JOURNAL OF ECONOMIC THEORY 18, 301–317 (1978)

Monopoly and Product Quality

MICHAEL MUSSA AND SHERWIN ROSEN

Graduate School of Business, Department of Economics, University of Chicago, Chicago, Illinois 60637

Received June 19, 1975; revised August 2, 1978

1. INTRODUCTION

This paper considers a class of monopoly pricing problems involving what a businessman might call a product line, a quality-differentiated spectrum of goods of the same generic type. While the goods are similar, they are not perfect substitutes because all customers do not place the same valuations on all attributes of the goods. The seller knows the general distribution of tastes and demands in the market, but cannot distinguish among buyers prior to an actual sale and cannot prevent resale in other markets. Therefore, the monopolist cannot engage in the usual sort of price discrimination. Instead, the goods are offered in an impersonal market on a take-it-or-leave-it basis and the seller exploits the possibilities for a pricing policy (a price–quality schedule) to allocate customers along the quality spectrum by a process of self-selection. The optimal policy "smokes out" consumer preferences, separates markets, and assigns different customer types to different varieties of goods, thereby permitting partial discrimination among consumers of varying intensities of demand.

Assuming that buyers purchase one unit of the good and that there are constant costs of producing a given variety and increasing marginal costs of higher quality items, it is established that the monopolist almost always reduces the quality sold to any customer compared with what would be purchased under competition. Generally speaking this is done by increasing the slope of the price–quality gradient offered relative to marginal cost. Furthermore, the monopolist frequently prices customers with the least intensities of demand out of the market, while at the same time selling a broader range of qualities than would be offered in competitively organized markets. Finally, demand conditions may be such that it does not pay the seller to separate all markets completely, but rather to bunch customers of different tastes onto the same product. This maneuver is accomplished by imparting corners in the price–quality schedule, so that customers with

0022-0531/78/0182–0301$02.00/0

somewhat different marginal valuations of quality find it in their interest to purchase the same quality item.

When goods are related along a quality or similar dimension, it is intuitively clear that monopoly pricing rules should serve a self-selection role of discriminating among consumers and separating markets. Yet, the point is hard to pin down with existing theory of discriminating monopoly because conventional demand functions do not conveniently characterize substitutions along the product line when there are a large number of closely related goods. We handle this issue by analyzing underlying utility functions, on the hypothesis that differentiated products can be compared by decomposing them into more fundamental measurable attributes (see [7]). In fact, the problem bears some resemblance to spatial equilibrium models (e.g. [4]), and also to some normative questions in the theory of optimal taxation [5, 8]. Appeal to that literature suggests considerable technical difficulties with extremely general models. Befitting a new way of looking at an old problem, our strategy is to employ as many simplifying assumptions as necessary to make the problem manageable and to reveal its fundamental structure.

The basic ingredients of the analysis are discussed in the next section. A useful example illustrating the logic of the monopoly solution is presented in Section 3. The formal solution and properties of the monopolist's profit maximization problem are developed and illustrated in Sections 4 and 5. Implications of the present analysis for recent discussions of product durability and market organization, and some possible extensions, are considered in the concluding section.

2. Demand and Cost Conditions

Consider the market for a commodity which can be produced in a number of different varieties. Let q represent the underlying hedonic attributes that characterize a particular variety. Referred to as "product quality," q is restricted to one dimension, with larger values of q indicating higher quality varieties. All varieties are sold on an impersonal market and the same price is charged to all buyers of a given quality. Market equilibrium is described by a set of prices, one for each quality, $P(q)$; the number of each q transacted, $N(q)$; and the observed "breadth" of the product line, $[\underline{q}, \bar{q}]$. All buyers are offered the same price–quality schedule and different varieties are produced to cater to different types of customers.

The unit cost for any particular quality, $C(q)$, is assumed to be constant, independent of the number of units of that or any other variety. This assumption is useful because it permits the analysis to focus clearly on the role of demand conditions and subsitution among product varieties from self-selection, without complications arising from interactions among the costs of producing different varieties.[1] Unit and marginal cost are assumed to be increasing functions of quality, $C'(q) > 0$, $C''(q) > 0$ for all feasible qualities $q \geqslant 0$.

Each consumer has a utility function $U(x, q; \theta)$, where x is a composite commodity other than the generic type in question and θ indexes customer types. Consumers choose their optimal variety by maximizing utility subject to the budget constraint, $P(q) + x \leqslant y$, where y is income, and y and $P(q)$ are measured in terms of x. As pointed out above, some representation of taste differences among consumers is essential for the main questions to be analyzed. Assuming that each consumer spends a small fraction of total expenditure on the type of goods in question and ignoring income effects, U may be approximated by[2]

$$U = x + \theta q, \tag{1}$$

where $\theta > 0$ is a parameter that measures intensity of a consumer's taste for quality. Consumers' valuations of quality vary in proportion to θ, so that the preferences of the set of potential consumers are described by the distribution of θ, a density $f(\theta)\, d\theta$ defined on the interval $[\underline{\theta}, \bar{\theta}]$.

Given this specification of technology and preferences, the competitive solution is easily described. Competition and free entry require that profits must be zero for all q: Hence, $P(q) = C(q)$. Given this price–quality schedule, each consumer is self-assigned to the quality that maximizes utility, requiring, for consumers who purchase positive qualities, the first-order condition

$$P'(q) = \theta. \tag{2}$$

Let $q(\theta)$ denote the quality purchased by a consumer of type θ. Then

$$q(\theta) = C'^{-1}(\theta) \equiv J(\theta)$$

is the assignment of customer type to product type in the competitive market. All consumers buy the good if $C'(0) < \underline{\theta}$ and the product line extends from $\underline{q} = J(\underline{\theta})$ to $\bar{q} = J(\bar{\theta})$. If $C'(0) > \underline{\theta}$, only consumers for whom $\theta \geqslant C'(0)$ purchase the good, and the product line extends from $q = 0$ to $\bar{q} = J(\bar{\theta})$. The density of varieties actually sold is given by $N(q)\, dq = f[C'(q)]C''(q)\, dq$.

