Rand Journal of Economics
Vol. 16, No. 3, Autumn 1985

Price discrimination in free-entry markets

Severin Borenstein*

Using a spatial model of monopolistic competition, we investigate price discrimination in free-entry, zero-profit markets. We show that when brands are heterogeneous, competition does not prevent discrimination. The power to earn economic profits is not necessary for a firm to maintain discriminatory prices. Our model treats formally the fact that consumers differ not only in the utility they derive from a good, but also in how strongly they prefer one brand over all others. In markets where firms are very competitive, sorting consumers on the strength on brand preference produces larger price differentials between groups than sorting on the basis of consumers' reservation prices for the good. When firms sort customers on the basis of strength of brand preference, however, we find that the output and welfare effects are generally less favorable than when sorting is more closely related to consumers' reservation prices.

1. Introduction

■ Since Pigou (1932) and Robinson (1933), economists have associated price discrimination with monopoly or oligopoly markets. Yet, much pricing that appears to be discriminatory occurs in markets where entry and exit are commonplace and there is little coordination among firms. Magazine and journal subscriptions are sold at reduced rates to students; hotels let "kids stay free," yet charge a considerable fee for extra adults; drugstores advertise discounts on prescriptions for customers 59 years or older; airlines charge lower fares to passengers whose trip lasts more than seven days.

Despite the relative ease of entry and apparent rarity of long-run economic profits in these industries, it should not surprise economists to see performance differ from the perfectly competitive ideal. Product homogeneity, a necessary condition for perfect competition, is not a realistic approximation in these markets. Chamberlin (1931) demonstrated that heterogeneity in free-entry markets can lead to a persistent differential between price and marginal cost. In this article we see that heterogeneity can also lead to persistent price discrimination.

Using a spatial model of monopolistic competition, we examine third-degree price discrimination in free-entry, zero-profit markets. Consumers are sorted into two groups according to criteria that are imperfectly correlated with willingness-to-pay. If firms have access to a usable, though perhaps noisy, signal of consumers' willingness-to-pay for their

---

* University of Michigan.

This article is based on my dissertation. Financial support for this research was provided by the Sloan Foundation. I would like to thank Richard Schmalensee, Mark Bagnoli, Nicholas Economides, Paul Joskow, Garth Saloner, Steve Salop, and two anonymous referees for helpful comments. Any remaining errors are my own.

brand, we find that equilibrium prices will almost certainly be discriminatory.[1] The primary barrier to discrimination, it thus appears, is the absence of a usable sorting mechanism. Competition among heterogeneous brands and the absence of entry barriers will almost never prevent price discrimination, even when they cause long-run profits to be driven to zero. In fact, when a usable sorting mechanism exists, a firm could be forced to discriminate to avoid losses when competing with other discriminating firms.

Having established the possibility of price discrimination in free-entry markets, we focus on two broad questions. First, how do equilibria when discrimination is prohibited differ from those when discrimination is allowed? What effect does the change in price structure have on the quantities sold, the equilibrium number of brands, welfare measures, and price levels? Second, how do changes in the sorting mechanism affect the discriminatory equilibria? In particular, what factors increase or decrease the price differential between groups? We find that extrapolations from theories of monopoly price discrimination ignore an important basis for sorting customers in heterogeneous product markets: strength of preference among brands.

In monopoly markets we often observe value-of-service pricing with the firm's charging higher prices to consumers who receive higher utility from the product.[2] When differentiated brands exist, however, there is another basis for sorting buyers. People with very strong preferences among brands can be separated from those who are not so particular. We conclude that in competitive heterogeneous product markets, sorting based on strength of brand preference is often more effective than sorting based on the value consumers place on the basic product; in many cases the former leads to larger price differentials between the sorted groups. This concept of brand-preference sorting, previously discussed by Frank (1983), is, of course, absent in traditional models of monopoly price discrimination.

Section 2 presents a circular characteristic-space model of monopolistic competition based on the seminal work of Lerner and Singer (1937) and a more recent article by Salop (1979). Salop's model is generalized in a way conducive to considering price discrimination. Consumers are assumed to differ not only in the location of their most preferred point in characteristic space, but also in the basic utility they derive from the good (measured as reservation price for a hypothetical brand at the consumer's most preferred point in characteristic space) and in their strength of preference among brands (measured as the decline in reservation price, per unit of distance, for a brand away from the consumer's most preferred point). We compare sorting of consumers on the basis of their reservation prices with sorting based on the strength of their brand preference. We do not study sorting based on the location of the consumer's most preferred point in characteristic space. There are two reasons for this. First, there are already many studies of such spatial discrimination (see, for instance, Greenhut and Ohta (1975) and Norman (1981)). Second, they do not seem to apply well to the consumer markets we are particularly interested in explaining. While in input markets the seller may know the geographic location of the buyer, in consumer markets a firm is unlikely to have any reliable signal of which customers are located close to it in geographic or characteristic space.[3] Signals of the customer's general flexibility (strength of brand preference) or value of the product (reservation price) are more available and seem to be more commonly used in consumer markets.

Unfortunately, the generalized spatial model is not tractable. Still, we make some analytical observations in Section 3 by examining the effect that a group of consumers' res-

---

[1] Usable, in this context, means that the standard problems with resale between groups, costs of imposing the sorting mechanism, or legal barriers are not so great as to prevent profitable use of the sorting mechanism.

[2] In this article "product" refers to the good produced by all firms in the market, and "brand" refers to a particular firm's output.

[3] It would be quite difficult, for instance, for an airline to determine which passengers on a 6 P.M. flight wanted to leave at precisely that time and which would have preferred a 5 P.M. departure.

ervation prices and strengths of brand preference have on their price elasticity of demand. One can find situations in which sorting does not yield demand functions with different elasticities and, thus, results in no price discrimination.

Section 4 explains the problems of tractability and outlines the structure of the computer simulations used to find equilibria for most parameter values. It describes briefly the computational procedure used to find zero-profit symmetric Nash-in-price (Bertrand) equilibria both when there is one only price and when consumers are sorted and discrimination is allowed. We also consider appropriate models of sorting in this section. Because the ways in which firms can effectively segment the market are limited, we do not model a profit-maximizing sorting mechanism. Rather, we look at different exogenously determined binary groupings of the population and their resulting equilibria. Thus, the results do not reveal which groupings firms would choose (though they do indicate which of the sorting mechanisms, *if usable,* would be more profitable in the short run). Instead, the simulations show the effects a certain kind of sorting would have if it were chosen.

We discuss the results of these simulations in Section 5. We see that both sorting on reservation price and sorting on strength of brand preference can lead to persistent price discrimination. The two approaches, however, differ significantly in their effects on the quantities sold, the number of brands, and welfare.

