with Incomplete Information
Author(s): Eric Maskin and John Riley
Source: *The RAND Journal of Economics*, Vol. 15, No. 2 (Summer, 1984), pp. 171-196
Published by: Blackwell Publishing on behalf of The RAND Corporation
Stable URL: http://www.jstor.org/stable/2555674
Accessed: 18/05/2009 10:57

Rand Journal of Economics
Vol. 15, No. 2, Summer 1984

Monopoly with incomplete information

Eric Maskin*

and

John Riley**

Recent theoretical research on principal-agent relationships has emphasized incentive problems that arise when the parties involved are constrained by either asymmetric information or their inability to monitor each other's actions. Here we concentrate on the former constraint and consider the case of a single principal (the monopolist). The main contribution is to show that, under a separability assumption, strong conclusions can be drawn about the nature of optimal incentive schemes. Although the primary focus is on optimal quantity discounts in a monopolized market, the results also shed light on related topics, such as optimal income taxation and commodity bundling.

1. Introduction

■ In the last few years much theoretical work has considered incentive schemes (or "principal-agent" relationships) where the parties involved are constrained either by asymmetric information or by their inability to monitor each other's actions. In this article we concentrate on the former constraint, known in the literature as "adverse selection," rather than on the latter, often called "moral hazard."

We show that a variety of issues can be viewed formally as members of a single family of principal-agent problems. In this article we consider in detail: (i) price discrimination via quantity discounts (Goldman, Leland, and Sibley, 1980; Oi, 1971; Roberts, 1979; Spence, 1977)[1] and (ii) monopoly pricing of products of differing quality (Mussa and Rosen, 1978). In a companion piece, Maskin and Riley (1983a), we study (iii) the choice of an auction for selling a fixed number of identical items. (See Harris and Raviv (1981), Maskin and Riley (1980, 1984), Milgrom and Weber (1982), Myerson (1981), and Riley and Samuelson (1981), as well as the seminal paper by Vickrey (1961).)

For each of these problems, the central issue is how to construct a sorting mechanism

---

* Massachusetts Institute of Technology.

** University of California at Los Angeles.

The helpful comments of the participants in numerous workshops and, in particular, of David Kreps, David Levine, Paul Milgrom, and Roger Myerson are gratefully acknowledged. This research was supported by the National Science Foundation, the U.K. Social Science Research Council, and the Sloan Foundation.

[1] There is also a much earlier literature which recognizes the potential gains to differential unit pricing for different sized purchases. Pigou (1930), for example, discusses "block pricing." He does not, however, consider its role as a screening mechanism in the absence of complete information. Instead he views it as an approximation to perfect ("first-degree") price discrimination.

to extract the greatest possible private gain. Our main contribution is to show that, under a separability assumption, we can draw strong conclusions about the nature of optimal incentive schemes. Our results also shed new light on a wide range of closely related topics. These include the theory of optimal income taxation (Mirrlees, 1971), the monopoly pricing of insurance (Stiglitz, 1977), the monopoly provision of excludable public goods (Brito and Oakland, 1980), multiproduct monopoly pricing (Mirman and Sibley, 1980), the regulation of a producer with unknown costs (Baron and Myerson, 1982), and monopoly pricing of a product line (Oren, Smith, and Wilson, 1984).

In the following section, we begin with a heuristic discussion of the issues involved for the simplest application: nonlinear pricing. Then, in Section 3 we examine the self-selection constraints that the monopolist faces and show that the choice of an optimal selling procedure reduces to a tractable variational problem. In Section 4 we present the fundamental characterization theorem for the optimal selling strategy of the monopolist. Using this result, we are able to show that, in a broad class of cases, the optimal selling strategy involves pricing larger quantities at successively lower unit prices; that is, using quantity discounts. We also provide a strong result on the effect of a change in the distribution of preferences on the profit-maximizing quantity discount scheme.

Finally, in Section 5, we extend the Mussa and Rosen (1978) analysis of the pricing of quality to preferences that are defined over both quality and quantity. In this case a seller's optimal strategy is to sell multiple units in bundles. Higher quality units are sold in packages of a different size from those of lower quality units. For a simple parameterization of preferences, we provide a complete characterization of the optimal bundling strategy. We conclude in Section 6 with a discussion of further applications of the theory and with some comments on the central assumptions.

2. Description of the problem

■ To introduce the general principles involved we focus on the simplest application—the use of a nonlinear price schedule to discriminate among customers. A monopolist produces a single product at constant marginal cost, c. A buyer of type i has preferences represented by the utility function[2]

$$U_i(q, -T) = \int_0^q p(x; v_i)dx - T, \tag{1}$$

where q is the number of units purchased from the monopolist and T is total spending on these units. That is, we take the standard consumer surplus approach and assume that differences in tastes are captured by the single parameter v. The seller does not observe v, but knows $F(v)$, the distribution of buyers' preferences.[3] Throughout we shall assume that higher levels of v are associated with a higher demand. We also assume that the demand price $p(q; v)$ is decreasing in q and that there is some maximum quantity $q^e(v)$ for which

[2] Thus we exclude income effects. Notice that we do not rule out the possibility that differences in income *across* consumers may account for differences in demand; such effects can be embodied in the parameter v_i. What we are assuming, in effect, is that the proportion of any *single* consumer's income spent on the good in question is so small that variations in p have a negligible effect on income.

[3] Throughout this section, we shall interpret $F(v)$ to be the underlying probability distribution from which the parameters of a population of buyers are drawn independently. Therefore, in considering expected seller profit, we are assuming risk neutrality on the part of the seller. Alternatively we could think of $F(v)$ as the actual—i.e., the realized—distribution of the v's. The analysis would remain unchanged except that profit-maximization would replace expected-profit-maximization. Actually, there is a slight difficulty with the latter (traditional) interpretation of $F(v)$. If all buyers but one revealed their v's through their purchasing behavior, the seller could infer the parameter value of the remaining buyer without observing it directly. This problem arises, however, only if the seller knows the actual distribution *exactly,* which seems improbable in practice.

demand price exceeds marginal cost. For each v, $q^e(v)$ is thus the efficient consumption level.

To be precise we impose the following restrictions.

Assumption 1. (i) For all feasible v the demand price function $p(q; v)$ is nonincreasing in q and nonnegative, and there exists $q^e(v) \geqslant 0$ such that $p(q; v)$ is decreasing in q for $q \leqslant q^e(v)$, and $p(q; v) \geqslant c$ if and only if $q \leqslant q^e(v)$. (ii) $p(q; v)$ is twice continuously differentiable for $q \leqslant q^e(v)$. (iii) $p(q; v)$ is strictly increasing in v whenever $p(q; v)$ is positive.

A selling procedure is then a schedule of pairs $\langle \hat{q}_s, \hat{T}_s \rangle_{s \in S}$, which the seller offers to the buyers. If a buyer chooses s, he receives $\hat{q}_s$ and pays a total of $\hat{T}_s$. The profit or "return" to the seller is then

$$\hat{R}_s = \hat{T}_s - c\hat{q}_s. \tag{2}$$

Throughout we shall assume that any selling procedure includes the pair $\langle 0, 0 \rangle$, that is, the buyer always has the option of buying (and paying) nothing.

Combining (1) and (2), we can rewrite the utility of a buyer of type i as

$$U(q, R; v_i) = \int_0^q p(x; v_i)dx - cq - R \equiv N(q; v_i) - R, \tag{3}$$

where $N(q; v_i)$ is the social surplus generated by the sale. Thus, we can think of the trades between the seller (the "principal") and buyers ("agents") as giving each buyer the entire surplus less a fee R. The selling procedure is then a schedule of pairs $\langle q_s, R_s \rangle_{s \in S}$ offered to each of the buyers. This latter formulation proves more convenient.[4]

Before considering the choice of a selling procedure, we note that our formulation can be reinterpreted, without modification, as the choice of an optimal piece-rate schedule. Let $C(q; v_i)$ be the opportunity cost to a worker of type v_i of producing q units for a potential employer with some monopsony power. Suppose that the employer (the principal) knows the distribution of v_i but cannot observe v_i directly. If m is the value of each unit of output to the employer and the worker is offered a payment Ω for q units, the net gain to the worker is

$$U(q, \Omega; v_i) = \Omega - C(q; v_i), \tag{4}$$

whereas the employer's return is

$$R = mq - \Omega. \tag{5}$$

As before, we define $N(q; v)$ to be the social surplus, that is,

$$N = mq - C(q; v_i). \tag{6}$$

Substituting (5) and (6) in (4), we find that the net gain to an agent (worker) of type i is given once again by (3).[5]

---

[4] Our formulation embraces a vast array of selling procedures that the notation perhaps conceals. We allow, for example, for selling procedures involving multiple rounds of moves by buyer and seller, as long as the only things that ultimately matter to the buyer and seller are the quantity sold and the payment. Two possibilities that our formulation *does* exclude are (i) allowing one buyer's quantity-payment pair to depend on the choices of other buyers and (ii) permitting $q(s)$ and $\hat{T}(s)$ to be random functions of s. As we argue below, however, neither possibility is advantageous for the seller, given our assumptions (see, however, the minor qualification in footnote 3.)

[5] To make the problem exactly equivalent, we also introduce the constraint that the worker will participate only if his utility exceeds some reservation level $\bar{u}$. This description suggests that the relationship between an employer and worker begins only after the worker has observed his v_i (an *ex post* relationship), but we can equally well suppose that the two parties sign a contingent contract before v_i is realized. By making the payment and output contingent on v_i, such a contract serves to share risk. If we assume that the employer is risk neutral and the worker infinitely risk averse and maintain the hypothesis that only the worker can observe the realization of v_i, then this *ex ante* contract generates the same contingent payments and outputs as the *ex post* relationship. This results because, given the worker's extreme risk aversion, the constraint on his utility in the contract is simply that net gain should exceed some minimum, $\bar{u}$, which is the same as in the *ex post* relationship. Hence, the literature on labor contracts with asymmetric information (Hart, 1983) can also be fit into our framework.

Returning to the original interpretation, we begin with a diagrammatic derivation of the seller's optimal price schedule. A buyer's utility from any pair $\langle q, R \rangle$ is, from (3), just the social surplus $N(q; v_i)$ less the seller's profit R. Given our definition of $q^e(v)$ as the efficient level of consumption by a buyer with parameter v_i, it follows that, for any R, $U(q, R; v_i)$ increases with q until it reaches a maximum at $q = q^e(v_i)$.