[1] This assumption rules out indivisibilities or cost savings from large production runs or from piling up production at isolated nodes along the quality spectrum. These phenomena lead to monopolistic competition in which not all possible goods are produced, and need to be analyzed in a rather different manner (see [3]).

[2] The specification of the utility function could be generalized to $U = x + \theta \cdot h(q)$, with $h' > 0$ and $h'' < 0$. This generalization simply involves a redefinition of the units in which "quality" is measured. The example in Section 3 suggests that the alternative route of specifying a more general utility function with different marginal valuations induced by a distribution of income would lead to similar conclusions.

3. An Example

Many of the essential features of both the competitive and monopolistic solutions can be illustrated with a simple example. Figure 1 translates the origin of a standard indifference map to show trade-offs between quality and expenditures on quality (measured in the negative direction) and consumer surplus, z (measured in the positive direction). The curve $C(q)$ shows the schedule of prices that must be paid in the competitive market to purchase alternative qualities. Consumers of type θ^1 maximize utility at point A, where $C'(q^1) = \theta^1$, purchasing quality q^1, paying price $p^1 = C(q^1)$ and enjoying consumer suplus of $z^1 = \theta^1 q^1 - p^1$. Similarly, consumers of type θ^2 maximize utility at point B where $q = q^2$, $C'(q^2) = \theta^2, p^2 = C(q^2)$ and $z^2 = \theta^2 q^2 - p^2$.

It is useful to consider how a perfectly discriminating monopolist would behave with regard to these two types of consumers. A perfectly discriminating monopolist does not sell in an impersonal market and can charge different prices to different consumers, even for the same quality item, in order to extract all potential consumer surplus. In the present case each consumer would be sold the same quality as would be purchased under competition, but at a much higher price: Consumers of type θ^1 would be offered the quality q^1 at a price just under $\theta^1 q^1$ (point D in Fig. 1), and consumers of type θ^2 would be offered the quality q^2 at a price just under $\theta^2 q^2$ (point E in Fig. 1).

Now, suppose that an impersonal monopolist offers all consumers the price–quality combinations corresponding to points D and E in Fig. 1. Consumers of type θ^1 still choose point D, but consumers of type θ^2 are not satisfied with the point E. They, too, choose D (a choice not available to them if perfect discrimination is feasible) because their utility is larger at D than at E, and the two markets are not separated. The offer of point D to θ^1 consumers *interferes* with the monopolist's ability to charge higher prices to θ^2 consumers, who, because of their higher taste for quality, have more potentially extractable consumer surplus.[3]

Stacking both types of consumers at D is not optimal for an impersonal monopolist. Given that point D continues to be offered, markets can be separated and the monopolist can increase profit by reducing price for quality q^2 sufficiently to induce consumers of type θ^2 to buy it. Thus, points D and F could be offered, instead of D and E, where at F, $P(q^2)$ is slightly below $\theta^1 q^1 + \theta^2(q^2 - q^1)$. Then θ^2 customers choose F rather than D. However, there is nothing to compel the monopolist to offer point D and it is not optimal to do so. If the quality of the low quality variety is reduced by Δq and its price is reduced by $\theta^1 \Delta q$, consumers of type θ^1 are content to buy this new

[3] A geometry similar to that of Fig. 1 has been employed by Wilson (1975) in studying adverse selection in insurance markets. That work has been helpful to us in uncovering the structure of the present problem.

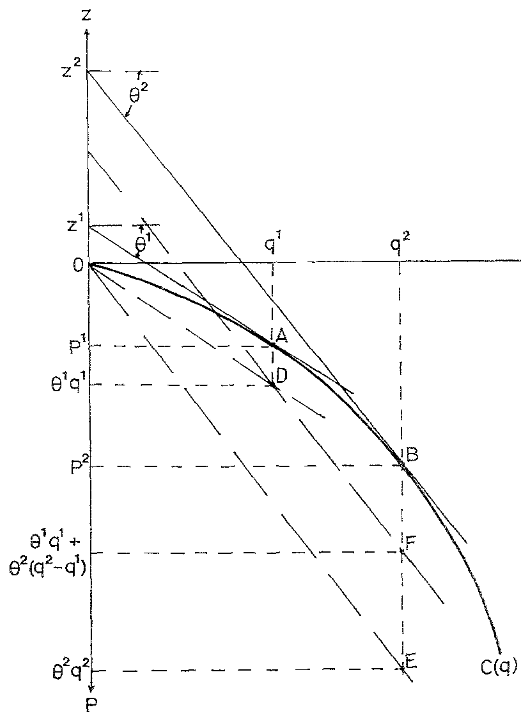

FIG. 1. Competitive and monopoly allocations to two consumer types.

lower quality item and markets are still separated. Since $C'(q^1) = \theta^1$ at D, this shift of θ^1 customers results in a profit reduction which is "second-order of smalls." But by reducing the quality sold to θ^1, the monopolist can increase the price charged for q^2 by an amount $(\theta^2 - \theta^1)\,\Delta q$, without inducing consumers of type θ^2 to shift to the low quality variety. The gain in profit from sales of q^2 at the higher price is "first-order of smalls." Hence, it always pays the monopolit to reduce quality sold to low θ consumers in order to be able to charge higher prices to high θ consumers. The extent to which it pays depends on the relative numbers of consumers of each type. If θ^1 consumers are relatively numerous, the reduction in quality is small, and, conversely, if θ^1 consumers are not very numerous, it may not pay to serve them at all. In either case the quality consumed by θ^1 consumers must fall, compared with the competitive outcome.[4]

