

## Price Discrimination

For most of chapters 1 and 2 we assumed that a monopolist produces a single good and sells it at a uniform price (per unit). In general, uniform pricing leaves some surplus to the consumers. (Indeed, we saw in chapter 2 that this nonappropriability of the consumer surplus is a force pushing toward the introduction of too few products by a monopolist.)

Uniform pricing may be a good generalization for most retail markets. However, examples abound in which the same economic good is sold at different prices to different consumers. A doctor may charge a rich patient more than a poor one, or may charge an insured patient more than an uninsured one. The price of a consumer good may vary across areas of a city or a country, even in the absence of any significant cost differential. The same economic good may also be sold at two different prices to the same consumer. This occurs, for instance, when the producer practices quantity discounts, in which the marginal unit can be purchased at a lower price than the inframarginal ones. Such examples may be viewed as attempts by the producer to capture a higher fraction of consumer surplus than he would if he charged a uniform price.

It is hard to come up with a satisfactory definition of price discrimination. Roughly, it can be said that the producer price-discriminates when two units of the same physical good are sold at different prices, either to the same consumer or to different consumers.

This definition is unsatisfactory, and sometimes it must be amended or extended. First, consider the case of a cement producer serving a geographic area. To the producer's costs must be added the freight costs. Suppose that the cement producer is vertically integrated and thus provides his own transportation. In such a case, a uniform delivered price is discriminatory, whereas delivered prices that respond fully to transportation-cost differentials among consumers located at different distances from the factory are not. Hence, we will say that there is no price discrimination if differences in prices between consumers

exactly reflect differences in the costs of serving these consumers (this amounts to considering the *net* cost of serving a consumer). Second, it should not be inferred that price discrimination does not occur when differentiated products are sold to different consumers.<sup>1</sup> The use of different qualities of services (e.g., classes in trains and airplanes) is also partly an attempt to capture consumer surplus by separating consumers into different groups, as we shall see. Thus, it is difficult to offer an all-encompassing definition. A general-equilibrium theorist might rightly point out that goods delivered at different dates, at different locations, in different states of nature, or of different quality are distinct economic goods and thus that the scope of "pure" discrimination is very limited.<sup>2</sup>

The possibility of price discrimination is linked to the possibility of arbitrage. It is conventional to distinguish between two types of arbitrage.

The first type of arbitrage is associated with the *transferability of the commodity*. It is clear that if the transaction (arbitrage) costs between two consumers are low, any attempt to sell a given good to two consumers at different prices runs into the problem that the low-price consumer buys the good to resell it to the high-price one. For instance, the introduction of quantity discounts (which, as we will see below, is often optimal) implies that, in the absence of transaction costs between consumers, only one consumer buys the product and resells it to other consumers. For example, if each consumer buys according to the "two-part tariff"  $T(q) = A + pq$  (where  $A > 0$  is a fixed fee and  $p$  is the marginal price), only one consumer will pay the fixed fee. Hence, if there are many consumers, everything is almost as if the manufacturer sold at the linear (or uniform) price. If consumers can arbitrage perfectly, the producer is generally forced to charge a uniform or fully linear price:  $T(q) = pq$ .

Transaction costs offer a clue as to when price discrimination is feasible. Services such as medical treatment and travel are much less transferable than most retail commodities. Similarly, the consumer can hardly engage in arbitrage for electricity or telephone calls.

Of course, perfect (costless) arbitrage and no arbitrage

are two extreme cases. In general, some limited arbitrage may occur, depending on the relative cost and benefit: Consider the use of a false student ID card in order to enjoy a student discount. An interesting case of partial arbitrage, and therefore partial discrimination, is that of a manufacturer who sells his product to several retailers. The retailers may arbitrage if the manufacturer tries to charge different marginal prices to different retailers. This prevents the manufacturer from imposing general nonlinear tariffs  $T(q)$  to his retailers. However, even though he may not be able to observe the exact quantity sold by each retailer, he may observe that a retailer carries his product. Then (ignoring legal constraints) he can charge a two-part tariff  $T(q) = A + pq$ , where  $A$  is a fixed premium (a "franchise fee," in this case).

Two-part tariffs will also be applicable whenever the (variable) consumption of the arbitrated good is linked with the (fixed) consumption of a complementary product, as in the case of razor blades and razors or that of Polaroid film and Polaroid cameras. If pictures (rather than film and camera) are thought of as the final consumption good, the producer can manipulate the relative prices of the two inputs that produce pictures to discriminate (i.e., can charge two different prices for different units of the good "picture"). The analogy with the franchising example is that the fixed fee must in practice be paid by every consumer, while the good which is consumed in variable proportions may be subject to arbitrage.

The second type of arbitrage is associated with the *transferability of demand* between different packages or bundles offered to the consumers. Here there is no physical transfer of good between consumers. The consumer simply chooses between the different options offered. For instance, a consumer may have a choice between buying two units of the good at total price  $T(2)$  and buying one unit at price  $T(1)$  (this is known as a price-quantity package) or between first-class and second-class accommodations on a train (a price-quality package). As we will see, if consumers' tastes differ, the producer generally wants to target a specific package for each consumer. However, in the absence of information about the identity of each

1. In the example of the cement producer, consumers indeed buy spatially differentiated goods. However, each consumer can consume only a well-defined commodity ("cement delivered at the consumer's location"). We want to extend the notion of discrimination to cases in which the consumer is given a choice among several differentiated goods.