Thus, indifference curves must be as depicted in Figure 1. Note that at $\langle q, R \rangle$ the slope of the corresponding indifference curve is

$$\left.\frac{dR}{dq}\right|_{dU=0} = -\frac{\partial U}{\partial q}\bigg/\frac{\partial U}{\partial R} = p(q; v) - c.$$

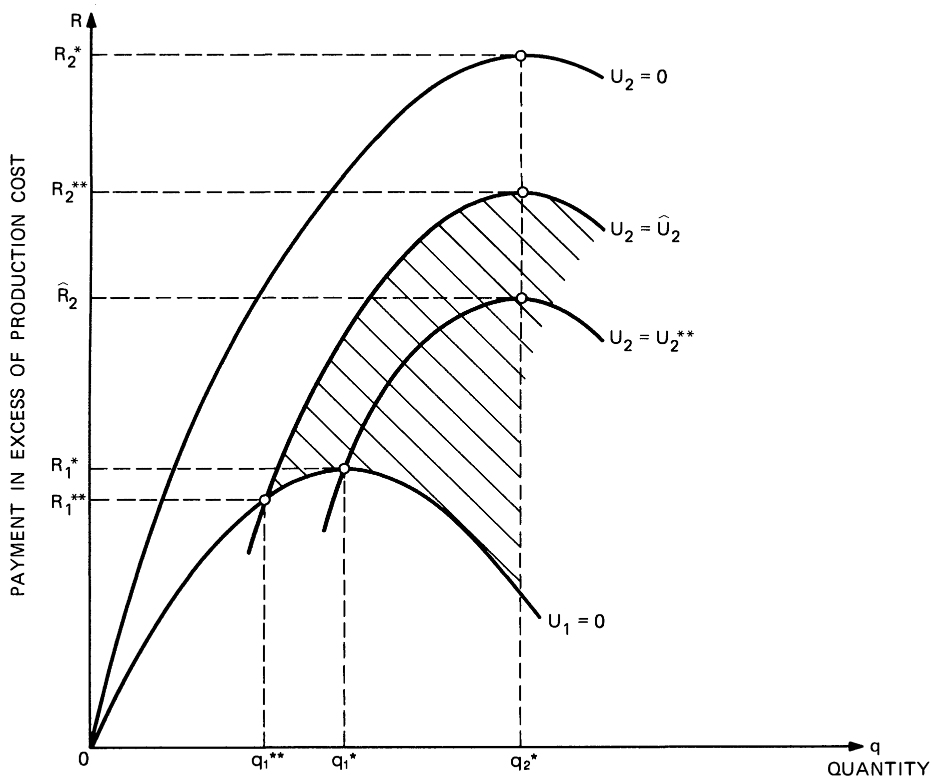
Therefore, at any point $\langle q, R \rangle$ the indifference curve for a buyer with a higher parameter value has a greater slope. Sorting is feasible precisely because different individuals have different marginal rates of substitution between the commodity and income. Our assumption that one individual's marginal rate of substitution is everywhere higher than another's is an important analytical simplification.

For the simplest case of two buyer types we can illustrate the profit-maximizing selling strategy with the help of Figure 1. If the seller had complete information about buyer types, he could extract all buyer surplus by introducing the schedule $I^* = \{\langle q_1^*, R_1^* \rangle, \langle q_2^*, R_2^* \rangle\}$. But since we assume that he has no direct means of distinguishing buyer types, this selling procedure will not extract all surplus. Indeed, high demanders ($v = v_2 > v_1$) are strictly better off buying q_1^* units at a total cost of $R_1^* + cq_1^*$.

From the figure it is easy to see, moreover, that the seller can do strictly better than I^*. Consider the indifference curve for a high demander through $\langle q_1^*, R_1^* \rangle$. Any such

FIGURE 1

PRICE DISCRIMINATION WITH TWO TYPES OF BUYER



buyer cannot be dissuaded from choosing $\langle q_1^*, R_1^* \rangle$ if available, unless also offered an alternative on or below this curve. Thus, assuming he also offers $\langle q_1^*, R_1^* \rangle$, we find that the seller maximizes his return by offering the alternative pair $\langle q_2^*, \hat{R}_2 \rangle$. Note that at the points chosen by each type, the associated indifference curves have zero slope. That is, the pairs $\langle q_1^*, R_1^* \rangle$ and $\langle q_2^*, \hat{R}_2 \rangle$ are efficient.

We next establish, however, that the monopolist can do better than $\hat{I} = \{\langle q_1^*, R_1^* \rangle; \langle q_2^*, \hat{R}_2 \rangle\}$ by introducing inefficiency. Consider the alternative $\langle q_1^{**}, R_1^{**} \rangle$, depicted in Figure 1, which also extracts all the surplus from low demanders. Much as before, the monopolist maximizes his return from high demanders, given that he offers $\langle q_1^{**}, R_1^{**} \rangle$, by also offering $\langle q_2^*, R_2^{**} \rangle$. By presenting buyers with $I^0 = \{\langle q_1^{**}, R_1^{**} \rangle, \langle q_2^*, R_2^{**} \rangle\}$ rather than $\hat{I}$, the monopolist gains relatively from high demanders ($R_2^{**} > \hat{R}_2$) and loses from low demanders ($R_1^{**} < R_1^*$). But observe that at least for small moves to the left of q_1^*, the slope of the indifference curve for type 1 buyers is approximately zero. Therefore, the fall in the return from type 1 buyers is approximately zero. To be precise, there is a first-order rise in the return from high demanders and only a second-order decline from low demanders.

As the ratio of high to low demanders increases, the offer $\langle q_1^{**}, R_1^{**} \rangle$ moves further to the left until eventually $\langle q_1^{**}, R_1^{**} \rangle = \langle 0, 0 \rangle$. Then, if the ratio of high to low demanders is sufficiently great, I^0 in effect becomes simply $\{\langle q_2^*, R_2^* \rangle\}$. Regardless of whether I^0 or $\{\langle q_2^*, R_2^* \rangle\}$ is optimal, however, only high demanders purchase the efficient quantity $q_2^* = q^e(v_2)$, and the low demanders' demand price for the last unit purchased exceeds marginal cost.

A straightforward generalization of this argument establishes that, with more than two types of buyer, demand price will exceed marginal cost for all except the highest demanders. To summarize, we have:

Proposition 1: Inefficiency of monopoly price discrimination.[6] Under Assumption 1, if $V = \{v_1, \ldots, v_n | v_i < v_{i+1}\}$ is the set of possible parameter values, then the profit-maximizing selling procedure of the form $\langle q_s^*, R(q_s^*) \rangle_{s \in S}$ will result in purchases $q(v_i)$ by a type i buyer such that $p(q(v_i); v_i) \geqslant c$ for all i, and the inequality is strict for $i < n$.

This proposition is an instance of a very general principle that has emerged from the rapidly growing literature on incentive schemes (e.g., income taxation and piece-rate schemes); *viz.*, the optimal incentive schedule typically involves distortions (that is, deviations from the first-best) at all but one point.

The primary goal of this article is to show that by introducing only mild restrictions, we can provide a *complete* characterization of the expected profit-maximizing nonlinear price schedule for any smooth distribution of demand curves. In so doing we lay a broad foundation for future empirical investigations of monopolistic pricing practices.

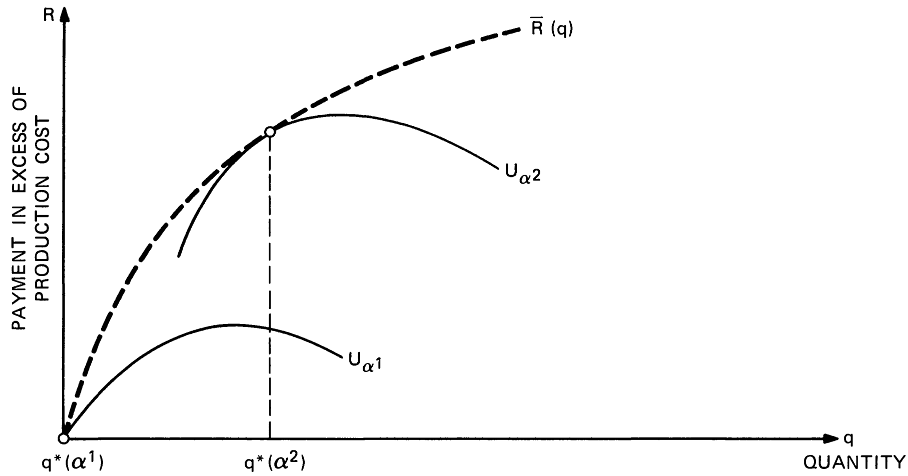
To preview the issues involved and the results to follow, consider a selling procedure $\langle q(s), R(s) \rangle = \langle q(s), \bar{R}(q(s)) \rangle$, where $\{q(s)\}_{s \in S}$ consists of all nonnegative quantities, and the return function $\bar{R}(q)$ is continuously differentiable, as depicted in Figure 2a. Because indifference curves are steeper for larger v, the response $q^*(v)$ (that is, the purchase by a type v buyer) will be a strictly increasing function for all v above some minimum α^1 where purchases are zero. Thus, the buyer types who actually make purchases are completely "sorted out."

Figure 2a is somewhat misleading, however, since it is not generally optimal for the seller to sort completely. One alternative possibility is illustrated in Figure 2b. Instead of being continuously differentiable, the return function, $\bar{R}^*(q)$, has an upward kink at A. As depicted, any buyer with $v > \alpha^3$ chooses $q^*(v) > q_A$. Those with $v \in [\alpha^2, \alpha^3]$ all choose

[6] The graphical derivation of this proposition is easily formalized.

FIGURE 2a

COMPLETE SORTING OF ALL BUYERS MAKING PURCHASES



$q^*(v) = q_A$. By using the function $\bar{R}^*(q)$ rather than $\bar{R}(q)$, a seller obtains a lower return from those buyers for whom $v \in [\alpha^2, \alpha^3]$. But the kinked schedule generates a greater return from high values of v. It is therefore intuitively plausible that if the proportion of the buyer population with intermediate values of v is sufficiently low relative to the proportions with high or low v's, seller profit is increased by the introduction of such a kink. As we shall see, kinked nonlinear pricing schemes are optimal under weaker conditions as well.

A third possibility is depicted in Figure 2c. Instead of the continuously differentiable $\bar{R}(q)$, the seller constructs $\bar{R}^{**}(q)$ so that a buyer of type α^2 is indifferent between the points A and A'. The resulting $\bar{R}^{**}(q)$ lies below $\bar{R}(q)$, and hence relative to $\bar{R}(q)$, the seller has a lower return if v is sufficiently high. But he now gains from buyers with parameter

FIGURE 2b

KINKED PRICE FUNCTION

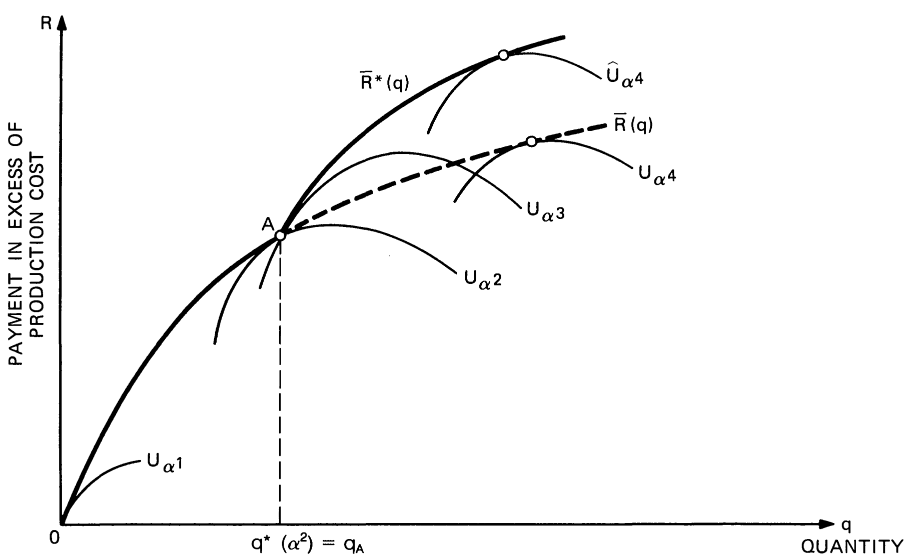
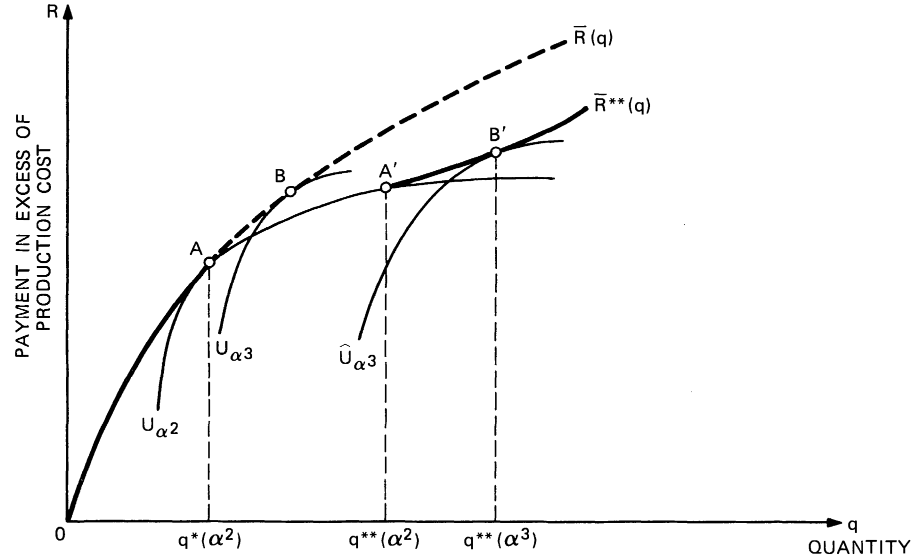


FIGURE 2c

DISCONTINUOUS PRICE FUNCTION



value α^3 somewhat greater than α^2, since such buyers select points like B' rather than B. This suggests that if the fraction of buyers with a parameter value near α^3 is sufficiently high, discontinuities will be optimal. Of course, the argument is incomplete, since it ignores the shape of the optimal return function below point A. Indeed we shall see that, given our other assumptions, discontinuities in the return function are optimal if and only if the density function for v has an upward discontinuity.