[4] If there are two types of consumers, θ^1 and θ^2, with quality q^a sold to θ^1 consumers and quality q^b sold to θ^2 consumers, then the monopolist's price quality offers must satisfy $p^a = \theta^1 q^a$ and $p^b = p^a + \theta^2 \cdot (q^b - q^a)$. Total profits are $n^1 \cdot (\theta^1 q^a - C(q^a)) + n^2 \cdot (\theta^2 q^b - (\theta^2 - \theta^1) q^a - C(q^b))$, where n^1 and n^2 are the numbers of type 1 and type 2 consumers. Provided that $q^a > 0$, the necessary conditions for a maximum are $n^1 \cdot [\theta^1 - C'(q^a)] - n^2 \cdot (\theta^2 - \theta^1) = 0$ and $n^2 \cdot [\theta^2 - C'(q^b)] = 0$. The second condition implies that $q^b =$

Though the above example is simplified to an extreme, its logic applies in general. Serving customers who place smaller valuations on quality creates negative externalities for the monopolist that limit possibilities for capturing consumer surplus from those who do value quality highly. In the nature of the case, these external effects all go in an "upstream" direction. The monopolist internalizes them by inducing less enthusiastic consumers to buy lower quality items, opening the possibility of charging higher prices to more adamant buyers of high quality units. Thus, the impersonal monopolist achieves some, but imperfect, discrimination by taking advantage of the natural tendency of consumers to sort themselves out along the quality spectrum.

4. The Monopoly Solution

The example of Section 3 is difficult to generalize when there are more than a few types of consumers. It is better to go to the limit and adopt the fiction that there are an indefinite number of consumers with taste parameters distributed in accord with a continuous density function $f(\theta)\, d\theta$ in the range $[\underline{\theta}, \bar{\theta}]$. Under competition, each consumer purchases the variety $J(\theta)$ for which the marginal cost, $C'(J(\theta))$, is equal to the consumer's incremental demand price, θ. The problem for the monopolist is to find an assignment of qualities, $q(\theta)$, and associated prices, $p(\theta)$, which maximize profit subject to the constraints implied by consumer behavior.[5] The choice of $p(\theta)$, of course, implies the price–quality schedule $P(q)$.

Consider the constraints in more detail: First, it is apparent from Fig. 1 that $q(\theta)$ must be a nondecreasing function of θ. In choosing between two qualities, q^a and $q^b < q^a$, a consumer of type θ chooses q^b only if $P(q^a) - P(q^b) > \theta(q^a - q^b)$. The larger is θ, the greater the reduction in price required for a consumer to choose the lower quality. Hence, it is impossible to induce a consumer of type θ^i to purchase a lower quality item than would be purchased by a consumer of type $\theta^j < \theta^i$. From this constraint and from the fact that the monopolist can make positive profits from serving at least the high-θ consumers, it follows that the monopolist serves all consumers in some interval $[\theta^*, \bar{\theta}]$, where $\underline{\theta} \leqslant \theta^* < \bar{\theta}$.

$J(\theta^2)$; whereas the first condition implies that $q^a < J(\theta^1)$. Therefore, in contrast to the usual discriminating monopoly problem (e.g., the loss leader problem of Allen [1] or Edgeworth's taxation paradox discussed by Hotelling [2] in which cross elasticities of demand enter into the marginal conditions for the determination of both prices, this problem has a natural recursive structure: Interference runs from low- to high-θ consumers.

[5] $p(\theta)$ indicates price as a function of θ, while $P(q)$ indicates price as a function of q. Derivatives taken with respect to θ are indicated with a "dot" superscript, while derivatives with respect to q are indicated with a "prime" superscript.

Second, given the quality allocation or assignment function $q(\theta)$, (2) implies that the price function $p(\theta)$ must satisfy

$$p(\theta) = \theta \cdot q(\theta) - \int_{\underline{\theta}}^{\theta} q(s)\, ds. \tag{4}$$

To establish this fact it is useful to define

$$z(\theta) \equiv \theta \cdot q(\theta) - p(\theta) \tag{5}$$

as the consumer surplus accruing to a consumer of type θ. Consumer surplus must be nonnegative, since consumers have the option of not buying the good. Differentiating (5) and substituting (2) yields

$$\dot{z}(\theta) = dz(\theta)/d\theta = q(\theta). \tag{6}$$

Integrating (6),

$$z(\theta) = z(\theta^*) + \int_{\theta^*}^{\theta} q(s)\, ds. \tag{7}$$

However, it is obvious that at the "extensive" margin, θ^*, it is always optimal for the monopolist to make the marginal customer indifferent between buying and not buying. Clearly at this margin (which may possibly occur at $\theta^* = \underline{\theta}$), $z(\theta^*) = 0$, and (5) and (7) together imply (4).

Constraint (4) indicates precisely how servicing customers with small values of θ affects the monopolist's profitability of selling to customers with larger values of θ. Customers of type θ^i are assigned a quality $q(\theta^i)$ and would be willing to pay as much as $\theta^i q(\theta^i)$ for it. However, given the assignments to consumers with smaller values of θ, (7) shows that consumer surplus of type θ^i cannot be reduced below $\int_{\theta^*}^{\theta^i} q(s)ds$ because of the market availability of lower quality goods. Therefore, the maximum price that the seller can actually charge to θ^i, given the lower qualities offered on the market, is only $\theta^i q(\theta^i) - z(\theta^i) = p(\theta^i)$. A price higher than that would induce type θ^i to buy a lower quality good.

The extent of upstream interference from servicing consumer type θ is seen by calculating the effect on revenue of selling an additional increment of quality in a small neighborhood of some θ. There are $f(\theta)$ customers in this neighborhood and their incremental reservation price is θ. Hence revenue increases by $\theta f(\theta)$. However, (4) and (7) show that the price that can be charged to each and every customer with a larger value of θ falls by the increment sold to type θ. There are $1 - F(\theta)$ higher taste customers, where $F(\theta) = \int_0^{\theta} f(s)\, ds$. Therefore, the gain in revenue from the whole operation is $\theta f(\theta) - (1 - F(\theta)) \equiv MR(\theta) \cdot f(\theta)$, where

$$MR(\theta) \equiv \theta - [(1 - F(\theta))/f(\theta)] \tag{8}$$

is the marginal revenue associated with an increment of quality sold to consumers of type θ. We assume that $MR(\theta)$ is a well-defined smooth and differentiable function throughout $[\underline{\theta}, \bar{\theta}]$, i.e., $f(\theta)$ is a continuous and differentiable function over the interval with a positive density of consumers everywhere.