In Section 6 we extend the model to examine the effect of sorting by self-selection. A firm offers two prices, but places restrictions on the low price that make it costly for customers to obtain the discount. Minimum-stay requirements for discount air fares are an example of this behavior. Sellers design the restrictions so that the cost of qualifying for the discount is positively correlated with a buyer's willingness-to-pay for the brand. Usually, the cost is inconvenience, such as rearranging travel plans or driving out to the company warehouse. Sorting criteria that are not self-selective, such as student and senior citizen discounts, are called "index" sorting in this article. The results of self-selective sorting are similar to those from index sorting. The primary difference seems to be that self-selection dampens the effectiveness of discrimination: changes in the number of firms, the quantities sold, and welfare that result from price discrimination are smaller with self-selection than with index sorting.

2. Characteristic-space model of monopolistic competition

■ In our model brand characteristics and consumer preferences over those characteristics are assumed to differ in only one dimension. Thus, we can represent differentiated brands and preferences in a one-dimensional space, in this case as points on a unit-circumference circle. We assume that the N brands in the market are evenly spaced so that the arc distance between brands is $1/N$. All brands have the same costs of production regardless of location on the circle, a positive fixed cost, F, plus constant marginal cost, m:

$$\text{total cost} = F + mq. \tag{1}$$

This implies economies of scale in production of a brand. There are, however, no economies of scope: costs of producing different brands are entirely separate. The absence of scope economies allows us to treat each brand as a separate firm. On the other hand, declining average costs prevent an equilibrium with an infinite number of firms, together providing each customer with exactly the product he would most prefer.

A consumer is completely described by the triple (z_i, A_i, c_i), where:

z_i = consumer i's most preferred point on the characteristic circle;
A_i = his reservation price for a hypothetical brand located at z_i; and
c_i = the decline in his reservation price for a brand per unit of arc distance the brand is from z_i.

A "type" of consumer is defined to be a group of all potential buyers with the same A and c values. We assume that the preferred points, z's, of the consumers of any (A, c) type are distributed uniformly around the circle. This definition is useful because we assume that sorting of consumers is based on their reservation prices (A's) and/or strengths of brand preference (c's). Thus, all consumers of a given type are treated symmetrically in the sorting process.

Without loss of generality, we can restrict each consumer to a single purchase/no purchase decision.[4] If at least one brand yields positive consumer surplus for him, the individual buys one unit of the brand that gives him the highest surplus. If he would receive negative (or zero) surplus from all existing brands, he buys none. The consumer surplus buyer i receives from brand X is given by

$$CS_i = A_i - P_x - c_i|z_i - X|, \tag{2}$$

where P_x is the price of brand X and $|z_i - X|$ is the arc distance from z_i to X.

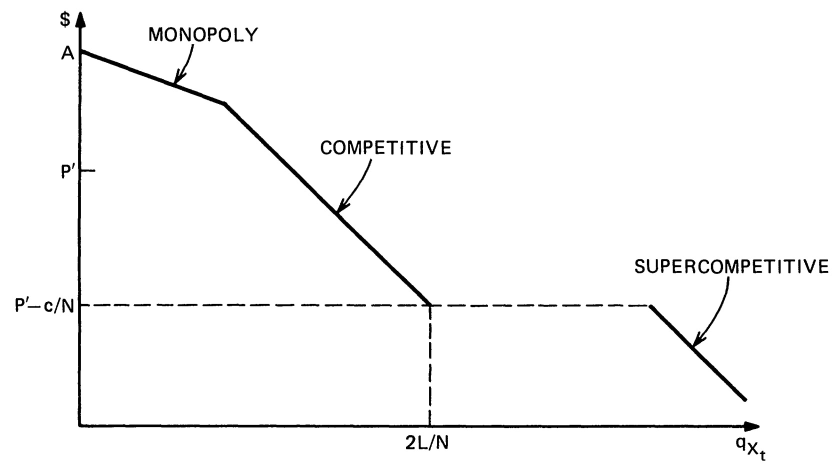
□ **Demand from a single (A, c) type.** Competitive interactions in this model are most easily demonstrated by focusing on the consumers of a single (A, c) type. Figure 1 shows the demand of a given type for brand X when each of the other $N - 1$ brands is priced at P'. When the price of brand X, P_x, is high, the only people in the type who buy it are consumers who are very close to brand X.[5] So long as P_x is in this region, brand X is essentially a monopoly. A consumer who is located at the edge of its market receives zero surplus from brand X and negative surplus from any other brand. This implies that there are "gaps" between markets. Consumers with z_i's in these characteristic-space gaps do not purchase the good at all.

We determine the quantity of brand X sold to consumers in this type, q_{x_t}, by calculating the market boundaries for the type. With L_t consumers in the type,

$$q_{x_t} = L_t \times \text{the arc distance between firm } X\text{'s market boundaries,} \tag{3}$$

FIGURE 1

DEMAND OF A GIVEN TYPE FOR BRAND X WITH OTHER BRANDS PRICED AT P′



[4] A simple generalization is to assume that one individual is the source of many of the (z_i, A_i, c_i) triples. With discrimination, however, this implies that resale between groups can be prevented, even if the same person is sometimes sorted into one of the groups and sometimes into the other.

[5] Throughout this article a customer's "location" refers to z_i, his most-preferred point in product space.

since the consumers in the type are uniformly distributed around the unit-circumference circle. In the monopoly region (Figure 2a) we can calculate the distance, d, from the brand X location to either boundary from the zero-surplus condition for buyers at the monopoly market boundary:

$$A - P_x - cd = 0 \quad \Rightarrow \quad d = (A - P_x)/c, \tag{4}$$

$$q_{x_t} = L_t 2d = 2L_t(A - P_x)/c. \tag{5}$$

As P_x declines, brand X's market boundaries eventually meet the boundaries of neighboring brands. The brands are competitive below this price in that marginal customers are choosing between brand X and a neighboring brand, rather than between brand X and no purchase. The P_x at which the switch to competition occurs is the one that causes the monopoly markets of brand X and its neighbors exactly to fill the space between brands (Figure 2b):[6]

$$(A - P_x)/c + (A - P')/c = 1/N \tag{6}$$

$$P_x = 2A - P' - c/N. \tag{6a}$$

In this competitive region all customers with z_i's between brand X and its neighbors receive positive surplus from at least one brand, and therefore make a purchase. The demand function in this region is again calculated from the market boundaries of brand X (Figure 2c). A customer on the boundary between brand X and a neighboring brand receives equal surplus from the two brands:

$$A - P_x - cd = A - P' - c(1/N - d) \tag{7}$$

$$d = (P' - P_x + c/N)/2c. \tag{7a}$$

Therefore, we obtain[7]