3. Self-selection constraints

■ Rather than proceeding directly to the general model with an arbitrary continuous cumulative distribution function $F(v)$, we first examine the "self-selection" constraints for the discrete case. These constraints require that each consumer type be as happy with the $\langle q, R\rangle$ pair assigned to him as with that of any other type. The crucial step is to show that the optimization problem faced by the seller—choice of a schedule $\langle q_s, R(q_s)\rangle_{s\in S}$, subject to self-selection constraints—reduces to a tractable variational problem. This is done by first showing that for each type, only *one* of the constraints is actually binding; *viz.*, the "local downward" constraint.

Consider Figure 1 once again, but now suppose that there are n different types with $v_i < v_{i+1}$, $i = 1, \ldots, n$. First we note that, because the indifference curves of an individual with a higher v are everywhere steeper, the following syllogism must hold:

$$\text{For all} \quad q_2 \geqslant q_1 \quad \text{and} \quad \langle q_1, R_1\rangle \neq \langle q_2, R_2\rangle, \tag{7}$$

$$\langle q_1, R_1\rangle \underset{v_i}{\precsim} \langle q_2, R_2\rangle \Rightarrow \langle q_1, R_1\rangle \underset{v}{\prec} \langle q_2, R_2\rangle, \qquad \text{for} \qquad v > v_i.$$

We next define $\langle q(v_i), R(v_i)\rangle_{v_i\in V}$ to be the schedule of pairs optimal for each buyer type, given some schedule of offers by the seller. Suppose that the pair $\langle q_1^{**}, R_1^{**}\rangle$ depicted in Figure 1 is optimal for type 1. Then all other offers $\langle q_s, R_s\rangle$ must lie on or above the indifference curve for type 1 through $\langle q_1^{**}, R_1^{**}\rangle$. Given Assumption 1, we know that the indifference curve for type 2 through this same point is steeper at any level of q. Then either $q(v_2) = q(v_1)$ or a buyer of type 2 chooses a pair $\langle q_s, R_s\rangle$ on or below this steeper indifference curve. But we have already argued that all the alternatives must lie above the flatter indifference curve for type 1. Then either $q(v_2) = q(v_1)$ or $\langle q(v_2), R(v_2)\rangle$ must lie in the shaded region, that is,

$$q(v_2) \geqslant q(v_1).$$

Since exactly the same logic applies for any two neighboring types, we can conclude that the response function

$$q(v) \text{ is a nondecreasing function.} \tag{8}$$

A further important implication of Assumption 1 is that, for each type, the "local downward" constraint is strictly binding. That is, there is no type v_i such that

$$\langle q(v_i), R(v_i)\rangle \underset{v_i}{\succ} \langle q(v_{i-1}), R(v_{i-1})\rangle. \tag{9}$$

To confirm this, suppose (9) holds for some i. Since $\langle q(v_{i-1}), R(v_{i-1})\rangle$ is optimal for type v_{i-1}, we also have

$$\langle q(v_{i-1}), R(v_{i-1})\rangle \underset{v_{i-1}}{\succsim} \langle q(v_k), R(v_k)\rangle, \qquad k = 1, \ldots, i-1.$$

Then appealing to (7) and (8)

$$\langle q(v_{i-1}), R(v_{i-1})\rangle \underset{v_i}{\succ} \langle q(v_k), R(v_k)\rangle, \qquad k = 1, \ldots, i-1. \tag{10}$$

Now consider the alternative scheme

$$\hat{q}_k = q(v_k)$$

$$\hat{R}_k = \begin{cases} \tilde{R}(y_k), & k < \mathrm{i} \\ R(v_k) + \delta, & k \geqslant \mathrm{i}. \end{cases}$$

For sufficiently small δ, (9) continues to hold. Then (10) holds also. That is, type v_i will not switch to a lower level of q. An almost identical argument establishes that no type v_j, $j > i$, will do so either. Finally we note that, since indifference curves are vertically parallel, rankings over $\langle \hat{q}(v_j), \hat{R}(v_j)\rangle_{j\geqslant i}$ are identical to the rankings over $\langle q(v_j), R(v_j)\rangle$ for all types v_j, $j \geqslant i$.

Thus, the new scheme results in the same choices of q by each type and a strictly greater return to the monopolist. But this is impossible since, by assumption, $\langle q(v_i), R(v_i)\rangle_{i=1,\ldots,n}$ is the monopolist's optimal scheme. Therefore, (9) is false, and the local downward constraint is binding; that is,

$$\langle q(v_i), R(v_i)\rangle \underset{v_i}{\sim} \langle q(v_{i-1}), R(v_{i-1})\rangle, \qquad i = 2, \ldots, n. \tag{11}$$

This argument implies that a buyer with the lowest v obtains no surplus, that is,

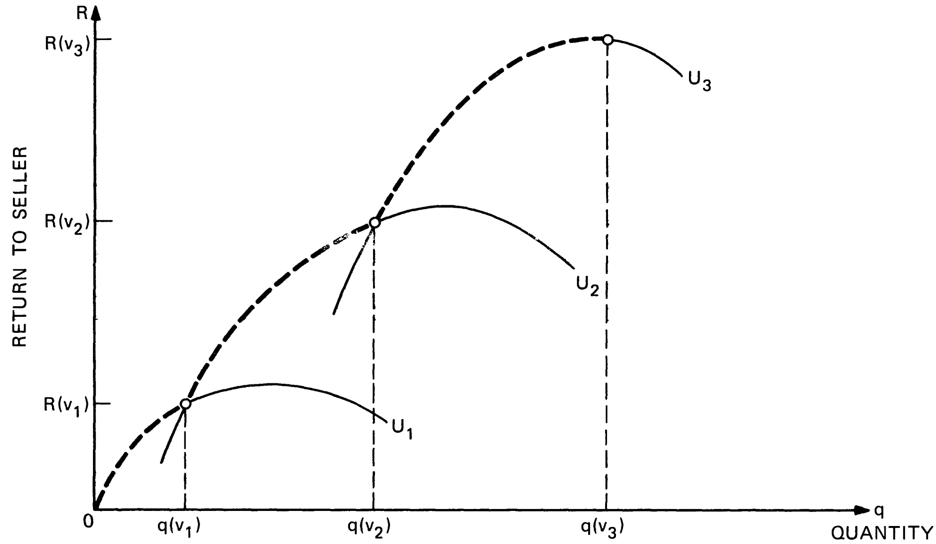
$$\langle q(v_1), R(v_1)\rangle \underset{v_1}{\sim} \langle 0, 0\rangle. \tag{12}$$

Figure 3 depicts a possible optimal selling procedure with three different types. All the pairs $\langle q(v_i), R(v_i)\rangle$ lie along the dashed curve, where each segment corresponds to an indifference curve for a different demander. From this figure, it should be clear that the indifference curve for each type v_i through $\langle q(v_i), R(v_i)\rangle$ never lies above the dashed curve. Thus, if conditions (8) and (11) are satisfied, $\langle q(v_i), R(v_i)\rangle$ is indeed the global optimum for type v_i. That is, we assert that besides being necessary properties of a monopolist's optimum, (8) and (11) are also conditions sufficient for *all* self-selection constraints to be satisfied.

From this analysis of the finite case we can now move easily to the limit with the buyer's demands continuously distributed. In what follows we shall make use of the next assumption.

FIGURE 3

COMPLETE SORTING WITH THREE DIFFERENT BUYER TYPES



Assumption 2. The cumulative distribution function for v, $F(v)$, is a strictly increasing, continuously differentiable function on the interval $[0, \bar{v}]$ with $F(0) = 1 - F(\bar{v}) = 0$.

With $\langle q(v_i), R(v_i)\rangle$ optimal for a buyer with parameter v_i, we can write maximized utility as

$$U^*(v_i) \equiv N(q(v_i); v_i) - R(v_i). \tag{13}$$

From (11)

$$U^*(v_{i+1}) \equiv N(q(v_{i+1}); v_{i+1}) - R(v_{i+1}) = N(q(v_i); v_{i+1}) - R(v_i).$$

Therefore,

$$\begin{aligned} U^*(v_{i+1}) - U^*(v_i) &= N(q(v_i); v_{i+1}) - N(q(v_i); v_i) \\ &= \int_{v_i}^{v_{i+1}} \frac{\partial N}{\partial v}(q(v_i); x)dx. \end{aligned} \tag{14}$$

From (12), we know that the utility of the lowest demanders must equal the utility of not participating, that is,

$$U^*(0) = 0.^7 \tag{15}$$

Combining (14) and (15), we obtain

$$U^*(v_i) = \sum_{j=0}^{i-1} \int_{v_j}^{v_{j+1}} \frac{\partial N}{\partial v}(q(v_j); x)dx, \qquad \text{where} \qquad v_0 = 0.$$

Substituting this expression in (13), we can write the expected seller revenue from a buyer of type v_i in terms of $q(\cdot)$:

$$R(v_i) = N(q(v_i); v_i) - \sum_{j=0}^{i-1} \int_{v_j}^{v_{j+1}} \frac{\partial N}{\partial v}(q(v_j); x)dx.$$

[7] Oren, Smith, and Wilson (1984) consider a case in which there is an additional fixed cost, k, of supplying each customer. Then (15) becomes $U^*(0) = k$. Their analysis is easily generalized by using our methods.

Taking the limiting case of a continuous distribution of types, we have

$$R(v) = N(q(v); v) - \int_0^v \frac{\partial N}{\partial v}(q(x); x)dx. \tag{16}$$

From the monopolist's viewpoint, v and, hence, $R(v)$ are random variables. He, therefore, chooses $q(v)$ to maximize the expectation of $R(v)$ subject to the constraint (8) that $q(v)$ is nondecreasing. The expectation of $R(v)$ is

$$\int_0^{\bar{v}} [N(q(v); v) - \int_0^v \frac{\partial N}{\partial v}(q(x); x)dx]dF(v),$$

or, after integration by parts,

$$\int_0^{\bar{v}} [N(q(v); v) - \frac{\partial N}{\partial v}(q(v); v)/\rho(v)]dF(v),$$

where $\rho(v) \equiv F'(v)/(1 - F(v))$ is the hazard rate of F. Summarizing, we can state the following proposition.