It is apparent from (8) that the marginal revenue associated with quality increments sold to consumers of type θ is less than incremental demand price, θ, for all values of θ except $\theta = \bar{\theta}$. This divergence is similar to the difference between price and marginal revenue in the standard monopoly problem and in fact is the key to the result that our monopolist restricts quality, analogous to the restriction of quantity that occurs in the standard case.

Indeed, it might be thought intuitively that equating marginal revenue at θ (Eq. 8) with marginal cost there, $C'(q(\theta))$, must characterize the monopoly solution, for all values of θ. In fact this is only partially correct, and the solution is rather more complicated: Such an assignment need not obey the constraint imposed by consumer behavior that $\dot{q}(\theta) \geqslant 0$, because $MR(\theta)$ need not be a monotonically increasing function of θ. Since $MR'(\theta) = 2 + (1 - F(\theta)) \cdot f'(\theta)/f(\theta)^2$, it follows that $MR(\theta)$ will be a decreasing function of θ in any range where $f'(\theta)$ is strongly negative. In such a range the monopolist cannot equate marginal revenue and marginal cost. Nor can he exclude the consumers in such a range, unless he also excludes all consumers with lower θ's, an action that may be unprofitable. In this situation, it is shown below that the optimal policy is to assign the same quality to a bunch of consumers in an interior subinterval of $[\theta^*, \bar{\theta}]$, and to equate marginal cost and marginal revenue elsewhere. The need to deal with this possibility accounts for much of the complexity of the following analysis.

The monopolist seeks to maximize profit

$$\Pi(q) = \int_{\underline{\theta}}^{\bar{\theta}} [\theta \cdot q(\theta) - z(\theta) - C(q(\theta))] \cdot f(\theta)\, d\theta$$

by choice of a nondecreasing assignment function $q(\theta)$, subject to constraint (7) and $z(\theta^*) = 0$. It is assumed that $q(\theta)$ is piecewise differentiable. If an assignment $q(\theta)$ is optimal then there cannot be any admissible deformation of $q(\theta)$ that increases profit. A deformation $h(\theta)$ is admissible if $q(\theta) + h(\theta)$ is piecewise differentiable and nondecreasing in θ. Clearly, if $h(\theta)$ is admissible then so is $\alpha \cdot h(\theta)$, $0 < \alpha \leqslant 1$, and $q(\theta)$ maximizes Π only if $\Pi(q + \alpha \cdot h) - \Pi(q) \leqslant 0$ for all admissible $h(\theta)$. Dividing by α and taking the limit as α approaches 0, gives the requirement that

$$\Lambda(h; q) \equiv \int_{\underline{\theta}}^{\bar{\theta}} \left[\theta \cdot h(\theta) - \int_{\underline{\theta}}^{\theta} h(s)\, ds - C'(q(\theta)) \cdot h(\theta)\right] \cdot f(\theta)\, d\theta \leqslant 0 \quad (9)$$

for all admissible $h(\theta)$. The problem is to use (9) to determine economically meaningful constraints on $q(\theta)$.

Consider the deformation $v(\theta; t; \Delta)$ that increases the quality allocated to consumers of type $\theta \geqslant t$ by a constant increment Δ, while leaving the quality allocated to consumers of type $\theta < t$ unchanged. Substituting into (9), the effect on profit is given by $\Lambda(v; q) = \Delta \cdot \mu(t)$, where

$$\begin{aligned} \mu(t) &= \int_t^{\bar{\theta}} \left[\theta - \int_t^{\theta} ds - C'(q(\theta))\right] \cdot f(\theta)\, d\theta \\ &= t \cdot (1 - F(t)) - \int_t^{\bar{\theta}} C'(q(\theta) \cdot f(\theta)\, d\theta \end{aligned} \tag{10}$$

measures the effect on the monopolist's profits of a unit increase in quality for all consumers of type $\theta \geqslant t$. Evidently $\mu(\bar{\theta}) = 0$. Differentiate $\mu(t)$ with respect to t and substitute (8):

$$\dot{\mu}(t) = [MC(t) - MR(t)] \cdot f(t), \tag{11}$$

where $MC(t) \equiv C'(q(t))$ is the marginal cost of producing an increment of quality for consumers of type t and $MR(t)$ is the marginal revenue defined above. Equation (11) and the boundary condition $\mu(\bar{\theta}) = 0$ imply an equivalent expression for $\mu(t)$, namely,

$$\mu(t) = \int_t^{\bar{\theta}} [MR(\theta) - MC(\theta)] \cdot f(\theta)\, d\theta. \tag{12}$$