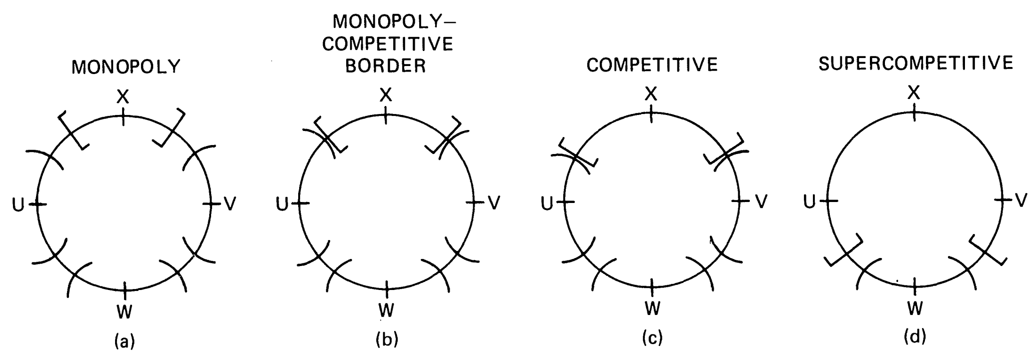
$$q_{x_t} = 2L_t d = L_t(1/N + (P' - P_x)/c). \tag{8}$$

Finally, if P_x drops far enough that

$$P_x < P' - c(1/N), \tag{9}$$

FIGURE 2

MARKET BOUNDARIES



[6] The kink at the monopoly-competitive switch point can lie above or below P', depending on whether neighboring firms' distances to their monopoly borders, $(A - P')/c$, are greater or less than half the distance between firms.

[7] One implication of equation (8) is that if $P_x = P'$, then $q_{x_t} = L_t/N$. That is, if all brands sell for the same price and the price is in the competitive region, then all customers in the type purchase the good, and each brand gets $1/N$ of the market sales.

then even consumers whose preferred point is exactly at the location of the neighboring brand will get greater consumer surplus from brand X. So will all the customers on the far side of the neighboring brand who had been buying it. This is the point of discontinuity in Figure 1 and the jump to the supercompetitive region (Figure 2d).[8] Similarly, sales of brand X will be zero if $P_x \geq P' + c/N$.

□ **Demand from all (A, c) types.** When there are many types of consumers with differing reservation prices and strengths of brand preference, service to all consumers will not lie in the same region of the Figure 1 demand curve. Figure 3 illustrates the simplest case, when firm X sets its price equal to the price charged by all other firms, $P_x = P'$. Types with reservation prices less than this price (e.g., type R) fall in the no-purchase region; none of these people buys the product, even if there is a brand located at exactly z_i, the person's most preferred point in product space.

Types with slightly higher A's, but whose reservation prices are still low relative to their strengths of brand preference, are served monopolistically (e.g., type S). The consumers in these types who are located close to a firm buy the product. For others of the same type, however, no brand is close enough to their preferred point to give positive surplus at the going price.[9]

Types with large A values relative to c are served competitively (e.g., type T). Everyone in these types buys the good. Equation (6), which defines the monopoly-competitive border, can be rewritten as

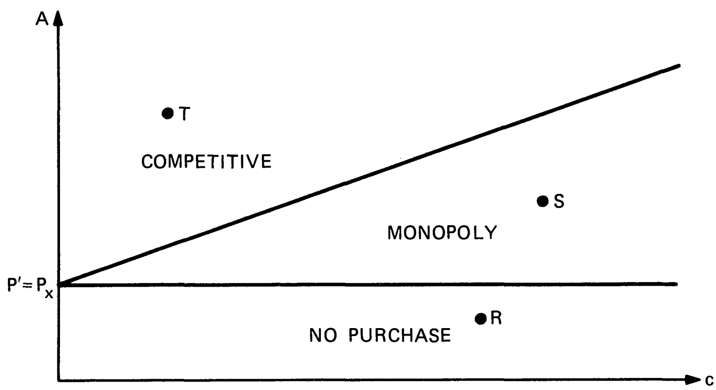
$$A = (P_x + P' + c/N)/2. \tag{6b}$$

The slope of the border is $1/(2N)$. As N increases, the monopoly region shrinks. This reflects the fact that a characteristic space more crowded with brands will satisfy more people's tastes.

We find the quantity sold by firm X (and all other firms in this symmetric case) by integrating the monopoly and competitive demand functions (5) and (8) multiplied by the population density, $f(A, c)$, over the appropriate regions of the (A, c) space. With L consumers in the population,[10]

FIGURE 3

DEMAND REGIONS WHEN FIRM X SETS PRICE EQUAL TO THE PRICE CHARGED BY ALL OTHER FIRMS



[8] There is another jump at $P_x = P' - 2c/N$ as firm X also takes over the market of brands adjacent to its neighbors. Another jump exists at $P_x = P' - 3c/N$, etc.

[9] The proportion of a monopolistically served type that falls into these gaps between markets depends on the (A, c) pair. Nearly everyone in a type near the monopoly-competitive border buys the good, while in a type near the no-purchase line only those very close to a firm in characteristic space will buy.

[10] The demand functions faced by firms in this model are well behaved (i.e., continuous and smooth) under reasonably general conditions. It is necessary, however, that $0 \leq c < t$, which means that $f(A, c) = 0 \; \forall \; A$ for some arbitrarily small, positive t. If there were individuals with $c = 0$, they could be won away from competitors by

$$q_x(P_x|P', N, f(A, c))|_{P_x=P'}$$

$$= 2L \int_0^\infty \int_{P_x}^{P_x+c/2N} \frac{(A - P_x)}{c} f(A, c)dAdc + L \int_0^\infty \int_{P_x+c/2N}^\infty \frac{1}{N} f(A, c)dAdc. \quad (10)$$

Calculation of the quantity sold when P_x is not equal to P' may be complicated by supercompetition for some types. From (9) we can rewrite the supercompetitive condition when $P_x > P'$ as

$$c < N(P_x - P'). \quad (9a)$$

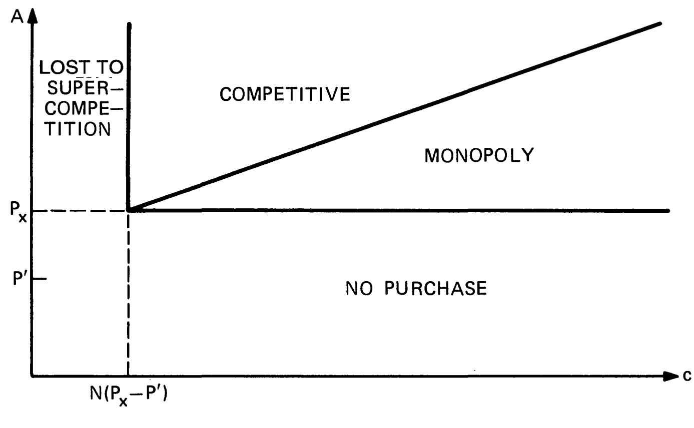
Any consumer whose strength of brand preference, c_i, satisfies this condition would prefer a neighboring brand over brand X, even if the person's preferred point, z_i, is exactly at the location of brand X. It is clear that for types with very weak preferences among brands (low c's), one brand will dominate a neighboring brand if there is even a small price differential. Capturing the strong-preference customers of a neighbor requires a larger differential.