Proposition 2: Expected profit from second-degree price discrimination. If Assumptions 1 and 2 are satisfied, the expected profit obtainable from a selling procedure $\langle q(v), R(v)\rangle$, where $q(v)$ is nondecreasing, can be expressed as

$$\int_0^{\bar{v}} I(q(v); v)dF(v),$$

where

$$I(q; v) = N(q; v) - \frac{\partial N}{\partial v}(q; v)/\rho(v), \tag{17}$$

$$N(q, v) \equiv \int_0^q p(x, v)dx - cq \text{ is the net social benefit,}$$

and

$$\rho(v) \equiv F'(v)/(1 - F(v)).^8$$

To maximize expected revenue, the seller thus chooses $q^*(v)$ to solve

$$\max_{q(v)} \left\{\int_0^{\bar{v}} I(q(v); v)dF(v) | q(v) \text{ is nondecreasing}\right\}. \tag{18}$$

Assume that $q^*(\cdot)$ is piecewise differentiable. (We shall see below that this assumption is justified.) Then, from (16), the return to the seller from a buyer of type v is

$$\begin{aligned} R^*(v) &= N(q^*(v); v) - \int_0^v \frac{\partial N}{\partial v}(q^*(x); x)dx \\ &= \int_0^v \frac{\partial N}{\partial q}(q^*(x); x)\frac{dq^*(x)}{dx}dx. \end{aligned}$$

Let $\phi(q) = \min\{v | q^*(v) = q\}$. Because $q^*(v)$ is nondecreasing, we can write $v = \phi(q)$. We can, therefore, write

[8] In deriving the formula for expected profit in the case of only finitely many values of v, we invoked the assumption that $\langle q(v), R(v)\rangle$ is optimal for the monopolist. In particular, we used this assumption to show that the local downward incentive constraint is binding. One can show, however, (see the lemma following Proposition 3) that formula (17) is valid even for nonoptimal selling schemes.

$$R^*(\phi(q)) = \int_0^q \frac{\partial N}{\partial q}(z; \phi(z))dz$$
$$= \int_0^q p(z; \phi(z))dz - cq.$$

Hence we have the following proposition.

Proposition 3: Optimal nonlinear pricing. Under Assumptions 1 and 2 expected monopoly profit is maximized by introducing the nonlinear total payment schedule

$$\bar{T}(q) = \int_0^q p(z; \phi(z))dz, \tag{19}$$

where $\phi(z)$ is defined by $\phi(z) = \min\{v|q^*(v) = z\}$ and $q^*(v)$ is the solution to (18).

We have derived Propositions 1–3 somewhat informally to emphasize the ideas behind the proofs. In fact, we have left certain gaps in the analysis. First of all, our derivation of (16) by taking limits was rather loose. In Maskin and Riley (1983b) we provide a formal proof. In fact, we establish the following lemma.

Lemma. For any $q(v)$, a nondecreasing function on $[0, \bar{v}]$, there exists a unique return function, $R(\cdot)$, given by (16), such that $\langle q(\cdot), R(\cdot)\rangle$ satisfies all the self-selection constraints and (15).

We should note that the uniqueness of $R(\cdot)$ in this lemma depends crucially on the continuity of $F(\cdot)$. If $F(\cdot)$ is, by contrast, a discrete distribution, then corresponding to each $q(\cdot)$ there will be a continuum of $R(\cdot)$'s such that the self-selection constraints and (15) are satisfied (although only one that is profit-maximizing).

Another point we ignored is the possibility that the seller might offer a schedule $\langle \tilde{q}_s, \tilde{R}_s\rangle_{s\in S}$, where $\tilde{q}_s$ and $\tilde{R}_s$ are random rather than deterministic functions. We consider this possibility in Maskin and Riley (1983a), wherein it is established that gains from randomization are ruled out if the following assumption is satisfied.

Assumption 3: Nondecreasing price elasticity. Demand elasticity is nondecreasing in the demand price. That is,

$$\frac{\partial}{\partial v}\left(\frac{-q}{p}\frac{\partial p}{\partial q}\right) \leq 0. \tag{20}$$

For a buyer of type v, utility is

$$U(q, T; v) = \int_0^q p(x, v)dx - T.$$

Hence, for such a buyer,

$$\frac{\partial}{\partial v}\left(\frac{-\dfrac{\partial^2 U}{\partial q^2}}{\dfrac{\partial U}{\partial q}}\right) = \frac{1}{q}\frac{\partial}{\partial v}\left(\frac{-q}{p}\frac{\partial p}{\partial q}\right). \tag{21}$$

Assumption 3 implies, therefore, that absolute risk aversion is nonincreasing in v, if we interpret U as a von Neumann-Morgenstern utility function.

We should note that we have a great deal of flexibility in our choice of a parameterization. In particular, if $p(q; v)$ represents a family of inverse demand curves satisfying Assumptions 1, 2, and 3, then $p(q, \omega(v))$ represents the same family and also satisfies these three assumptions if $\omega(\cdot)$ is strictly increasing. For convenience, we shall choose, without loss of generality, a parameterization for which the increases in demand price are diminishing as v rises.

Assumption 4. $p_{22}(q; v) \leqslant 0$.

Since we shall appeal to Assumptions 1, 2, 3, and 4 in much of what follows, we observe that they are satisfied for large classes of preferences. For example, they are satisfied if

$$p(q; v) = \alpha(v)a(q) - \beta(v)h(q), \tag{22}$$

where $a(q)$ is positive and nonincreasing, $h(q)$, $\alpha(v)$, and $\alpha(v)/\beta(v)$ are all nonnegative and nondecreasing, $\alpha''(v) \leqslant 0$, $\beta''(v) \geqslant 0$, $\alpha(v)a(0) - \beta(v)h(0) \geqslant 0$, and $\alpha(v)a'(0) - \beta(v)h'(0) < 0$.

One special case of (22) is the family of linear demand curves. Without loss of generality, we can set $\alpha(v) = v$, $a(q) = 1$, and $h(q) = q$ so that

$$p(q; v) = v - \beta(v)q.$$

A second important case is the family of constant elasticity demand functions $p(q; v) = \alpha(v)q^{-1/\eta}$. As we shall see, the seller's optimization problem is especially straightforward for this family.

4. Characterization of the optimal selling procedure

■ We now turn to the solution of the seller's optimization problem, (18). First, we provide conditions under which there is complete sorting of all those buyers actually making purchases.

Proposition 4: Complete sorting optimum. Suppose Assumptions 1, 2, 3, and 4 hold. Then the $\bar{q}(v)$ that solves $\max_q I(q; v)$ is nondecreasing and, whenever $\bar{q}$ is positive, strictly increasing if either

$$J(v) \equiv v - \frac{1 - F(v)}{F'(v)} \text{ is increasing} \tag{23}$$

or

$$\frac{1}{p_2(\bar{q}(v); v)} \frac{\partial}{\partial v}\left(\frac{p_2(\bar{q}(v); v)}{\rho(v)}\right) < 1. \tag{24}$$

Proof. We first establish that Assumptions 1, 2, and 3 together imply that $I(q; v)$ is a strictly quasi-concave function of q. From (17)

$$\frac{\partial I}{\partial q} = p(q; v) - p_2(q; v)/\rho(v) - c.$$

Hence

$$\frac{\partial I}{\partial q} \geqslant 0 \Rightarrow 1/\rho < p/p_2. \tag{25}$$

Also

$$\frac{\partial^2 I}{\partial q^2} = p_1 - p_{12}/\rho. \tag{26}$$

If p_{12} is nonnegative, $\partial^2 I/\partial q^2$ is negative since, by Assumption 1, $p_1 < 0$ for $q > 0$. If p_{12} is negative, then by (25)

$$\frac{\partial I}{\partial q} \geqslant 0 \Rightarrow \frac{\partial^2 I}{\partial q^2} < p_1 - \frac{p_{12}p}{p_2} = \frac{p^2}{qp_2}\frac{\partial}{\partial v}\left(\frac{-qp_1}{p}\right).$$

By Assumption 3 the final term is nonpositive. Thus $I(q, v)$ is indeed strictly quasi-concave.

Furthermore,

$$\frac{\partial^2 I}{\partial q \partial v} = p_2\left(1 - \frac{1}{p_2}\frac{\partial}{\partial v}\left(\frac{p_2}{\rho}\right)\right) = \frac{p_2}{\rho}\left(\rho + \frac{\rho'}{\rho} - \frac{p_{22}}{p_2}\right). \tag{27}$$

Since $\bar{q}(v)$ maximizes $I(q, v)$, either $\bar{q}(v) = 0$ or $\bar{q}(v)$ satisfies the first-order condition $\partial I/\partial q = 0$. Totally differentiating the latter equation, we obtain

$$\frac{d\bar{q}}{dv} = -\frac{\partial^2 I}{\partial q \partial v}\bigg/\frac{\partial^2 I}{\partial q^2}.$$

We have just argued that the denominator is negative, hence the derivative exists. Also, if (24) holds, then, from (27), $\partial^2 I/\partial q\partial v$ is positive. Alternatively, since

$$J'(v) = \frac{1}{\rho}\left(\rho + \frac{\rho'}{\rho}\right),$$

and by hypothesis $p_{22} \leqslant 0$, it follows from (27) that if (23) is satisfied, $\partial^2 I/\partial q\partial v$ is again positive.

Hence $\bar{q}(v)$ is either zero or strictly increasing. Thus $\bar{q}(v)$ solves the monopolist's optimization problem (18). *Q.E.D.*

Although condition (24) has the advantage of being invariant to parameterization, the somewhat stronger condition (23) is easier to interpret. It requires that the hazard rate $\rho(v)$ not decline too rapidly with v, a requirement that accords well with our heuristic discussion of complete sorting in Section 2. Certainly, this condition is satisfied for a large subclass of distribution functions. It is not difficult, however, to produce instances in which it is violated. For example, if, holding the support of v fixed, the variance of v is made sufficiently small, $J(v)$ is nonmonotonic.[9] We shall see below that whenever there exists a v such that $J(v)$ is positive and decreasing, complete sorting is no longer optimal.

Example 1: Vertically parallel demand:

$$p(q; v) = v - h(q),\ h'(\cdot) > 0.$$

For this simple case

$$\frac{\partial I}{\partial q}(q; v) = J(v) - h(q) - c,$$

where

$$J(v) = v - \frac{1 - F(v)}{F'(v)}.$$

Then if $J(v)$ is increasing,

$$\bar{q}(v) = h^{-1}(\max\{0, J(v) - c\}) \text{ is nondecreasing, and}$$

by Proposition 3 the optimal nonlinear price function is

$$\bar{T}(q) = \int_0^q [J^{-1}(h(z) + c) - h(z)]dz.$$

Example 2: Constant price elasticity of demand:

$$p(q; v) = vq^{-1/\eta}, \qquad \eta > 1$$

$$\frac{\partial I}{\partial q}(q; v) \equiv J(v)q^{-1/\eta} - c.$$

---

[9] It is easy to give other examples for which $J(v)$ is not monotonic. For instance, if

$$F(v) = v(7 - 9v + 4v^2)/2$$

so that $\bar{v} = 1$, it may be confirmed that $J(1/2) = J(5/8) = J(3/4) > 0$.

Then, once again, $\bar{q}(v)$ is monotonic if $J(v)$ is increasing. Substituting into (19), we obtain

$$\bar{T}(q) = \int_0^q J^{-1}(cz^{1/\eta})z^{-1/\eta}dz.$$

Under the hypothesis of Proposition 4, it is in principle possible to compute the optimal nonlinear pricing scheme for any family of demand curves. One important qualitative issue is how a change in the underlying distribution affects the optimal price function. Suppose that there is a rightward shift in the distribution, so that the new distribution $G(v)$ is strictly less than $F(v)$ for all $v < (0, \bar{v})$. Since higher values of v are associated with higher demanders, keeping the same pricing scheme will make the seller better off. Less obvious, however, is the effect on the optimal payment function $\bar{T}(q)$.

One condition sufficient to ensure that $G(v) < F(v)$ is that the hazard rate for F exceeds the hazard rate for G. We now show that this condition ensures that the optimal payment function $\bar{T}_G(q)$ is strictly steeper for all q than the corresponding function for F.