Thus, the effect on the monopolist's profits of a unit increase in $q(\theta)$ for $\theta \geqslant t$, is the sum of the difference between the marginal revenue and the marginal cost of an increment assigned to all consumers for whom $\theta \geqslant t$, weighted by their density in the market. Since v is admissible, it follows that the monopolist's optimal quality allocation function $q(\theta)$ must be consistent with the condition

$$\mu(\theta) \leqslant 0 \qquad \text{for all } \theta. \tag{13}$$

With slight modification, the argument that proves that optimal $\mu(\theta)$ cannot be positive can be used to show that $\mu(\theta)$ cannot be negative, except at values of θ where $\dot{q}(\theta) = 0$. Specifically, if $\mu(t)$ is negative at some t, it appears that the monopolist can reduce costs by more than revenue by reducing $q(\theta)$ by an amount Δ for all $\theta > t$, thereby implying that $q(\theta)$ was not optimal. However, such a deformation would violate the restriction that $q(\theta)$ must be nondecreasing in θ, in the neighborhood of $\theta = t$. The constraint is satisfied by extending the deformation to some consumers with $\theta \leqslant t$. Specifically, define $S(t, \Delta)$ as the set of consumers for whom $q(\theta) < q(t+) - \Delta$, where $q(t+)$ is the right-hand limit of $q(\theta)$ at t. Consider the deformation

$$\begin{aligned} w(\theta; t, \Delta) &= -\Delta && \text{for } \theta > t, \\ &= q(t+) - \Delta - q(\theta) && \text{for } \theta \text{ in } S(t, \Delta), \\ &= 0 && \text{for all other } \theta. \end{aligned} \tag{14}$$

$q(\theta) + w(\theta; t, \Delta)$ is nondecreasing in θ; specifically, for θ in $S(t, \Delta)$, $q(\theta) + w(\theta; t, \Delta) = q(t+) - \Delta$. If $\dot{q}(t) > 0$, then $S(t, \Delta)$ is an interval of (approximate) length $\Delta/\dot{q}(t)$. Therefore,

$$\lim_{\Delta \to 0} (1/\Delta) \cdot \Lambda(w; q) = -\mu(t).$$

Since $w(\theta; t, \Delta)$ is admissible, (9) implies that $-\mu(t)$ must be $\leqslant 0$ at all t where $\dot{q}(t) > 0$. This and (13) imply that the optimal quality allocation function must be consistent with the condition

$$\mu(\theta) = 0 \qquad \text{whenever } \dot{q}(\theta) > 0. \tag{15}$$

Further, if $q(\theta)$ has a discontinuity (jump) at $\theta = t$, $S(t, \Delta)$ consists of at most one value of θ, namely t, for small enough values of Δ. It follows that $\mu(\theta)$ must also be zero at any θ where $q(\theta)$ is discontinuous.

This latter fact assists in establishing that jumps in $q(\theta)$ are not optimal. Suppose a jump occurs at θ^j. Let $\bar{q}(\theta^j)$ denote the limit of $q(\theta)$ as θ approaches θ^j from above and let $\underline{q}(\theta^j)$ denote the limit of $q(\theta)$ as θ approaches θ^j from below. Then at least one of two conditions must hold at θ^j: Either $MR(\theta^j) < C'(\bar{q}(\theta^j))$ or $MR(\theta^j) > C'(\bar{q}(\theta^j))$. However, either condition implies $\mu(\theta) > 0$ for some θ, which contradicts (13). To show this, note that

$$\begin{aligned} \mu(\theta) &= \int_\theta^{\bar{\theta}} [MR(s) - MC(s)]\, f(s)\, ds \\ &= \int_\theta^{\theta^j} [MR(s) - MC(s)]\, f(s)\, ds + \int_{\theta^j}^{\bar{\theta}} [MR(s) - MC(s)]\, f(s)\, ds \\ &= \int_\theta^{\theta^j} [MR(s) - MC(s)]\, f(s)\, ds, \end{aligned} \tag{16}$$

where the second equality follows from (12) and from the fact that $\mu(\theta^j)$ must be 0. Suppose $MR(\theta^j) < C'(\bar{q}(\theta))$. Then, since $MR(\theta)$ is continuous and $q(\theta)$ is nondecreasing, there must be some interval of θ just above θ^j where $MR(\theta) < MC(\theta)$. It follows from (16) that $\mu(\theta)$ is positive for θ in this interval, contradicting (13). If $MR(\theta^j) < C'(\underline{q}(\theta^j))$, the same logic leads to the conclusion that there must be some interval just below θ^j where $MR(\theta) > MC(\theta)$. For any θ in this interval (16) shows that $\mu(\theta)$ is positive, again contradicting (13).

The economic rationale for the conclusion that jumps in $q(\theta)$ are not optimal is that the monopolist would not be making full use of his power to discriminate among different types of buyers. Jumps in $q(\theta)$ imply corresponding "holes" in the spectrum of varieties offered. However, if there is a positive density of customers everywhere on $[\underline{\theta}, \bar{\theta}]$, such holes would represent an extreme and suboptimal market separation. Thus under the smoothness and continuity assumptions employed here, the spectrum of varieties appearing in the market is dense and continuous under both forms of market organization.[6] Since the monopolist achieves the optimal assignment by pursuing a corresponding price–quality schedule $P(q)$, the absence of jumps in $q(\theta)$ rules out "flats" in the optimal price–quality schedule $P(q)$.

Since $q(\theta)$ is nondecreasing and does not jump, it must be characterized by connected segments of two fundamental types: segments where $\dot{q} > 0$ and segments where $\dot{q} = 0$:

(i) In intervals where $\dot{q} > 0$, (15) applies and $\mu(\theta) = 0$. Therefore $\dot{\mu}(\theta) = 0$ and (11) requires that the optimal assignment equates the marginal cost and marginal revenue of increments of quality. Therefore $q(\theta) = G(\theta)$ in these intervals, where $G(\theta)$ is defined by