Figure 4 illustrates the role of the supercompetitive effect. It shows the change in regions from Figure 3 as P_x increases while P' is unchanged. Besides causing firm X to lose all customers whose reservation prices are between P' and P_x, the price increase also causes firm X to lose all sales to types that have weak preferences among brands. Furthermore, equations (5) and (8) show that smaller proportions of the people in each type remaining in the monopoly and competitive regions buy brand X as P_x increases.

In a symmetric equilibrium, however, the supercompetitive region vanishes because the prices of all brands are the same. In fact, so long as the proportion of the population with low c values is relatively small, supercompetition plays no role in the equilibrium analysis.[11] Still, the supercompetitive effect highlights the importance of strength of brand preference in the price elasticity of the demand for a given brand. This is also clear from equation (8), the one-type competitive quantity equation. Thus, one should expect that firms that price discriminate might try to sort out individuals with weak brand preferences (low c types) so price cuts could be targeted at them.

FIGURE 4

ROLE OF THE SUPERCOMPETITIVE EFFECT



charging a price slightly less than any other firm. In this case no firm would ever be content to charge the same price as all others; no symmetric pure strategy Bertrand equilibrium would exist. The existence and uniqueness of an equilibrium are discussed in Appendix A of Borenstein (1983), which is available from the author.

[11] More precisely, a sufficient condition for supercompetition to be insignificant in the equilibrium analysis is that sales do not increase so rapidly by charging a $P_x < P'$ that the marginal revenue curve slopes upward (or jumps discontinuously) to a level above marginal cost (see footnote 10). We assume this to be the case in the theoretical analysis and tested for it successfully in all of the simulations.

The demand functions faced by each of the firms along with their cost functions described above provide sufficient information to solve (numerically) for the symmetric one-price Bertrand equilibrium. To examine the effects of price discrimination, we also find the symmetric Bertrand equilibrium when every firm sorts the population into the same two groups and is allowed to charge different prices to the groups. In either case the number of firms can be adjusted so that all firms earn zero profits in the equilibrium. These are the one-price and two-price, zero-profit symmetric Bertrand equilibria that are compared in the following sections.[12]

3. Analytical results of the model

■ Unfortunately, for reasons discussed in Section 4, there is no general closed-form solution for the symmetric Bertrand equilibria in this model. This is the case whether or not firms are allowed to price discriminate. Still, there are some useful observations that can be made analytically. This section discusses the relationship between the (A, c) values of different types of consumers and their price elasticities of demand. From this analysis one can identify the extreme cases under which sorting consumers on the basis of either their reservation prices or their strengths of brand preference will be completely ineffective. That is, even after two distinct groups of consumers have been identified through some mechanism, they will still be charged the same price. This gives some intuition for the factors that affect the equilibria in the less extreme cases that were simulated.

The two groups that result from a single binary sorting device will be charged different prices only if their total demand curves display different price elasticities at the no-discrimination equilibrium price. If the demand elasticities of the two groups are the same, then the standard first-order condition for profit maximization,

$$\frac{P - MC}{P} = -\frac{1}{\epsilon},$$

indicates that the groups will be charged the same price. In perfect competition firms see each consumer's demand as infinitely elastic, so price is equal to marginal cost and the same for every buyer. In this free-entry spatial model, however, firms do not face infinitely elastic demand from any type of consumer. For discrimination to persist, then, it need only be the case that consumers can be sorted so that the resulting groups' demand curves have different elasticities at any given price.

From the demand equations derived in Section 2, we can calculate the elasticities for a single (A, c) type of consumers. For a type in the competitive region of its demand function, (8) gives us

$$\epsilon_{comp} = -\frac{P_x}{(P' - P_x + c/N)}. \tag{11}$$

The demand elasticity of a type served competitively depends on the strength of brand preference, but is independent of reservation price. Thus, if sorting takes place on A and *all* consumer types are served competitively before sorting, then the prices for the sorted groups will differ only if A is correlated with c. If A is uncorrelated with c, the sorting will have no effect since the two groups will have the same demand elasticity.

If sorting is based on c directly, or indirectly thorough correlation between A and c, we can also infer the direction of the price difference between sorted groups. The change in demand elasticity when strength of brand preference changes,

---

[12] These equilibria exist only if firms play a "location-first, price-second" game. If a firm can change location and price simultaneously, then any symmetric equilibrium can be disrupted by a firm's moving to its neighbor's location and charging a price slightly less than its neighbor. See Economides (1984).

$$\frac{d\epsilon_{comp}}{dc} = \frac{NP_x}{(P' - P_x + c/N)^2} > 0, \tag{12}$$

indicates that the demand elasticity will be smaller in absolute value from a group with generally stronger brand preferences. This would lead to a higher price's being charged to the group with stronger brand preferences.

The opposite bias in sorting would occur if *all* consumers were served monopolistically. For a type served monopolistically, (5) gives us

$$\epsilon_{monop} = -\frac{P_x}{A - P_x}. \tag{13}$$

In this case the elasticity of a type's demand is independent of its strength of brand preference. A criterion that sorts consumers on the basis of strength of brand preference will lead to different prices for the sorted groups only if c is correlated with A. Furthermore, differentiating ϵ_{monop} with respect to A,

$$\frac{d\epsilon_{monop}}{dA} = \frac{P_x}{(A - P_x)^2} > 0, \tag{14}$$

shows that elasticity of demand will be larger in absolute value at any given price for types with a smaller reservation price. Thus, if all consumers were served monopolistically, the profit-maximizing price would be higher for sales to the group with higher reservation prices.[13]

The fundamental difference between the monopoly and competitive regions is the choice facing the marginal consumers of a brand. In the competitive region a brand's marginal customers are just indifferent between two brands, both of which give positive surplus. For these types, responses to (infinitesimal) changes in one brand's price come entirely from interbrand switching.[14] Total market sales to these types are unchanged. On the other hand, for types in the monopoly region, changes in the quantity purchased of a given brand come entirely from changes in total market sales. Price increases lead some of the people in each monopolistically served type to cease purchasing any brand of the product, while price decreases expand total market sales to these types.

Although the mathematics we present are specific to the extreme cases of our model, the distinction between monopolistic and competitive service is more general. When any firm lowers the price for its product, for instance, there are two sources of changes in the quantity sold:

$$\frac{dq_x}{dP_x} = \text{increase in total market sales} + \text{sales that switch from competing brands.}$$

In regulated industries the first term is often referred to as "demand creation" and the second term is called "demand diversion." In the competitive region price changes result primarily in demand diversion. When a firm is trying to gain customers who are buying competing brands, or is concerned with losing customers to competing brands, reservation price is of little interest in pinpointing who those "vulnerable" consumers are. Thus, sorting on A is ineffective. The buyers most likely to switch brands are ones with weak brand preferences (low c values). If these consumers can be identified, competition for them will be greater, and the price they pay will be lower.