Proposition 5. If for all $v \in (0, \bar{v})$ the distribution $G(v)$ has a lower hazard rate than $F(v)$ and Assumptions 1, 2, and 3 are satisfied, then

$$\frac{d}{dq}\bar{T}_G(q) > \frac{d}{dq}\bar{T}_F(q).$$

Proof. We consider only the case in which the optimal response function, $\bar{q}(v)$, is the solution to $\max_q I(q; v)$. That the proof generalizes follows directly from the Characterization Theorem (Proposition 7).

From the proof of Proposition 4, $I(q; v)$ is strictly quasi-concave. Also, from (17), if $\rho_F > \rho_G$, then

$$\frac{\partial I_F}{\partial q} > \frac{\partial I_G}{\partial q}.$$

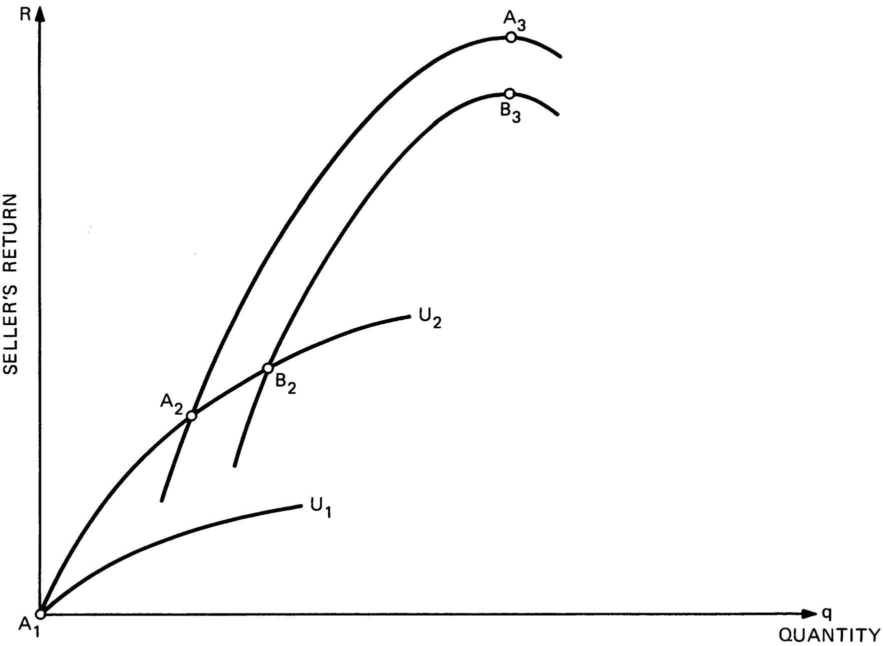
But if $\bar{q}_G(v) > 0$, $\frac{\partial I_G}{\partial q}(\bar{q}_G(v); v) = 0$. Thus, $\bar{q}_G(v) > 0$ implies that $\bar{q}_F(v) > \bar{q}_G(v)$, so that the two inverse functions satisfy $\phi_F(q) < \phi_G(q)$. Since $p_2 > 0$, the result then follows directly from (19), the definition of $\bar{T}(q)$. *Q.E.D.*

Intuitively, by making the payment function steeper, the monopolist is able to extract more revenue from the highest demanders. Thus, as the number of high relative to low demanders rises, a steeper price function becomes more desirable. This intuition, however, is incomplete, since it suggests that any rightward shift in the distribution will raise the entire payment schedule. A counterexample, in which there are just three types, is illustrated in Figure 4. Suppose that, for the initial probability weights, the optimal selling procedure is to offer the three pairs, A_1, A_2, A_3. Next suppose that the distribution is shifted to the right via a reduction in the probability weight for the lowest demanders and an increase in the weight on type 2, with intermediate demand. This raises the marginal benefit to increasing the payment by type 2. The new optimum is then A_1, B_2, B_3 with a lower profit from the type 3 demanders. Since the highest demanders are allocated the same output, $\bar{T}(q(v_3)) = \bar{R}(q(v_3)) + cq(v_3)$ is also lower.

One feature common to nonlinear pricing schemes in practice is the use of quantity discounts. We often see offers such as "one for a dollar, three for two dollars," and the more complicated multipart tariffs of utility companies almost invariably involve offers of lower unit prices for greater consumption. A reason why quantity premiums ("one for a dollar, two for three dollars") may not be used is that they may be difficult to enforce; a customer might be able to circumvent a premium by successively buying small quantities

FIGURE 4

IMPACT OF A SHIFT TO A MORE FAVORABLE DISTRIBUTION



under different guises.[10] There are many commodities, however, (e.g., utilities) for which enforcement is possible. We, therefore, consider whether or not premiums are desirable in the absence of enforcement costs. For a broad class of cases, we shall see, they are not.

The payment per unit purchased is, under the hypotheses of Proposition 3,

$$\frac{T(q(v))}{q(v)} = \int_0^v \frac{p(q(x); x)q'(x)dx}{q(v)}.$$

This is decreasing in v, and hence in q, if and only if

$$q(v)p(q(v); v) - \int_0^v p(q(x); x)q'(x)dx = \int_0^v [p(q(v); v) - p(q(x); x)]q'(x)dx < 0. \quad (28)$$

From Proposition 1, for all $x < \bar{v}$

$$p(q(x); x) > c = p(q(\bar{v}), \bar{v}). \quad (29)$$

Thus, quantity discounts are always optimal for buyers at the upper tail of the distribution. Invoking some further fairly mild restrictions, we can establish the considerably stronger result that quantity discounts are everywhere optimal.

Proposition 6: Quantity discounts. Suppose that Assumptions 1, 2, and 3 hold. Then the optimal unit price $T(q)/q$ is everywhere declining if

[10] Of course, quantity discounts can in principle be circumvented as well. Consumers can band together to buy in bulk or an entrepreneur can do the bulk buying for them. But such evasive actions, unlike those for quantity premiums, require the coordination of more than one agent.

$$\frac{\partial}{\partial v}\left(\frac{p_2(q; v)}{p_1(q; v)\rho(v)}\right) > 0 \tag{30}$$

for $q = q^*(v)$, where $q^*(v)$ is the optimal response function.

Proof. We consider only the case where the optimal response function $q^*(v)$ solves max $I(q, v)$. That the proposition holds in general, however, follows directly from the Characterization Theorem (Proposition 7).

From (28), quantity discounting is optimal for all q if $p^*(q(v); v)$ is a decreasing function of v for all $q^*(v) > 0$. From the proof of Proposition 4, $q^*(v) > 0$ implies that

$$\frac{dq^*}{dv} = \frac{-\dfrac{\partial^2 I}{\partial q \partial v}}{\dfrac{\partial^2 I}{\partial q^2}} = \frac{-p_2 + \dfrac{\partial}{\partial v}\left(\dfrac{p_2}{\rho}\right)}{p_1 - \dfrac{p_{12}}{\rho}}. \tag{31}$$

Furthermore, the denominator $\dfrac{\partial^2 I}{\partial q^2}$ is negative. Differentiating $p(q(v); v)$ and substituting from (31), we obtain

$$\begin{aligned}\frac{d}{dv}p(q^*(v); v) &= p_1\frac{dq}{dv} + p_2\\ &= p_1^2\frac{\partial}{\partial v}\left(\frac{p_2}{p_1\rho}\right)\Bigg/\frac{\partial^2 I}{\partial q^2},\end{aligned}$$

which, given (30), is negative. *Q.E.D.*

Sufficient conditions for condition (30) to be satisfied are that Assumption 3 hold ($p_{22} \leqslant 0$), that the hazard rate $\rho(v)$ be nondecreasing, and that $p_{12} \leqslant 0$.[11]

Finally, we turn to the general characterization of the optimal selling procedure when $\bar{q}(v)$ is not monotonic. Proposition 7 reveals that the optimal response function $q^*(v)$ is continuous. There are subintervals over which $q^*(v)$ is constant, whereas on all other subintervals $q^*(v) = \bar{q}(v)$. As depicted in Figure 2b, a transition from the former to the latter corresponds to an upward kink in the payment schedule, $\bar{T}(q) = cq + \bar{R}(q)$. Given Assumptions 1, 2, and 3, Proposition 7 also shows that the intuitively plausible alternative depicted in Figure 2c is never optimal.[12]

In the proof of Proposition 7, which appears in the Appendix, we require the following two restrictions.

Assumption 5. $I(q; v)$ is strictly quasi-concave.

Assumption 6.

$$\int_v^y \frac{\partial I}{\partial q}(q; z)dF(z) \geqslant 0 \rightarrow \int_v^y \frac{\partial^2 I}{\partial q^2}(q; z)dF(z) < 0, \qquad v < y.$$

Both assumptions are fairly weak. For example, Assumption 6 holds if

$$p(q; v) = \phi(v)a(q) - h(q),$$

and we have already seen that Assumption 3 is a sufficient condition for Assumption 5.

[11] It is easy to see that $p_{12} \leqslant 0$ is not necessary to establish the desirability of quantity discounts globally. For example, if $p(q; v) = 1 - \dfrac{q}{1+v}$ and $F(v) = v$, then (30) is satisfied while $p_{12} > 0$.

[12] Proposition 7 continues to hold if $F(v)$ is a strictly increasing continuous piecewise differentiable function. It then follows immediately that if the density rises discontinuously at some v', $q^*(v)$ will have an upward discontinuity at v'.

Proposition 7: General characterization of the optimal response function.

(A) Suppose Assumptions 1, 2, and 5 are satisfied. Assume that there exists an optimal response function, $q^*(v)$. Then there exists a set of subintervals

$$\{[x^i, y^i] \subseteq [0, \bar{v}] | x^{i-1} > y^i\}_{i \in I},$$

possibly empty, such that for all i

$$\text{(i)} \qquad \bar{q}(x^i) = \bar{q}(y^i),$$

and

$$\text{(ii)} \qquad \int_v^y \frac{\partial I}{\partial q}(\bar{q}(y); z)dF(z) \leqslant 0, \qquad v < y, \qquad y \notin \bigcup_{i \in I}(x^i, y^i)$$

with equality if $y = y^i$ and $v = x^i$ $(x^i > 0)$. Moreover,

$$\text{(iii)} \qquad q^*(v) = \begin{cases} \bar{q}(y^i) & \text{if} \quad v \in [x^i, y^i] \text{ for some } i \\ \bar{q}(v) & \text{otherwise.} \end{cases}$$

(B) Given Assumptions 1, 2, and 5, a collection of subintervals satisfying (i) and (ii) exists. If, in addition, Assumption 6 holds, then $q^*(v)$ defined by (iii) is an optimal response function.

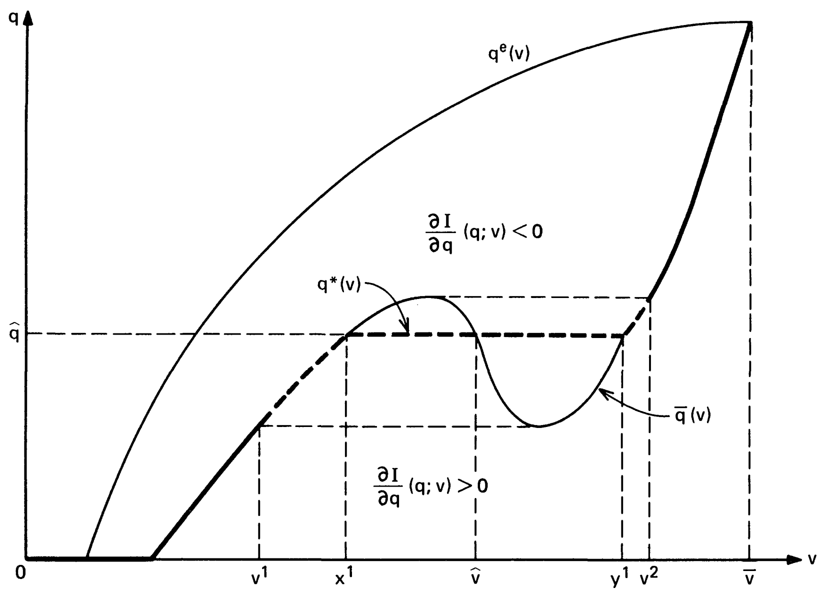
Proof. See the Appendix.