$$C'(G(\theta)) = MR(\theta). \tag{17}$$

(ii) There may be intervals $[\theta^u, \theta^v]$ where $\dot{q}(\theta) = 0$ and $\mu(\theta) < 0$. These intervals correspond to bunches of customers who are sold the same quality q^r. There are two possible subcases. First, if $q^r = 0$, then since $q(\theta)$ is nondecreasing, the bunch must be $[\underline{\theta}, \theta^*]$ and all customers for whom $\theta \leqslant \theta^*$ do not buy the good. The conditions of consumer behavior require $P'(0) = \theta^*$. Furthermore, since θ^* must occur at the lower limit of an interval where $\dot{q}(\theta) > 0$, the boundary condition for the extensive margin is $G(\theta^*) = \mu(\theta^*) = 0$. Second, there may also be a bunch where $q^r > 0$. If it is interior, that is, if $\theta^u > \underline{\theta}$ and $\theta^v < \bar{\theta}$ then its boundaries are determined by the requirement that $\mu(\theta^u) = \mu(\theta^v) = 0$ because its neighboring intervals have $\dot{q}(\theta) > 0$. However, it is possible that $\theta^u = \underline{\theta}$ and $\theta^v < \bar{\theta}$, with $q^r = G(\theta^v)$. Still $\mu(\theta^u) = \mu(\theta^v) = 0$, for if $\mu(\theta^u) < 0$, q^r could be reduced and profit increased. Further, q^r cannot exceed $G(\underline{\theta})$, for if it did there would be some θ in the neighborhood of $\underline{\theta}$ for which $q(\theta) > G(\theta)$. It follows from (12) and the boundary condition $\mu(\underline{\theta}) = \mu(\theta^u) = 0$ that $\mu(\theta)$ in such a neighborhood would be positive, violating optimality condition (13). Finally, it is required that $\theta^v < \bar{\theta}$, implying that the monopolist always differentiates among customers who have the largest valuations of quality. To prove this last

[6] Of course if $f(\theta)$ is not dense everywhere on the interior then there would be gaps in the spectrum observed in the competitive market. It is interesting to note that it may be worthwhile for the monopolist to fill in some of these holes and possibly eliminate them, as part of the general policy to reduce product quality.

assertion, suppose $\theta^v = \bar{\theta}$ and $q^r < G(\bar{\theta})$. Then for θ in the neighborhood of $\bar{\theta}$, $C'(q^r) < MR(\theta)$, and (12) implies that $\mu(\theta) > 0$, contradicting (13). On the other hand, if $\theta^v = \bar{\theta}$ and $q^r \geqslant G(\bar{\theta})$, then θ^u must be $\underline{\theta}$ since there exists no $\theta < \bar{\theta}$ for which $G(\theta) = G(\bar{\theta})$. Further, since $C'(q^r) \geqslant MR(\bar{\theta}) > MR(\theta)$ for $\theta < \bar{\theta}$, it follows from (12) that $\mu(\underline{\theta}) < 0$, implying that the monopolist could increase his profits by reducing the quality he sells to all consumers.

These conditions imply that the monopolist assigns a lower quality item to all consumers who purchase a good under competition, except to those for whom $\theta = \bar{\theta}$ where the assignment equating marginal cost and marginal revenue is the same as that which equates marginal costs and the incremental demand price. For proof, note that in intervals where $\dot{q} > 0$, the monopolist assigns $q(\theta) = G(\theta)$, where $C'(G(\theta)) = MR(\theta)$; whereas the competitive assignment $J(\theta)$ satisfies $C'(J(\theta)) = \theta > MR(\theta)$, except at $\bar{\theta}$ where $MR(\bar{\theta}) = \bar{\theta} = J(\bar{\theta})$. In intervals where $\dot{q} = 0$ and $q^r > 0$ the boundary conditions discussed above imply that $q(\theta) = q^r = q(\theta^u) \leqslant G(\theta^u) < J(\theta^u) \leqslant J(\theta)$ for $\theta^u \leqslant \theta \leqslant \theta^v$. Finally, it is clear from these results that consumers who are not served under competition are not served by the monopolist either.

5. Monopoly Solution: Examples

Two examples illustrate and clarify the nature of the optimal policy described in Section 4.

Let $C(q)$ be quadratic, $C'(q) = a + bq$, with $a \geqslant 0$ and $b > 0$, and suppose that $f(\theta)$ is uniform over $[\underline{\theta}, \bar{\theta}]$, with $\bar{\theta} > a$. $MR(\theta)$ is computed from (8) as $MR(\theta) = 2\theta - \bar{\theta}$, and from (17) $G(\theta) = [2\theta - (\bar{\theta} + a)]/b$, an increasing (linear) function of θ. However, $G(\theta)$ is defined only at θ's where $MR(\theta) \geqslant a$; i.e., for $\theta \geqslant \theta^*$, with $\theta^* = (\bar{\theta} + a)/2$. Letting $q^m(\theta)$ denote the monopolist's optimal assignment, then

$$q^m(\theta) = 0 \qquad \text{for } \theta \leqslant \theta^*,$$

$$q^m(\theta) = G(\theta) \qquad \text{for } \theta > \theta^*.$$

In this case $\dot{q}(\theta) = 0$ for $\theta < \theta^*$, and $\dot{q}(\theta) > 0$ elsewhere.

Since the conditions of consumer equilibrium require that $dP^m/dq = \theta$, it follows that $dP^m/dq = (b/2)q + (\bar{\theta} + a)/2$. In contrast, in the competitive solution, the price gradient is given by $dP^c/dq = a + bq$, and the quality allocation implied by it is $J(\theta) = (\theta - a)/b$ for $\theta > a$ and $q^c(\theta) = 0$ for $\theta \leqslant a$. Clearly $J(\bar{\theta}) = q^m(\bar{\theta}) = (\bar{\theta} - a)/b$ so that $\bar{\theta}$ customers buy the same good in either case. Also, $q^m(\theta) - J(\theta) = (\theta - \bar{\theta})/b < 0$ and $dP^m/dq \geqslant dP^c/dq$, for $\theta^* \leqslant \theta < \bar{\theta}$. The monopolist increases the price–quality gradient everywhere except at $q(\bar{\theta})$ and induces all customers for whom $\theta < \bar{\theta}$ and who would be served under competition to purchase smaller qualities. Some individuals who would buy an item under marginal cost pricing may not buy in the monopolistic market. For example, if $a = \underline{\theta} = 0$ the monopolist serves only those for whom $\theta > \bar{\theta}/2$, i.e., one-half of potential consumers, whereas the competitive market would serve all of them.