[13] Only with the restriction that there be an integer number of firms can all consumers be served monopolistically in a free-entry equilibrium. This would mean that there would be gaps between firms' largest markets. Except for the integer constraint or the zero-probability event that such firms are just breaking even, the gaps would bring new entry until firms' markets for some types were constrained by neighboring brands.

[14] Finite price changes require consideration of movements of some consumer types among the monopoly, competitive, and no-purchase regions. A consumer type could, for instance, go from the competitive region to the monopoly region and, if the price change was not infinitesimal, some in the type would stop buying.

In the monopoly region, on the other hand, all changes in the quantity sold come from changes in the total market sales; there is no diversion when consumers get positive surplus from only one brand. If a brand's marginal customers are deciding whether to purchase the product at all, we can expect their reservation prices to play a key role. By the same reasoning, since these consumers would derive negative surplus from all other brands in the market, their strengths of brand preference are irrelevant.

This intuitive argument suggests that the results from the extreme cases will carry over to cases in which most, but not all, consumers are in either the competitive or the monopoly region. When the no-discrimination equilibrium causes most people to be in the competitive region, for instance, sorting on strength of brand preference will tend to be more effective than sorting on reservation price. Such extensions to less extreme cases are supported by the simulation results.

4. Method of computer simulations

■ We now outline the method of computer simulation we used to find symmetric equilibria in the spatial model; details appear in the Appendix. Simulations were necessary to find equilibria, because the assumption of a continuous bivariate (nonuniform) distribution of A and c causes the model to be mathematically intractable (see equation (10)). When we assume a discrete number of types (or a uniform distribution), however, multiple kinks in the demand curve plague the analysis (e.g., Salop, 1979). Besides the loss of realism when we assume only a few types, the kinks induce so many pathological cases that the analysis ceases to be very informative.

In all of the simulations we assumed a bivariate normal distribution of A and c. In addition to the simplicity and familiarity of the normal density, we used this distribution because it allows one to study separately changes in the population means and variances.[15] For each simulation, we specified the parameters of the density function, the population size, and starting values for the prices and the number of firms.[16]

For the simulations in which firms were allowed to price discriminate, we also had to specify a sorting criterion that divided the population into two groups. The results presented for index sorting are based on simple divisions of the population according to whether a consumer's reservation price, A, or strength of brand preference, c, is greater or less than some critical value. For the simulations discussed in Section 5, the critical c value was the median of the distribution. The critical A value was the median of the population that had reservation prices greater than marginal cost, m $(=0)$.[17] More sophisticated index sorting criteria that take account of both A and c simultaneously and have varying degrees of noise in the sorting yield results completely consistent with those in Section 5.[18]

For each set of parameters we found the discrimination-allowed and discrimination-

---

[15] To avoid the problems with equilibrium nonexistence discussed in footnote 10, μ_c was always set at least three standard deviations above zero.

[16] All simulations presented here assume that a firm's conjectural variation in price was zero, a Bertrand assumption. Values from 0 to 1 were also simulated. Though values greater than 0 increased equilibrium prices and the number of firms, they had no important qualitative effect on the results discussed in Sections 5 and 6. We did not, however, vary the parameters of the cost function for the simulations. In the Appendix we show that the cost function parameters, F and m, could be held constant for all simulations without loss of generality. Fixed costs affect the results only through changes in the ratio of population to fixed costs. Marginal cost affects the outcome through the difference between m and reservation prices. Thus, a change in m is equivalent to a shift of the distribution along the A-axis.

[17] Simulations with critical values that divided the population less evenly tended to yield greater price differentials, but otherwise were qualitatively similar to those presented.

[18] See Borenstein (1984). As the noise in the sorting increases, for instance, the groups differ less in their (A, c) distributions, and the effects of discrimination lessen.

prohibited zero-profit symmetric equilibria. The output included the equilibrium prices, number of firms, the quantities sold, and total consumer surplus.[19]

5. Results from simulations of the model

■ To assess the effect of price discrimination on the basis of index-sorting criteria, we look at the impact that sorting has on the equilibrium price differential between groups, the number of brands (firms), the quantities sold to each group, and the total surplus of consumers in each group. A useful basis for comparison is the equilibrium that results with no discrimination, as would be the case if price discrimination were prohibited.

The price differential between groups is an obvious measure of the degree to which discrimination is actually occurring.[20] Because scaling alone can affect the absolute difference between a firm's high and low price, $P_H - P_L$, we look at changes in the ratio P_L/P_H or, equivalently,

$$\theta = 1 - (P_L/P_H).$$

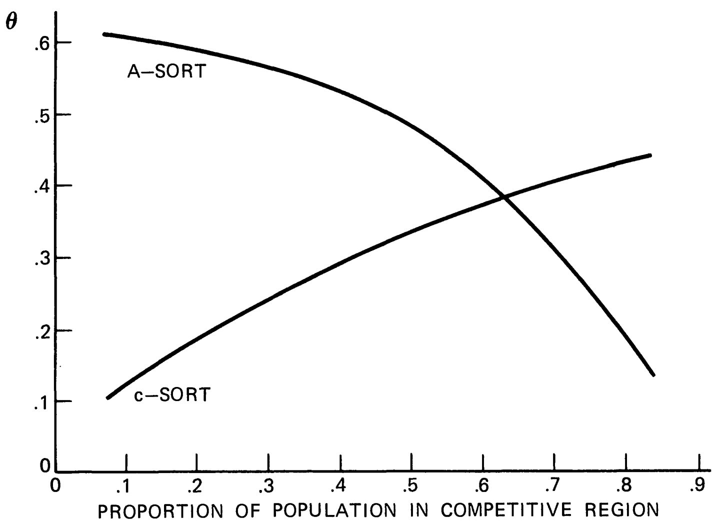
The parameter θ is the proportional discount to low-price buyers.

When discrimination is prohibited, or when no usable sorting mechanism exists, θ will be 0. As predicted in Section 3, this occurred in the simulations only in the extreme cases in which sorting was based on $A(c)$, A and c were uncorrelated, and all types were served competitively (monopolistically) before sorting.[21]

In less extreme cases, however, θ was also responsive to the proportion of consumers in the competitive and monopoly regions. Using one (A, c) distribution, Figure 5 illustrates

FIGURE 5

RELATIVE EFFECTIVENESS OF RESERVATION–PRICE AND STRENGTH–OF–BRAND–PREFERENCE SORTING



[19] Though one can never be certain that simulation results are entirely general, the conclusions presented are based on a very thorough scan of parameter values, over 1000 simulations of index sorting and 500 simulations of self-selection.

[20] This is true at least when group sizes resulting from different sorting criteria are about the same. Clearly, large differentials can result when a small tail of the distribution is sorted out as one group. It would only be useful to compare the price differential from such a sort with other sorts that yield equally imbalanced group populations.