Notice that in part (A) of Proposition 7, the existence of an optimal response function is *assumed*, whereas in part (B), existence is *asserted*. To understand this result consider Figure 5, which illustrates a simple case where $\bar{q}(v)$ is decreasing over one subinterval. Since $\bar{q}(v)$ solves $\max_q I(q; v)$ and since $I(q; v)$ is, by assumption, strictly quasi-concave, we have

$$\frac{\partial I}{\partial q}(q; v) \begin{cases} >0, & q < \bar{q}(v) \\ <0, & q > \bar{q}(v). \end{cases}$$

FIGURE 5

SOLVING FOR THE EXPECTED PROFIT—MAXIMIZING RESPONSE FUNCTION



The seller seeks to solve

$$\max_{q(v)}\left\{\int_0^{\bar{v}} I(q;\, v)dF(v)|q(v) \text{ nondecreasing}\right\}.$$

To derive the optimal response function, $q^*(v)$, we begin by defining v^1 as depicted, so that

$$v^1 = \min\{v|\exists\hat{v} > v \text{ with } \bar{q}(\hat{v}) \leqslant \bar{q}(v)\}.$$

We next establish that $q^*(v) = \bar{q}(v)$ for $v \leqslant v^1$.

Since $\partial I(q;\, v)/\partial q$ is positive for $q^*(v) < \bar{q}(v)$ and negative for $q^*(v) > \bar{q}(v)$, it follows immediately that

$$\int_0^{v^1} I(\bar{q}(v);\, v)dF(v) \geqslant \int_0^{v^1} I(q^*(v);\, v)dF(v).$$

Thus, $q^*(v^1) > \bar{q}(v^1)$ cannot be optimal, at least if q^* is continuous, since, without violating the constraint that q^* be nondecreasing, and without altering q^* for $v > v^1$, we can set $q^* = \bar{q}$ over $[0, v^1]$. If $q^*(v^1) < \bar{q}(v^1)$, define

$$\hat{q}(v) = \begin{cases}\bar{q}(v), & v \leqslant v^1\\ \max\{q^*(v),\, \bar{q}(v^1)\}, & v > v^1.\end{cases}$$

For $v > v^1$ and $q \in (q^*(v^1), \bar{q}(v^1))$, $\partial I/\partial q(q;\, v) > 0$. Then

$$\int_0^{\bar{v}} I(\hat{q}(v);\, v)dF(v) > \int_0^{\bar{v}} I(q^*(v);\, v)dF(v),$$

again contradicting the optimality of $q^*(v)$.

Similarly we can define v^2, as depicted, so that

$$\bar{q}(v^2) = \max\{\bar{q}(v)|v < v^2\}.$$

Arguing almost exactly as above, $q^*(v) = \bar{q}(v)$ over $[v^2, \bar{v}]$.

The next step is to note that with $q^*(v^i) = \bar{q}(v^i)$, $i = 1, 2$, there must be some $\hat{v}$ in the interval over which $\bar{q}(v)$ is decreasing where $q^*(v)$ and $\bar{q}(v)$ intersect. Define $\hat{q} = \bar{q}(\hat{v})$. For $v < \hat{v}$, $q^*(v) \leqslant \hat{q}$, since $q^*(v)$ is monotonically increasing. Then, since $\partial I/\partial q$ is positive for $q < \bar{q}(v)$, the integral is maximized by choosing

$$q^*(v) = \min\{\bar{q}(v), \hat{q}\}.$$

An almost identical argument establishes that $q^*(v) = \max\{\bar{q}(v), \hat{q}\}$ for $v > \hat{v}$. Thus, as depicted, $q^*(v)$ is constant over some subinterval $[x^1, y^1]$ and equal to $\bar{q}(v)$ otherwise, establishing (iii) in Proposition 7.

Notice that $\bar{q}(x^1) = \bar{q}(y^1) = q^*(v)$ for all $v \in (x^1, y^1)$, establishing (i). To determine the location of x^1 and y^1, consider the monopolist's total profit

$$\int_0^{x^1} I(\bar{q}(v);\, v)dF(v) + \int_{x^1}^{y^1(x^1)} I(\bar{q}(x^1);\, v)dF(v) + \int_{y^1(x^1)}^{\bar{v}} I(\bar{q}(v);\, v)dF(v),$$

where $y^1 = y^1(x^1)$ satisfies $\bar{q}(y^1(x^1)) = \bar{q}(x^1)$. Differentiating this expression with respect to x^1, we obtain the further necessary condition

$$\int_{x^1}^{y^1} \frac{\partial I}{\partial q}(\bar{q}(x^1);\, v)dF(v) = 0.$$

Given our assumptions, this uniquely defines the interval $[x^1, y^1]$.

5. Monopoly pricing of product quality and optimal bundling

■ Our analysis of optimal pricing tactics in Sections 2–4 can be readily adapted to determine the optimal selling strategy of a monopolist selling products of differing quality levels.

Indeed, in the simplest case we need only reinterpret the basic model. We assume that each consumer wishes to purchase only one quality level. Furthermore, following Mussa and Rosen (1978), we begin by assuming that consumers either do not buy the good or purchase just one unit (in this case, of course, the one quality level assumption is satisfied automatically).

Consider the Marshallian utility function

$$\hat{u}(y, q, z; v) = y + zB(q; v), \tag{32}$$

where y is spending on other goods, q is the quality level of the single unit purchased, v represents the strength of preference for quality, and z is a dichotomous variable equal to unity with purchase and zero otherwise. If a consumer with income level I pays T for a unit of quality level q, we can rewrite his indirect utility as

$$u(q, T, z; v) = z(B(q; v) - T) + I. \tag{33}$$

Assumption 1 then requires that the marginal utility of higher quality is, *ceteris paribus,* at least as high for consumers with higher levels of v.

With little loss of generality, we can define units of quality in such a way that the marginal cost of a unit of quality level q is cq.[13] Then the monopolist's problem is identical to the problem considered in Sections 2–4, except that q is now to be interpreted as quality rather than quantity.

The natural generalization of this problem is to incorporate the choice of both quality q and the number of units purchased, z. Then, instead of (33), we have the indirect utility function

$$u(q, T, z; v) = B(q, z; v) - T + I, \tag{34}$$

where T is the total payment by the consumer for z units of quality level q. Adopting only slightly the derivations of Propositions 2 and 3, we can state the following proposition.

Proposition 8: Optimal bundling. Let consumer surplus for an individual of type v, purchasing z units of quality level q at a cost of T, be

$$B(q, z; v) - T,$$

where v is an independent draw from the distribution $F(\cdot)$. Let the cost of producing each unit of quality level q be cp. Suppose, for simplicity, that unit cost is independent of quantity. If both $\partial B/\partial q$ and $\partial B/\partial z$ are increasing functions of v, and if $z^*(v)$, $q^*(v)$ which solve

$$\max_{z,q}\{B(q, z; v) - B_3(q, z; v)/\rho(v) - cq\}$$

are each either zero or strictly increasing in v, then the expected profit-maximizing selling strategy consists of the schedule $\langle q^*(x), z^*(x), T^*(x)\rangle_{x\in[0,1]}$, where

$$T^*(x) = B(q^*(x), z^*(x); x) - \int_0^x B_3(q^*(s), z^*(s); s)ds. \tag{35}$$

[13] Mussa and Rosen consider the special case in which preferences are given by

$$u = z(vr - T) + I$$

and $C(r)$, the cost of producing a unit of quality level r, is an increasing convex function. Defining $q \equiv C(r)$, we can rewrite preferences as

$$u = z(vA(q) - T) + I,$$

where $A(q) \equiv C^{-1}(q)$ is an increasing concave function. In the new quality units, marginal cost is constant and equal to unity.

As an immediate corollary, note that with $q^*(x)$ strictly increasing whenever q^* is greater than zero, we can define the inverse function $x = \phi(q)$ and hence the number of units and total payments as functions of q,

$$z^{**}(q) \equiv z^*(\phi(q)),\ T^{**}(q) = T^*(\phi(q)).$$

Then the optimal selling strategy can be reinterpreted as a schedule of offers of the form $\langle q, z^{**}(q), T^{**}(q)\rangle$. That is, the monopolist announces that quality level q will be sold in bundles of $z^{**}(q)$ units for a total cost of $T^{**}(q)$.

It is perhaps surprising that the monopolist does not gain by announcing a different quantity discount schedule for each quality level. But with only a single unobservable characteristic, v, the optimal schedule of bundles, $\langle q^*(v), z^*(v)\rangle$, traces out a curve in quality-quantity space. Given the hypotheses of Proposition 8, quantity increases with quality along this curve. Thus, both the unit price and the number of units in the bundle can be expressed as a function of product quality.

Space constraints preclude the derivation of conditions that ensure the monotonicity of $q^*(\cdot)$ and $z^*(\cdot)$ defined in Proposition 8 (that is, the counterpart of Proposition 4). Instead, we conclude with a simple example.

Example 3:

$$B(q, z; v) = 2vqz^{1/2} - q^3/3; \qquad F(v) = v.$$

Applying Proposition 8, we solve for $q^*(x)$ and $z^*(x)$ and obtain

$$q^*(x) = \begin{cases} 0, & x < 1/2 \\ (2x-1)/c^{1/2}, & x \geq 1/2. \end{cases} \tag{36}$$

$$z^*(x) = \begin{cases} 0, & x < 1/2 \\ q^*(x)^2/c, & x \geq 1/2. \end{cases} \tag{37}$$

Since both $q^*(x)$ and $z^*(x)$ satisfy the hypotheses of the proposition, we can substitute them into (35), the expression for $T^*(x)$, to obtain

$$T^*(x) = q^*(x)^3/3 + q^*(x)^2/c^{1/2}. \tag{38}$$

Combining (36)–(38), we therefore obtain

$$\langle q, z^{**}(q), T^{**}(q)\rangle = \langle q, q^2/c, q^3/3 + q^2/c^{1/2}\rangle,$$

the optimal bundling strategy.

6. Concluding remarks

■ In this article we have suggested a method for solving the choice problem of a principal (seller) who selects the rules of a game that is then played noncooperatively by a set of agents (buyers). We should repeat the two assumptions that are crucial to the success of this method.

First, buyers are assumed to exhibit no income effects and to be neutral toward income risk (although not necessarily neutral toward risk associated with the good being sold). With risk aversion, the analysis is much more complicated. For the relatively simple case where a single item is up for auction, several papers have compared specific auction rules such as the high bid and English auctions. See, for example, Holt (1980), Riley and Samuelson (1981), and Milgrom and Weber (1982). More recently, Matthews (1983) and Maskin and Riley (1984) have characterized the *optimal* auction, given risk aversion.

Second, we suppose that the underlying family of demand curves can be described by the variation of a single parameter. This is an assumption common to virtually all work to date on adverse selection. Unfortunately, work by Laffont, Maskin, and Rochet (1982) suggests that extensions to even two-parameter families will not be easy.

One restrictive feature of the models of Sections 2–5 is that the seller simply offers buyers a price schedule; there is no negotiation between the parties. That is, the seller effectively has all the monopoly power. In many cases, however, we might expect points on the Pareto frontier other than the seller's favorite to emerge as the outcome of contracting.