This example is well behaved because $G(\theta)$ is monotonically increasing over its domain. In general $G(\theta)$ is increasing in θ so long as $MR(\theta)$ increases in θ. However, the distribution of consumers may be such that $MR(\theta)$ is not increasing everywhere. This is an interesting case to analyze because it is the most striking illustration of the negative externalities that must be internalized by the optimal policy.

In Fig. 2, it is still assumed as in the uniform example that there is a unique θ^* at which $G(\theta) = 0$. However, $G(\theta)$ is not monotonically increasing for $\theta > \theta^*$; there is a wiggle in $G(\theta)$. Now the monopolist cannot allocate to each consumer $\theta > \theta^*$ the precise quality that equates $C'(q)$ to $MR(\theta)$, and satisfy $\dot{q} \geqslant 0$. Instead some group of consumers must be bunched, and the optimum assignment is

$$
\begin{aligned}
q(\theta) &= 0 && \text{for } \theta < \theta^*, \\
&= q^r && \text{for } \theta \text{ in } [\theta^u, \theta^v], \\
&= G(\theta) && \text{for } \theta \geqslant \theta^* \text{ and not in } [\theta^u, \theta^v].
\end{aligned}
$$

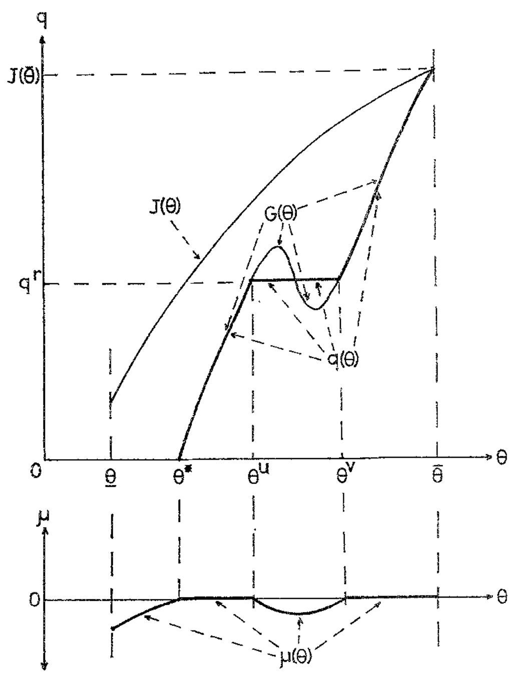

FIG. 2. Monopoly allocation with a "bunch" of consumers.

As indicated in Fig. 2, q^r, θ^u, and θ^v are chosen jointly so as to satisfy the boundary conditions discussed above:

$$G(\theta^u) = q^r = G(\theta^v), \tag{18}$$

$$\mu(\theta^u) - \mu(\theta^v) = \int_{\theta^u}^{\theta^v} [MR(\theta) - C'(q^r)] \cdot f(\theta)\, d\theta = 0. \tag{19}$$

Though the wiggle in $G(\theta)$ provides a range of possible choices of q^r (and of corresponding θ^u and θ^v) that satisfy (18) it is easily shown that the integral in (19) is a decreasing function of q^r. In fact it is positive for the lowest value of q^r consistent with (18) and negative for the highest value of q^r consistent with it, so that there is a unique choice of q^r (and θ^u and θ^v) that is consistent with both conditions. Outside this interval, and above θ^*, the optimal policy assigns customers the quality $G(\theta)$ that equates $C'(q)$ and $MR(\theta)$.

Whenever the monopolist sells the same positive quality q^r to a bunch of consumers $[\theta^u, \theta^v]$, condition (19) may be rewritten as

$$1/[F(\theta^v) - F(\theta^u)] \cdot \int_{\theta^u}^{\theta^v} MR(\theta) \cdot f(\theta)\, d\theta = C'(q^r).$$

The expression on the left-hand side is the average marginal revenue associated with the sale of an increment of quality to each consumer in the bunch. Hence, even though $MR(\theta)$ is not equal to $C'(q^r)$ for each consumer within the bunch, the average of marginal revenues is equal to marginal cost for the bunch as a whole. $MR(\theta)$ is not equated to $C'(q^r)$ for each individual consumer because $MR(\theta)$ does not fully reflect the extent to which the sale of an increment of quality to a consumer in $[\theta^u, \theta^v]$ interferes with the monopolist's ability to extract profit from consumers with more intense tastes: A change in $q(\theta^i)$ for some θ^i in $[\theta^u, \theta^v]$ requires the same change in $q(\theta)$ over a mass of customers in the interval and $MR(\theta)$ must be adjusted to reflect that fact.

6. Summary and Conclusion

We may now summarize the comparison of monopoly and competitive solutions. For this purpose, let a superscript "m" indicate a variable or function in the monopoly solution and a superscript "c" indicate the same variable or function in the competitive solution.

First, the highest θ consumer buys the same quality in both types of market organization: $q^{\mathrm{m}}(\bar{\theta}) = G(\bar{\theta}) = J(\bar{\theta}) = q^{\mathrm{c}}(\bar{\theta})$. For every other consumer (except those who do not buy under competition) the monopolist sells a lower quality than purchased under competition; $q^{\mathrm{m}}(\theta) < q^{\mathrm{c}}(\theta)$ for all $\theta < \bar{\theta}$ for which $J(\theta) > 0$. Indeed, some low-θ consumers who buy under competition may be priced out of the market by the monopolist. This situation is not unusual and arises because these customers impose the largest negative externalities on the monopolist's power to extract consumer's surplus from others. The effects may be so large that the marginal revenue from servicing low-θ consumers is negative.