[21] In fact, the extreme case with sorting on c and all customers in the monopoly region never occurred in the simulations because the number of firms was not restricted to be an integer. See footnote 13.

the relative effectiveness of reservation price and strength-of-brand-preference sorting when different proportions of the population are served in the competitive region before sorting. For a given (A, c) distribution, increases in the population size will lower prices and increase the number of firms, thereby causing a higher proportion to be served competitively before sorting. Thus, varying the population size allows one to study changes in the proportion served monopolistically or competitively without changing the (A, c) distribution.

If A and c are highly correlated, it is difficult (and, one might argue, pointless) to distinguish criteria that sort by reservation price from criteria that sort by strength of brand preference. Figure 6 shows that even when virtually all types are served competitively, sorting on A can be almost as effective as an equally noiseless c-sort if ρ, the correlation between A and c, is near 1 or -1.

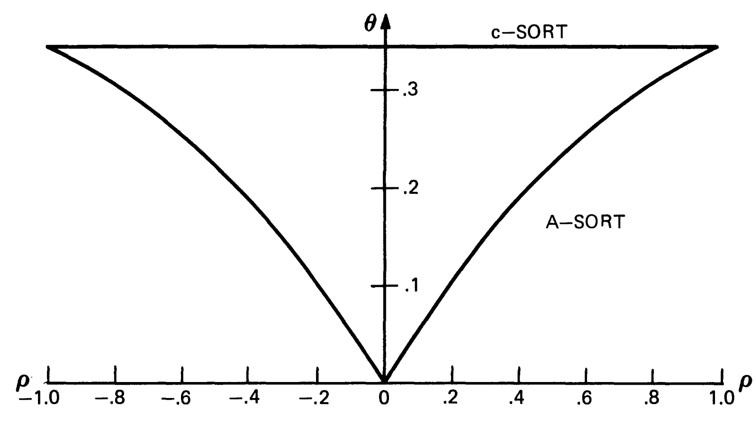
To allow a more rigorous comparison of sorting on A and c, we calculated equilibria with each kind of discrimination, as well as with no discrimination, for 75 sets of population parameters. The results strongly support the theoretical analysis of Section 3. In all cases in which less than 60% of the population was served competitively before sorting, θ was larger with discrimination based on reservation price. In all cases with more than 80% of the population in the competitive region, θ was greater with sorting on strength of brand preference. Between 60% and 80%, the results were less distinct.[22]

Price discrimination affects the equilibrium number of firms in an unsurprising way. Discrimination results in more firms than when the practice is prohibited. Equivalently, with the same number of firms, each earns positive profits. The strength of this effect depends on the effectiveness of discrimination. In the sample the correlation between θ and the change in N was .89 when sorting was based on reservation price and .80 when sorting was based on strength of brand preference.

Sorting based on A tends to have a greater impact on the equilibrium number of firms than sorting based on c. Though A-sorting yielded a smaller average θ, .27 versus .32 with strength-of-brand-preference sorting, it resulted in a much larger average increase in N, 6.2% versus 2.8%. (See Table 1.) The larger increase in N indicates that reservation-price sorting is likely to be more profitable than c-sorting before the entry of new firms. This is borne out in the simulations.

FIGURE 6

SORTING WHEN ALL TYPES ARE SERVED COMPETITIVELY



[22] In 46 of the parameter sets, A and c were uncorrelated. In these a smaller proportion of the population was in the competitive region for all cases in which A-sorting was more effective (larger θ) than for any case in which c-sorting was more effective.

TABLE 1 Measuring the Impact of Price Discrimination

	Sorting on Reservation Price (A)		Sorting on Strength of Brand Preference (c)	
	Index	Self-Selection	Index	Self-Selection
θ	26.8%	16.3%	32.5%	19.9%
Firms*	6.22%	4.22%	2.77%	1.51%
Quantity*	7.86%	4.92%	.29%	.20%
Consumer Surplus*	2.48%	1.07%	−.25%	−.08%

* Average change from the discrimination-prohibited case.

The change in total quantity sold when price discrimination is permitted has traditionally been of interest to economists studying this topic. In a monopolistic model (Robinson, 1933; Schmalensee, 1981; Yamey, 1974) quantity increase is a necessary, but not sufficient, condition for discrimination to improve resource allocation. In this model, where sorting need not be based upon reservation price, the condition is neither necessary nor sufficient.

Nevertheless, it is clear that in the simulations, sorting on A has a more positive effect on the quantity sold than sorting on c. When sorting was based on reservation price in the 75 sample parameter sets, the quantity sold increased an average of 7.86% versus a .29% average increase when sorting was based on strength of brand preference. In all cases A-sorting raised the quantity sold by a greater amount or lowered the quantity sold by less.

In the extreme case in which nearly all of the population is served competitively, c-sorting has virtually no effect on the quantity sold since few, if any, consumers are on the margin of not purchasing the product at all. At the opposite extreme, with nearly all of the population in the monopoly or no-purchase regions, the situation closely resembles monopoly in the short run; A-based discrimination can raise or lower the quantity sold. The profits from discrimination, however, attract entry in this model, and thus increase variety and put downward pressure on prices. Therefore, discrimination would be more likely to increase the quantity sold in these instances than in a monopoly model or any model in which the number of firms is fixed. In fact, sorting on reservation price lowered the quantity sold in only 11 of the 75 cases.

Summing consumers' surplus yields a result similar to that found in the comparison of the quantity sold.[23] As with the quantity sold, this measure of welfare can increase or decrease with the introduction of price discrimination. In the sample studied the average change was a 2.5% increase with A-sorting and a .3% decline with c-sorting. The only cases in which sorting on strength of brand preference increased consumers' surplus were ones in which A and c were correlated.

Of course, the distributional effects of discrimination are always an important issue in policy making, so this simple welfare analysis does not provide sufficient information for choosing among policies. It is clear, though, that the increase in N and decrease in price for the low-price group imply that the sum of consumers' surplus in the low-price group always increases with discrimination. In a few cases discrimination also lowered the price charged to the high-price group. When this happened, the lower price and greater variety necessarily increased total surplus in that group as well. Generally, though, discrimination brought price increases to the high-price group that outweighed the welfare gain from increased

[23] The usual caveat regarding interpersonal utility comparisons is, of course, relevant to this discussion of welfare. Since there are no economic profits, however, welfare analysis can be focused entirely on the consumer surplus of the people in each group. Calculation of the total surplus of each group is a straightforward extension of the quantity equation. See Borenstein (1983).

variety, and thus caused total surplus to fall. It is noteworthy that *all* cases in which discrimination benefitted the high-price group occurred when sorting was based on reservation price either directly or indirectly through strong correlation.