For example, consider the problem facing a manager (the "principal") who must engage an expert (the "agent") to perform some service. The principal's payoff is an increasing function of the quality, q, of the service provided and a decreasing function of the payment, Ω, to the agent. In fact, suppose that the principal is neutral toward income risk, with payoff function

$$R = B(q) - \Omega.$$

Quality is an increasing function of the agent's effort, e, and a random exogenous parameter $\tilde{v}$:

$$q = D(e; \tilde{v}).$$

The principal observes neither v, the realization of $\tilde{v}$, nor the agent's effort. He does, however, know the distribution, $F(\cdot)$, of $\tilde{v}$. A contract, negotiated before the realization of $\tilde{v}$, specifies a payment by the principal to the agent contingent on the quality, q, that is observed by both parties. The agent can observe v and choose an effort level after making this observation. We suppose further that, after observing v, the agent has the option of terminating the contract.[14] In that event, both the principal and agent obtain a zero payoff. Finally, we suppose that the agent is also neutral towards income risk, with payoff function

$$U = \Omega - A(e).$$

Because the principal can observe neither the random shock nor effort, it may appear as though this model, unlike those of the previous sections, exemplifies moral hazard as well as adverse selection (see the Introduction). But because we suppose that the agent chooses an effort level only *after* he observes v, the problem reduces, in effect, to one of pure adverse selection. This results because, assuming a monotonic production function, the principal could deduce the agent's effort if he knew the realization of $\tilde{v}$. Thus, only the asymmetric information about the value of $\tilde{v}$ prevents an efficient allocation. Indeed, we can "eliminate" effort from the model by defining the net social benefit function

$$N(q; v) = \max_{e}\{B(q) - A(e) | q \leqslant D(e; v)\}$$

and reexpressing the agent's payoff function as

$$U = N(q; v) - R.$$

Thus, the problem has exactly the same structure as that of monopoly pricing. Assuming that there are no other individuals in the economy with the same skills as the expert, both parties have monopoly power. Thus, the outcome of efficient contracting will be some point on the Pareto frontier.

Under the assumptions of Sections 2–4, it is possible to say quite a bit about the shape of this frontier (for the details, consult Maskin and Riley (1983b)). For example, under Assumptions 1, 2, and 3, a movement along the Pareto frontier in favor of the agent lowers the return function $\bar{R}(q)$ for each value of q. Furthermore, such a movement reduces q for each value of v, unless q equals zero. Under this same set of assumptions, we can also conclude that the Pareto frontier is concave, i.e., bowed outwards.

---

[14] If an agreement between the principal and agent prohibiting the agent from later terminating can be costlessly enforced, the incentive problem is readily solved. Since both parties are assumed to be neutral towards income risk, the principal simply offers to pay the total benefit $B(q)$ less some fixed sum P. The agent thus receives the full marginal benefit of his action.

Appendix

■ Proposition 7 and its proof follow.

Proposition 7: General characterization of the optimal response function.

(A) Suppose Assumptions 1, 2, and 5 are satisfied. Then if there exists an optimal response function, $q^*(v)$, there exists a set of subintervals $\{[x^i, y^i] \subseteq [0, \bar{v}] | x^{i-1} > y^i\}_{i \in I}$, possibly empty, such that for all i

$$\text{(i)} \qquad \bar{q}(x^i) = \bar{q}(y^i)$$

$$\text{(ii)} \qquad \int_v^y \frac{\partial I}{\partial q}(\bar{q}(y), z)dF(z) \leqslant 0, \qquad v < y \qquad y \notin \bigcup_{i \in I} (x^i, y^i)$$

with equality if $y = y^i$ and $v = x^i$ ($x^i > 0$). Moreover,

$$\text{(iii)} \qquad q^*(v) = \begin{cases} \bar{q}(y^i) & \text{if} \quad v \in [x^i, y^i] \text{ for some } i \\ \bar{q}(v), & \text{otherwise.} \end{cases}$$

(B) Given Assumptions 1, 2, and 5, a collection of subsets satisfying (i) and (ii) exists. If, in addition, Assumption 6 holds, then $q^*(v)$, defined by (iii), is an optimal response function.

Proof of (A). Assume that Assumptions 1, 2, and 5 hold and suppose that $q^*(v)$ is an optimal response function. That is, $q^*(v)$ solves

$$\max_{q(v)} \left\{ R_o = \int_0^{\bar{v}} I(q(v); v)dF(v) | q(v) \text{ nonnegative and nondecreasing} \right\}. \tag{A1}$$

We note first that for v such that $\bar{q}(v) > 0$, $\frac{\partial I}{\partial q}(\bar{q}(v); v) = 0$. Therefore, for such v, Assumption 5 implies that $\frac{\partial^2 I}{\partial q^2}(\bar{q}(v); v) < 0$. Hence,

$$\frac{d\bar{q}}{dv} = -\frac{\partial^2 I}{\partial q \partial v} \Big/ \frac{\partial^2 I}{\partial q^2},$$

and we conclude that $\bar{q}(v)$ is differentiable.

We next establish that, for any $\hat{v}$,

$$\left. \begin{array}{lllll} \text{if} & q^*(\hat{v}) \neq \bar{q}(\hat{v}), & \text{there exists a nonempty interval } [z_1, z_2] & & \\ \text{such that} & \hat{v} \in [z_1, z_2] & \text{and} \quad q^*(v) = q^*(\hat{v}) & \text{for all} & v \in [z_1, z_2] \end{array} \right\}. \tag{A2}$$

Suppose that this is false and there exists $\hat{v}$ such that $q^*(\hat{v}) \neq \bar{q}(\hat{v})$, but q^* is strictly increasing at $\hat{v}$ both from the left and to the right. If $\bar{q}^*(\hat{v}) < \bar{q}(\hat{v})$, then, because $\bar{q}(v)$ is continuous, there exists $v_* < \hat{v}$ such that

$$q^*(v) < q^*(\hat{v}) < \bar{q}(v) \qquad \text{for all} \qquad v \in (v_*, \hat{v}).$$

Define

$$\hat{q}(v) = \begin{cases} q^*(\hat{v}), & v \in (v_*, \hat{v}) \\ q^*(v), & \text{otherwise.} \end{cases}$$

Then from Assumption 5, $\int_0^{\bar{v}} I(\hat{q}(v); v)dF(v) > \int_0^{\bar{v}} I(q^*(v); v)dF(v)$, a contradiction of the optimality of q^*. A similar contradiction results if $q^*(\hat{v}) > \bar{q}(\hat{v})$. Hence, (A2) holds.

We next show that

$$q^*(v) \text{ is a continuous function.} \tag{A3}$$

Let $\{(x^i, y^i) | y^{i-1} > x^i\}_{i \in I}$ be the set of "maximal" intervals on which $q^*(v)$ is constant. That is,

$$\forall v < x^i, \qquad q^*(v) < q^*(x^i), \qquad \text{and} \qquad \forall v > y^i, \qquad q^*(v) > q^*(y^i).$$

From (A2) and the continuity of $\bar{q}(v)$, $q^*(v) = \bar{q}(v)$ for $v \notin \bigcup_{i \in I} [x^i, y^i]$, and it remains to show that $q^*(v)$ is continuous at x^i, if $x > 0$, and at y^i. Suppose there is an upward discontinuity at y^i. Since $\bar{q}(v)$ is continuous, there exists $z < y^i$ such that

$$\bar{q}(v) > \lim_{v \uparrow y^i} q^*(v), \qquad v \in [z, y^i).$$

Suppose that over $[z, y^i]$, $\bar{q}(v)$ takes on its minimum at $\hat{v}$. Then

$$q^*(v) = \lim_{v \uparrow y^i} q^*(v) < \bar{q}(\hat{v}), \qquad v \in [z, y^i).$$

Define

$$\hat{q}(v) = \begin{cases} q^*(\hat{v}), & v \in (z, y^i); \\ q^*(v), & \text{otherwise.} \end{cases}$$

Arguing exactly as before, one can show that expected seller revenue is greater under $\hat{q}$ than q^*, thereby contradicting the hypothesis that q^* is an optimal response function. Then $q^*(v)$ is continuous at y^i. Similarly, $q^*(v)$ is continuous at x^i as well, if $x^i > 0$.

Given (A2) and (A3), we can readily establish that the following conditions are also necessary:

$$\int_{\hat{v}}^{\bar{v}} \frac{\partial I}{\partial q}(q^*(v); v)dF(v) \leqslant 0, \qquad \hat{v} < \bar{v} \tag{A4}$$

and

$$\int_{x^i}^{y^i} \frac{\partial I}{\partial q}(q^*(v); v)dF(v) = 0, \qquad i \in I, \qquad x^i > 0. \tag{A5}$$

For any $\hat{v} \in [0, \hat{v}]$, define the nondecreasing function

$$q^+(v, s) = \begin{cases} q^*(v); & v < \hat{v} \\ q^*(v) + s; & v \geqslant \hat{v} \end{cases} \tag{A6}$$

for $s \geqslant 0$. Replacing $q^*(v)$ by $q^+(v, s)$, we can write expected seller revenue $\bar{R}_o(s)$ as a function of s:

$$\bar{R}_o(s) = \int_0^{\bar{v}} I(q^+(v, s); v)dF(v).$$

At $s = 0$, the right derivative of $\bar{R}_o(s)$ is

$$\frac{d\bar{R}_o^+}{ds}(0) = \int_{\hat{v}}^{\bar{v}} \frac{\partial I}{\partial q}(q^*(v); v)dF(v).$$

Since $q^*(v)$ is optimal, $\dfrac{d\bar{R}_o^+}{ds}(0)$ must be nonpositive; hence, we obtain (A4).

Because $\dfrac{\partial I}{\partial q}(q^*(v); v) = \dfrac{\partial I}{\partial q}(\bar{q}(v); v) = 0$ on $[y^1, \bar{v}]$,

$$\int_{x^1}^{y^1} \frac{\partial I}{\partial q}(q^*(v); v)dF(v) = \int_{x^1}^{\bar{v}} \frac{\partial I}{\partial q}(q^*(v); v)dF(v). \tag{A7}$$

Define the nondecreasing function

$$\bar{q}(v, s) = \begin{cases} \min\{q^*(v), q^*(x^1) - s\}, & v < x^1 \\ q^*(v) - s & v \geqslant x^1. \end{cases}$$

Because $q^*(v)$ is strictly increasing in a left neighborhood of x^1, there exists a function $v(s)$ such that for small nonnegative s,

$$\bar{q}(v, s) = \begin{cases} q^*(v), & v < v(s) \\ q^*(x^1) - s, & v(s) \leqslant v \leqslant x^1 \end{cases},$$

and $v(s)$ is differentiable, and tends to x^1 as s tends to zero. Replacing $q^*(v)$ by $\bar{q}(v, s)$, we obtain expected seller revenue

$$\bar{\bar{R}}_o(s) = \int_0^{\bar{v}} I(\bar{q}(v, s); v)dF(v).$$

At $s = 0$, the right derivative of $\bar{\bar{R}}(s)$ is

$$\frac{d\bar{\bar{R}}(0)^+}{ds} = -\int_{x^1}^{\bar{v}} \frac{\partial I}{\partial q}(q^*(v); v)dF(v). \tag{A8}$$

The optimality of $q^*(v)$ implies that $\dfrac{d\bar{\bar{R}}^+}{ds}(0)$ must be nonpositive. Thus (A8) implies

$$\int_{x^1}^{\bar{v}} \frac{\partial I}{\partial q}(q(v); v)dF(v) \geqslant 0. \tag{A9}$$

Formula (A9), together with (A4) and (A7), establishes that (A5) must hold for $i = 1$. Virtually the same argument establishes that (A5) holds for $i > 1$ as well.