Second, the loss of consumer surplus due to monopoly increases with θ and is greatest for type $\bar{\theta}$ customers who suffer no reduction in quality. From (7) and its boundary condition, consumer surplus is

$$z(\theta) = \int_{\underline{\theta}}^{\theta} q(s)\, ds.$$

$q^{\mathrm{m}}(s) < q^{\mathrm{c}}(s)$ for $s < \bar{\theta}$ for which $q^{\mathrm{c}}(s) > 0$, so that $z^{\mathrm{m}}(\theta) < z^{\mathrm{c}}(\theta)$ for all θ for which $z^{\mathrm{c}}(\theta) > 0$. Further, since $dz/d\theta = q(\theta)$, it follows that $d(z^{\mathrm{c}}-z^{\mathrm{m}})/d\theta = q^{\mathrm{c}} - q^{\mathrm{m}} > 0$.

Third, for the set of consumers $[\theta^*, \bar{\theta}]$ served by the monopolist, the range of quality is always greater than the range of quality sold to these consumers under competition. Moreover, unless the minimum feasible quality is sold under competition, the monopolist will increase the total quality range relative to competition, with the broadening occurring at the low quality end. This occurs as part and parcel of the monopolists' power to use both prices and qualities to discriminate the most against customers with more intense demands.

Fourth, for any quality $q^h > 0$ sold both by the monopolist and under competition, the monopoly price is larger than the competitive price, and the price differential increases with q^h. Since the competitive price is equal to the unit cost of producing the given quality, $C(q^h)$, the preceding statement also characterizes the relationship between the monopoly price and unit cost. The first part follows from the general loss in consumer surplus under monopoly pricing and the second part from the fact that the reduction in surplus increases with θ.

Fifth, in the ordinary theory of monopoly, it is shown that the monopolist reduces quantity, relative to marginal cost pricing. In the present situation, it is not true, in general, that the monopolist reduces the quantity of any given quality which is sold to consumers. The monopolist may sell some qualities which do not even appear under competition. Further, if a bunching occurs there is at least one quality which the monopolist sells to a finite mass of consumers, something that does not happen under competition. However, there is a general sense in which the monopolist does reduce quantity. Specifically, consider any quality q^h sold under both monopoly and competition. Suppose that $q^{\mathrm{c}}(\theta^j) = q^h$ and that $\theta^i > \theta^j$ is the smallest θ for which $q^{\mathrm{m}}(\theta^i) = q^h$. The number of units of quality q^h or better sold under monopoly, $1 - F(\theta^i)$, is smaller than the number, $1 - F(\theta^j)$, sold under competition.

One interesting application of the present model is to the question of depreciation rates for durable goods under alternative market organizations. The leading result (Swan, 1970) that durability should be the same under monopoly or competition is unquestionably correct, given the assumptions of perfect capital markets and no transactions costs. All that matters to a consumer under those circumstances is the cost of achieving a given level of service flow. In our terminology, if T is the length of service life of a particular variety, q may be defined as $q = (1 - \exp(rT))/r$, where r is the rate of interest common to all consumers. Therefore, all consumers will have the same value of θ and $\theta = \bar{\theta}$. Evidently only one quality will be sold under both monopoly and competition and it will be the quality at which marginal cost equals marginal value. As pointed out by Parks (1974), if traveling to the stores takes time and the distance or the value of time varies among consumers, or if capital markets are imperfect and rates of interest differ among them, then the competitive and monopolistic outcomes are generally different. This situation is the essence of our model. Different marketing or interest costs provoke a distribution of attribute valuation comparable to $f(\theta)$ that allows the monopolist to engage in a form of price discrimination.

One possible extension of the model developed here is to treat the case where individual consumers decide both on the quality of units they will consume and on the quantity of such units. Intuition suggests that if quantity and quality both enter the utility function, some of the strong conclusions reached here may no longer hold because additional constraints arising out of quantity–quality substitution must be considered. Another possible extension is to treat the case where quality has more than one dimension. If consumer preferences are similarly ordered in all dimensions, virtually all of the present results should continue to hold. If tastes follow different orders in different dimensions, then covariances in valuations between attributes will be important. However, these same factors count in both the monopoly and competitive outcomes, and there would appear to be little reason for altering the general conclusion that the monopolist tends to move consumers toward the origin.

More general assumptions about cost provide another possible area for extensions of the present analysis. It seems plausible that there should be production externalities associated with qualities that are close together, though it is not clear precisely how to model them. Further, there could be a fixed cost associated with the production of any quality. Then the quality spectrum would collapse to a number of discrete points. However, the non-convexities introduced by fixed costs create well-known difficulties for the competitive solution that are beyond the scope of this paper. Nevertheless, the constant cost assumption allows us to focus on how the monopolist uses the price system itself to sort and stratify different types of customers among differentiated products. Surely this feature would appear in the decreasing cost case as well.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation. We would like to thank the referees and an associate editor for their helpful suggestions.

REFERENCES

1. R. G. D. ALLEN, "Mathematical Analysis for Economists," Chap. 14, Macmillan & Co. London, 1938.
2. H. HOTELLING, Edgeworth's taxation paradox and the nature of demand and supply functions, *J. Polit. Econ.* (1932), 577–616.
3. K. LANCASTER, Socially optimum product differentiation, *Amer. Econ. Rev.*, 1975.
4. E. MILLS, An aggregative model of resource allocation in a metropolitan area, *Amer. Econ. Rev.* **52**, No. 2 (1967), 197–210.
5. J. MIRLEES, An exploration in the theory of optimum income taxation, *Rev. Econ.* Studies **38** (1971), 175–208.
6. R. PARKS, The demand and supply of durable goods and durability, *Amer. Econ. Rev.* **64** (1974), 37–55.
7. S. ROSEN, Hedonic prices and implicit markets: Product differentiation in pure competition, *J. Polit. Econ.*, 1974.
8. MICHAEL SPENCE, Nonlinear prices and economic welfare, *J. Publ. Econ.*, 1977.
9. P. SWAN, Durability of consumption goods, *Amer. Econ. Rev.* **60** (1970), 884–894.
10. C. WILSON, "A Model of Insurance Markets," Ph.D. dissertation, University of Rochester, 1975.