In the context of demand creation and diversion, these welfare results are intuitively appealing. When discrimination only heightens the competition for purchases that were already being made, without creating any new sales, one would expect that it would hurt the groups not singled out for competition. On the other hand, sorting designed to expand total market sales—demand "creation" in the monopoly region (and for finite price cuts, also the no-purchase region)—allows firms to spread fixed costs over more buyers, and can thus end up helping all buyers. Therefore, sorting on strength of brand preference, which is generally demand diverting, is less likely to improve welfare than sorting on reservation price, which more often is demand creating.

6. Analysis of the model with self-selection

■ We now generalize the analysis to include self-selective sorting. The major conclusion of the previous section still holds with self-selection: sorting on strength of brand preference and sorting on reservation price still differ strongly in their effectiveness, depending on the proportion of potential purchases in the competitive versus the monopoly and no-purchase regions. The effects on the equilibrium number of firms, the quantities sold, and consumer surplus are changed somewhat by self-selection, however.

The self-selective sorting function focuses on the cost of qualifying for the low price. Receiving a discount price often requires advance purchase (airlines, stage performances) or greater time to make the transaction (warehouse sales, newspaper coupons). These requirements are useful to the firm when the cost they impose on buyers is positively correlated with their willingness-to-pay for the product.

Two important factors help to explain the differences between index and self-selective sorting. First, with self-selection, the difference between a firm's two prices determines how many people fall into each group. As the differential gets very small, for instance, almost no one will go to the trouble of qualifying for the lower price. Thus, the profit-maximizing prices for each group are not independent. Second, many or all of those who buy at the low price are incurring a total cost greater than the purchase price. Those who pay the low price plus a qualification cost slightly less than $P_H - P_L$ would incur a lower total cost without discrimination if the resulting single price were less than P_H. Unless the variety, N, increased substantially with price discrimination, these types are worse off with discrimination, even though they obtain the lower price.

We assume that a person's cost of qualifying for the discount price, K, is a function of his (A, c) pair:

$$K(A, c) = \text{Max}\left\{0,\ r + \frac{s(A - A_0)}{\sigma_A} + \frac{t(c - c_0)}{\sigma_c}\right\}, \tag{18}$$

where

r = the cost to the (A_0, c_0) type of qualifying for the discount,
s = the change in cost per unit difference between A_0 and a person's A value, and
t = the change in cost per unit difference between c_0 and a person's c value.

For the simulations we chose A_0 and c_0 to correspond to the critical values of the index sorting, the median positive values of these variables (for A_0, the median of those values greater than marginal cost). The differences $A - A_0$ and $c - c_0$ are scaled by their standard deviations, σ, to give meaning to the relative values of s and t.

It is clear that all types of people for whom the cost of qualifying for the discount is less than the difference between the high and low price, $K(A, c) < P_H - P_L$, will obtain the low price if they buy the product. The cost to these types is $P_L + K(A, c)$. These people are the newspaper coupon clippers and supersaver flyers. Those with $K(A, c) \geq P_H - P_L$ are the types who find it too inconvenient or otherwise costly to qualify for the discount. These are the full-fare airline passengers or the people who cannot wait for a sale to buy that new dining room set.

To aid in comparing self-selective with index sorting, we simulated the same 75 sets of population parameters that were studied in Section 5 with self-selection. Each set of population parameters was simulated with $r = 1.0$, $s = 5.0$, $t = .001$, and with $r = 1.0$, $s = .001$, $t = 5.0$.[24] Setting the larger value in each case at 5.0 is rather extreme since the price differential with index sorting averaged \$1.23 and was greater than \$5.00 in only one of the 150 cases. It is useful to note that as s gets large holding t near 0, or as t gets large holding s near 0, the sorting resembles more and more closely index sorting on reservation price or strength of brand preference, respectively. With $s = 5.0$ and $t = .001$, the qualification cost for types with A more than .2 standard deviations below A_0 would be 0 (since $r = 1$). If the price differential were \$1.23 in equilibrium, types with A more than .25 standard deviations above A_0 would choose to pay the high price instead. Only about 18% of the potential purchases would incur a positive qualification cost if purchase were made.[25] Thus, the high values of s or t are a way to examine the effect of a small perturbation of index sorting towards self-selection.

The effects of this small change are numerically small in some cases, but they are consistent across the various measures of the impact of price discrimination, as shown in Table 1. Self-selection sorting is consistently less effective than index sorting. The average discount to low-price buyers is smaller, as is the average change in the number of firms, the total quantity demanded, and consumer surplus. One possibly surprising result is that, while sorting on strength of brand preference lowers consumer surplus on average, the effect is weaker with self-selection than with index sorting. The fact that self-selection lowers the level of price discrimination that is viable seems to outweigh the deadweight loss from the qualification cost. This is not always the case, however. In three of the 57 c-based discrimination cases in which index sorting lowered total consumer surplus, self-selection lowered consumer surplus by a greater amount.

As was the case with index sorting, sorting based on strength of brand preference yielded larger price differentials on average in the sample, yet it had smaller effects on the number of firms, the quantity sold, and consumer surplus. The effectiveness of A-based and c-based sorting is again explained well by the proportion of potential purchases in the competitive, monopoly, and no-purchase regions. These simulations of price discrimination based on self-selective sorting indicate that its effects are qualitatively the same as index sorting, though they are quantitatively smaller.

7. Conclusion

■ This article has presented a spatial approach to modelling price discrimination in monopolistically competitive markets. Within this framework we have found that free entry alone will generally not prevent price discrimination. Differences among consumers in strengths of brand preference are incorporated and found to be an important factor in understanding the nature of discrimination in these markets.

Sorting based on strength of brand preference is not present in traditional models of price discrimination. In monopolistically competitive markets, however, it can be the most effective basis for sorting buyers. When a change in one brand's price results primarily in

[24] For technical reasons in the computer program, neither s nor t could be set equal to 0.

[25] This assumes that the distribution in the positive quadrant is normal so that the median is equal to the mean.

brand switching ("demand diversion"), rather than in expansion or contraction of total market sales ("demand creation"), sorting based on strength of brand preference will be more effective than reservation-price sorting; the former will yield larger price differentials, larger profits in the short run, and more firms in long-run equilibrium.

On the other hand, when changes in one brand's sales are primarily associated with changes in total market sales, competition among firms is less important. Hence, strength of brand preference will be a less effective basis for sorting customers. The situation is similar to monopoly models of price discrimination; effective discrimination is based on the value customers derive from the product.

Discrimination affects the total quantity sold more strongly when sorting is based on reservation price than when it is based on strength of brand preference. This is consistent with the "creation" and "diversion" terminology we have used to describe the effects of the sorting criteria. Sorting on the basis of reservation price can increase or decrease the total quantity sold as it can with monopoly discrimination. It appears, though, that increases are more likely under monopolistic competition than with monopoly as the short-run profits from discrimination bring entry and downward pressure on prices in the long run.