By assumption, $q^*(v) = \bar{q}(v)$ for $v \notin \bigcup_{i \in I} (x^i, y^i)$. By the continuity of q^* and from (A2), $q^*(v) = q^*(y^i)$ for all $v \in [x^i, y^i]$. Hence, (i) and (iii) of the statement of the Proposition hold. Conditions (A4) and (A5) imply that (ii) holds. *Q.E.D.*(A)

Proof of (B). An alternative way of deriving (A2)–(A5) is to formulate the seller's optimization as an exercise in optimal control, with the associated Lagrangian

$$\mathcal{L} = \int_0^{\bar{v}} [I(q(v); v)F'(v) + \lambda(v)q'(v)]dv.$$

If we take

$$\lambda(v) \equiv -\int_v^{\bar{v}} I(q(z); z)dF(z), \tag{A10}$$

then it is straightforward to show that, given subintervals satisfying (i) and (ii), (A2)–(A5), together with the requirements that $q(v)$ be nonnegative and nondecreasing, are equivalent to the first-order conditions for a maximum corresponding to this Lagrangian. Hence, $q^*(v)$ defined by (iii) satisfies the first-order conditions.

As for the second-order conditions, sufficient conditions for a critical point (one satisfying the first-order conditions) to be a local maximum are that

$$\frac{\partial^2 I}{\partial q^2}(q(v); v) < 0 \quad \text{if} \quad \lambda(v) = 0, \tag{A11}$$

$$\int_\alpha^\beta \frac{\partial^2 I}{\partial q^2}(q(v); v)dF(v) < 0 \quad \text{if} \quad \lambda(\alpha) = \lambda(\beta) = 0 \quad \text{and} \quad \lambda(v) > 0 \quad \text{for all} \quad v \in (\alpha, \beta), \tag{A12}$$

and

$$\int_0^\beta \frac{\partial I}{\partial q}(q(z); z)dF(z) \leqslant 0 \quad \text{if} \quad \lambda(\beta) = 0 \quad \text{and} \quad \lambda(v) > 0 \quad \text{for all} \quad v \in (0, \beta). \tag{A13}$$

Now given (A10), $\lambda(v) = 0$ implies that

$$\int_v^{\bar{v}} \frac{\partial I}{\partial q}(q(z); z)dF(z) = 0. \tag{A14}$$

If $q(v) \neq \bar{q}(v)$, then $v \in (x^i, y^i)$ for some i. Furthermore, Assumption 5 implies that $\frac{\partial I}{\partial q}(q(v); v) \neq 0$. Therefore, in view of (A14), there exists $\tilde{v}$ in a neighborhood of v such that

$$\int_{\tilde{v}}^{\bar{v}} \frac{\partial I}{\partial q}(q(z); z)dF(z) > 0, \tag{A15}$$

a contradiction of (A4). Hence $\lambda(v) = 0$ implies that $\frac{\partial I}{\partial q}(q(v); v) = 0$. Thus, if Assumption 5 is satisfied, (A11) holds.

If $\lambda(\alpha) = \lambda(\beta) = 0$ and $\lambda(v) > 0$ for all $v \in (\alpha, \beta)$, then $(\alpha, \beta) = (x^i, y^i)$ for some i. Hence $q(v) = \bar{q}(\beta)$ for all $v \in (x^i, y^i)$, and $\int_\alpha^\beta \frac{\partial I}{\partial v}(\bar{q}(\beta); v)dF(v) = 0$. Therefore, Assumption 6 implies that (A12) holds. Similarly, it also implies (A13). Hence, under Assumptions 1, 2, 5, and 6, any $q(v)$ satisfying (A2)–(A5) is a global maximum.

It remains to show that there exists a collection of subintervals satisfying (i) and (ii) (since, from the above argument, $q^*(v)$ defined by (iii) will then be an optimal response function). Let

$$Y^1 = \left\{ v \,\middle|\, \int_x^y \frac{\partial I}{\partial q}(\bar{q}(y); z)dF(z) \leqslant 0,\ \forall x \leqslant y \forall y \geqslant v \right\}. \tag{A16}$$

For $\bar{q}(v) > 0$, we have

$$0 = \frac{\partial I}{\partial q}(\bar{q}(v); v) = p(\bar{q}(v); v) - c - \frac{p_2(\bar{q}(v), v)}{\rho(v)}.$$

By Assumption 1 and from the definition of $\rho(v)$, the last term is negative for $v < \bar{v}$ and zero at $\bar{v}$. Thus by Assumption 5,

$$\begin{aligned} \bar{q}(v) &< q^e(v), \qquad v < \bar{v} \\ &= q^e(v), \qquad v = \bar{v}, \end{aligned}$$

where $q^e(v)$ is the efficient quantity level. Therefore, from Assumption 5, $\frac{\partial I}{\partial q}(\bar{q}(\bar{v}); v) < 0$ for all $v < \bar{v}$, and so, for all v sufficiently close to $\bar{v}$, $v \in Y^1$.

Let $y^1 = \inf Y^1$. Then $y^1 < \bar{v}$. If $y^1 = 0$, we are done; the sequence is of zero length. Therefore, assume that $y^1 > 0$. By continuity

$$\int_x^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) \leq 0 \qquad \text{for all} \qquad x \leq y^1. \tag{A17}$$

We first argue that there exists $\epsilon > 0$ such that $\bar{q}$ is nondecreasing on $[y^1 - \epsilon, y^1]$. If not, then because $\bar{q}$ is piecewise differentiable, $\bar{q}$ must be decreasing in some interval $[y^1 - \epsilon, y^1]$. But then, from Assumption 5, $\int_{y^1-\epsilon}^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) > 0$, thereby contradicting (A17).

By definition of y^1, there exists an increasing sequence $\{v_n\} \subseteq [y^1 - \epsilon, y^1]$ such that for each n there exists $w_n < v_n$ with

$$\int_{w_n}^{v_n} \frac{\partial I}{\partial q}(\bar{q}(v_n); z)dF(z) > 0. \tag{A18}$$

Because $\bar{q}$ is nondecreasing on $[y^1 - \epsilon, y^1]$, $w_n < y^1 - \epsilon$ for all n. Let $w = \lim_{n\to\infty} w_n$ (if $\{w_n\}$ does not converge, let w be the limit of a subsequence that does). Then $w < y^1$ and in view of (A17) and (A18),

$$\int_w^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) = 0.$$

Thus, the set $X^1 = \left\{x \,\middle|\, \int_x^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) = 0, x < y^1\right\}$ is nonempty. Let $x^1 = \inf X^1$. Then

$$\int_{x^1}^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) = 0. \tag{A19}$$

If $x^1 = 0$, we are done. Therefore, assume that $x^1 > 0$.

We claim that $\bar{q}(x^1) = \bar{q}(y^1)$. First, suppose that $\bar{q}(x^1) < \bar{q}(y^1)$. From Assumption 5, $\frac{\partial I}{\partial q}(q; x^1) < 0$ for all $q > \bar{q}(x^1)$. Thus, $\frac{\partial I}{\partial q}(\bar{q}(y^1); x^1) < 0$, and so, from (A19), for small $\epsilon > 0$,

$$\int_{x^1+\epsilon}^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) > 0,$$

a contradiction of (A17). If, on the other hand, $\bar{q}(x^1) > \bar{q}(y^1)$, then from Assumption 5, $\frac{\partial I}{\partial q}(\bar{q}(y^1); x^1) > 0$, in which case $\int_{x^1-\epsilon}^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) > 0$ for $\epsilon > 0$ sufficiently small, again a contradiction. Thus,

$$\bar{q}(x^1) = \bar{q}(y^1).$$

Let

$$Y^2 = \left\{v \leq x^1 \,\middle|\, \int_x^y \frac{\partial I}{\partial q}(\bar{q}(y); z)dF(z) \leq 0, \forall x \leq y \,\forall y \in [v, x^1]\right\}.$$

For all $v \leq x^1$, $\int_v^{x^1} \frac{\partial I}{\partial q}(\bar{q}(x^1), z)dF(z) = \int_v^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) \leq 0$, from (A17) and (A19) and from the fact that $\bar{q}(x^1) = \bar{q}(y^1)$. Therefore, $x^1 \in Y^2$. If $x^1 = \inf Y^2$, then from the above argument, there exists $x < x^1$ such that $\int_x^{x^1} \frac{\partial I}{\partial q}(\bar{q}(x^1); z)dF(z) = 0$. But then $\int_x^{y^1} \frac{\partial I}{\partial q}(\bar{q}(y^1); z)dF(z) = 0$, thereby contradicting the fact that $x^1 = \inf X^1$. Thus, $\inf Y^2 < x^1$. If we take $y^2 = \inf Y^2$, we can continue, iteratively, to generate $x^2, y^3, \ldots$. The process terminates because $\bar{q}$ is piecewise differentiable. *Q.E.D.*(B)

References

BARON, D. AND MYERSON, R.B. "Regulating a Monopolist with Unknown Costs." *Econometrica,* Vol. 50 (1982), pp. 911–930.

BRITO, D.L. AND OAKLAND, W.H. "On the Monopolistic Provision of Excludable Public Goods." *American Economic Review,* Vol. 70, No. 4 (1980), pp. 691–674.

GOLDMAN, M.B., LELAND H.E., AND SIBLEY, D.S. "Optimal Nonuniform Prices." Mimeo, 1980.

HARRIS, M. AND RAVIV, A. "A Theory of Monopoly Pricing with Uncertain Demand." *American Economic Review,* Vol. 71 (1981), pp. 347–365.

HART, O.D. "Optimal Labour Contracts under Asymmetric Information: An Introduction." *Review of Economic Studies,* Vol. 50 (1983), pp. 3–36.

HOLT, C. "Competitive Bidding for Contracts under Alternative Auction Procedures." *Journal of Political Economy,* Vol. 88 (1980), pp. 433–445.

LAFFONT, J.J., MASKIN, E.S., AND ROCHET, J.C. "Optimal Nonlinear Pricing with Two-Dimensional Characteristics." Mimeo, 1982.

MASKIN, E.S. AND RILEY, J.G. "Auctioning an Indivisible Object." Kennedy School of Government Working Paper #87D, Harvard University, April 1980.

——— AND ———. "Optimal Multiunit Auctions." UCLA Discussion Paper, 1983a.

——— AND ———. "Monopoly with Incomplete Information." UCLA Discussion Paper, 1983b.

——— AND ———. "Optimal Auctions with Risk Averse Buyers." *Econometrica,* Vol. 52 (1984), forthcoming.

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

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

MIRMAN, L.J. AND SIBLEY, D. "Optimal Nonlinear Prices for Multiproduct Monopolies." *Bell Journal of Economics,* Vol. 11, No. 2 (Autumn 1980), pp. 659–670.

MIRRLEES, J. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies,* Vol. 38 (1971), pp. 175–208.

MUSSA, M. AND ROSEN, S. "Monopoly and Product Quality." *Journal of Economic Theory,* Vol. 18 (1978), pp. 301–317.

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

OI, W. "A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly." *Quarterly Journal of Economics,* Vol. 85 (1971), pp. 77–96.

OREN, S.S., SMITH, S.A., AND WILSON, R.B. "Pricing a Product Line." *Journal of Business,* Vol. 57, No. 1, Part 2 (1984), pp. S73–S100.

PIGOU, A.C. *The Economics of Welfare.* London: MacMillan, 1930.

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

ROBERTS, K.W.S. "Welfare Implications of Nonlinear Prices." *Economic Journal,* Vol. 89 (1979), pp. 66–83.

SPENCE, A.M. "Nonlinear Prices and Welfare." *Journal of Public Economics,* Vol. 8 (1977), pp. 1–18.

STIGLITZ, J.E. "Monopoly, Nonlinear Pricing, and Imperfect Information: The Insurance Market." *Review of Economic Studies,* Vol. 44 (1977), pp. 407–430.

VICKREY, W. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance,* Vol. 16 (1961), pp. 8–37.