Both sorting criteria increase the sum of consumers' surplus in the resulting low-price group. Only sorting based on reservation price, however, either directly or indirectly (through correlation between reservation price and brand preference), can raise the surplus of consumers in the high-price group. Total consumer surplus is also affected more favorably by reservation-price sorting.

The analysis concluded by examining the model with self-selective sorting. The results closely parallel the previous findings, but the changes brought about by discrimination are, in general, smaller. The fact that consumers can incur costs to affect the price that they must pay appears to moderate the price spread that is sustainable. Large differentials encourage a greater proportion of the population to pay the qualification cost and to obtain the lower price.

Appendix

■ This appendix describes the procedure used for finding symmetric spatial equilibria by computer simulation. The parameters specified were the fixed cost (F), marginal cost (m), the parameters of the normal density function (μ_c, σ_c, μ_A, σ_A, ρ), population size (L), the conjectural variation in price ($CNJV$), the values of the index-sorting function (A_0, c_0) or the values of the self-selective sorting function (A_0, c_0, r, s, t), and starting values for the number of firms (N) and the price that all other firms charge (P').

For a given N we computed equilibria in the two groups separately in the model of index sorting. Finding an equilibrium with self-selection required switching back and forth between the groups because of the relationship between the price differential and the sorting mechanism. In both cases the program calculated firm X's marginal revenue if it were to respond to all other firms' charging P' to a group by also charging P' to that group. The numerical approximation of MR at P' was

$$MR = \frac{(P'_+ q(P'_+, P'_x)) - (P'q(P', P'))}{q(P'_+, P'_x) - q(P', P')}, \tag{A1}$$

where $P'_+ = (1.001)P'$ and $P'_x = (1 + CNJV \cdot .001)P'$.

The starting values for P' and N were chosen arbitrarily at first and later from the solutions to simulations with fairly similar parameter values. If $MR > (<) m$, the price charged by all other firms, P', was decreased (increased). We determined the size of the change in P' by linear extrapolation from the P''s of the previous two steps and associated MR's. Eventually, the program arrived at the price for "all other firms" at which firm X's best response was to charge the same price. When we found this short-run equilibrium for both groups of consumers, we adjusted the number of firms up or down, depending on the sign of profits. Again, we determined the step size by linear extrapolation. With a new N we again began the search for a Bertrand equilibrium. We continued this until the absolute value of profits in the short-run equilibrium was less than the convergence criterion.

It was useful to note that fixed costs and population affected the results only through their ratio. This comes from two conditions of the model.

Condition 1. Profits are homogeneous of degree one in L and F:

$$\Pi = q_H(P_H - m) + q_L(P_L - m) - F, \tag{A2}$$

where q_H and q_L are homogeneous of degree one in L. Since equilibrium occurs when $\Pi = 0$, a zero-profit set of prices and N will not be affected by a scaling up or down of L and F.

Condition 2. MR is independent of L and F:

$$MR = P + \frac{q(P)}{q'(P)}. \tag{A3}$$

In this model $q(P_x)$ and $q'(P_x)$ are both homogeneous of degree one in L and, of course, independent of F. Thus, changes in L or F do not affect MR.

We also simplified the analysis by observing that an identical decrease (increase) in all A values (e.g., by changing μ_A) had the same effect on equilibria as an increase (decrease) in marginal cost. The only difference was in the level of the equilibrium prices. The quantities sold, the number of firms, and welfare were the same.

To prove this it is sufficient to show that only the level of prices is affected if MC and all A values are increased or decreased by the same amount. If this proposition is true, then starting from $m = 5$ and $\mu_A = 10$ for a type, a change to $m = 5$ and $\mu_A = 12$ (increasing all A values by 2) will be equivalent to a change to $m = 3$ and $\mu_A = 10$ (decreasing m by 2). A sketch of the proof is as follows.

For a type in the monopoly region,

$$\epsilon_{monop} = -P_x/(A - P_x)$$

so the markup equation is

$$\frac{P_x - m}{P_x} = \frac{A - P_x}{P_x}, \quad \text{or}$$

$$P_x = (A + m)/2. \tag{A4}$$

Price increases by exactly the increase in A and m. But, this implies that

$$q_{x_t} = L_t((A - m)/c). \tag{A5}$$

The quantity sold to that type is unchanged by the change in A and m.

In the competitive region

$$(P_x - m)/P_x = (P' - P_x + c/N)/P. \tag{A6}$$

In equilibrium $P' = P_x$. So $P = m + c/N$. Again, price increases by the increase in A and m:

$$q_{x_t} = L_t(1/N + (P' - P_x)/c). \tag{A7}$$

Again, quantity is unchanged.

The monopoly-no-purchase border, $A = P$, shifts up by the increase in P as does the monopoly-competitive border, $A = (P_x + P' + c/N)/2$. Since the borders shift up by the same amount as the A values, all types remain in the same region. Since A and P increase by the same amount, consumer surplus for any purchase is unchanged. This implies that total consumer surplus is unchanged, because total quantity is also constant. Finally, the markup, $P - m$, is unchanged, so firms' profits remain the same. A zero-profit equilibrium occurs with the same number of firms as before the shift in A and m.

References

BORENSTEIN, S. "Price Discrimination in Free-Entry Markets." Ph.D. Dissertation, M.I.T., 1983.

———. "Price Discrimination in Free-Entry Markets." Institute of Public Policy Studies Discussion Paper #206, University of Michigan, November, 1984.

CHAMBERLIN, E.H. *The Theory of Monopolistic Competition.* Cambridge: Harvard University Press, 1931.

ECONOMIDES, N. "Symmetric Equilibrium Existence and Optimality in Differentiated Product Markets." Discussion Paper #197, Columbia University, Department of Economics, February 1984.

FRANK, R.H. "When Are Price Differentials Discriminatory?" *Journal of Policy Analysis and Management,* Vol. 2 (Winter 1983).

GREENHUT, M.L. AND OHTA, H. *A Theory of Spatial Pricing and Market Areas.* Durham, N.C.: Duke University Press, 1975.

LERNER, A.P. AND SINGER, H.W. "Some Notes on Duopoly and Spatial Competition." *Journal of Political Economy,* Vol. 45 (April 1937).

NORMAN, G. "Spatial Competition and Spatial Price Discrimination." *Review of Economic Studies,* Vol. 48 (January 1981).

PIGOU, A.C. *The Economics of Welfare.* London: MacMillan, 1932.

ROBINSON, J. *The Economics of Imperfect Competition.* London: MacMillan, 1933.

SALOP, S.C. "Monopolistic Competition with Outside Goods." *Bell Journal of Economics,* Vol. 10 (Spring 1979).

SCHMALENSEE, R.L. "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination." *American Economic Review,* Vol. 71 (March 1981).

YAMEY, B. "Monopolistic Price Discrimination and Economic Welfare." *Journal of Law and Economics,* Vol. 17 (October 1974).