2. See Philips 1983 for an extensive discussion of the notion of discrimination. This introduction to chapter 3 is influenced by his introduction to the topic. Varian 1987 and Wilson 1985 are two other useful treatments of price discrimination.



consumer (the producer only knows the aggregate distribution of tastes), the producer must make sure that each consumer indeed chooses the package designed for him and not the package designed for another consumer. For instance, the traveler to whom first class is targeted should not want to travel second class because the price savings more than offset the reduction in quality. This, as we will see, puts "incentive compatibility" constraints on the set of packages offered by the producer. The producer must use "self-selection devices."

In terms of the consequences of discrimination, the two types of arbitrage are naturally very different. The transferability of the product tends to prevent discrimination, whereas the transferability of demand may induce the producer to increase discrimination. In the supplementary section we will see that, under reasonable assumptions, the producer enlarges the quantity spectrum or the quality spectrum when he has information about aggregate demand rather than about individual demands.

Following Pigou (1920), it is customary to distinguish three types of price discrimination. *First-degree* price discrimination is perfect price discrimination—the producer succeeds in capturing the entire consumer surplus. This occurs, for instance, when consumers have unit demands and the producer knows exactly each consumer's reservation price and (if these reservation prices differ) can prevent arbitrage between consumers. It then suffices for the producer to charge an individualized price equal to the consumer's reservation price. Perfect price discrimination is unlikely in practice, either because of arbitrage or because of incomplete information about individual preferences. In the case of incomplete information about individual preferences, the producer may still be able to extract consumer surplus imperfectly by using the self-selecting devices mentioned earlier. This is called *second-degree* price discrimination.<sup>3</sup> Also, the producer may observe some signal that is related to the consumer's preferences (e.g. age, occupation, location) and use this signal to price-discriminate; this is termed *third-degree* price discrimination. The important difference between second-degree and third-degree price discrimination is that

third-degree discrimination uses a direct signal about demand, whereas second-degree discrimination selects indirectly between consumers through their choice between different packages.

We will not quite follow the traditional ordering. After considering when perfect price discrimination is feasible (section 3.1), we will study third-degree price discrimination (section 3.2). The ideas contained in section 3.2 are direct applications of previous material on multi-product monopoly and the inverse elasticity rule (see section 1.1). To clearly differentiate this type of imperfect discrimination from second-degree price discrimination, we will assume that the monopolist is able to divide the market into  $n$  segments of customers on the basis of direct signals, but that he is unable to discriminate between consumers within a group (either because of unit demands or because of commodity arbitrage). We will then tackle the newer concept of second-degree price discrimination by assuming that the monopolist knows how consumers' preferences are distributed within a group but does not know each consumer's preferences. This will lead to a consideration of screening or self-selection devices, which is developed in more detail in the supplementary section. Our study of second-degree price discrimination will also be extended to exhibit the analogy between discrimination over quality spectra and nonlinear tariffs.

### 3.1 Perfect Price Discrimination

The simplest kind of perfect price discrimination occurs when a single consumer (or, equivalently, a number of identical consumers) has unit demand. Suppose that each consumer has  $v$  as his willingness to pay (valuation) for a good. The monopolist, by charging price  $p = v$ , extracts the whole consumer surplus.

Next, consider the case of identical downward-sloping demands. Suppose that the  $n$  consumers in the market all have the same demand,  $q = D(p)/n$ , for the monopolist's product, and that this demand function (and, therefore, the aggregate function  $q = D(p)$ ) is known by the mo-

3. We encountered an instance of second-degree price discrimination in the supplementary section of chapter 1, where we saw how time could screen consumers with different valuations for a durable good. High-valuation consumers are more eager to buy early. They pay a higher price than low-valuation buyers, who wait to purchase the good. A difference with the examples below

is that in the earlier example the monopolist actually loses from price discrimination. As we saw, he would prefer to commit to a fixed (uniform) price path (also, the model was constructed so that in a one-period framework the monopolist would be unable to discriminate). In section 3.3, we will assume the credibility issue away.



nopolist. By using an appropriate pricing schedule, the monopolist can increase his profits over those achieved with a linear pricing schedule (which yields  $p^m D(p^m) - C(D(p^m))$ ); he can even extract all of the potential social surplus. By *pricing schedule*, or *tariff*, we mean the total amount of money  $T$  to be paid by the consumer as a function of consumption  $q$ . A linear pricing schedule corresponds to a single price:  $T(q) = pq$ . An affine pricing schedule corresponds to a two-part tariff:  $T(q) = A + pq$ .

Imagine first that the monopolist adopts the competitive pricing schedule, i.e.,  $T(q) = p^c q$ , where  $p^c$  is the competitive price (see figure 3.1). Let  $S^c$  be the corresponding net consumer surplus:

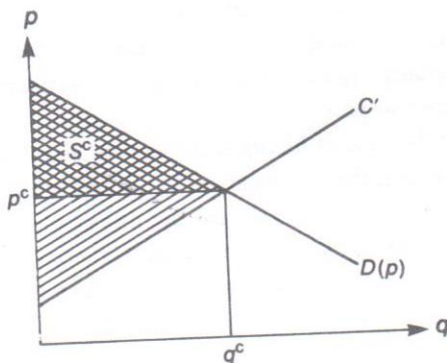
$$S^c = \int_0^{q^c} [P(q) - p^c] dq,$$

where  $P(q) \equiv D^{-1}(q)$  is the inverse demand function. Now suppose that, in order to have the right to buy at  $p^c$ , the consumer must pay a fixed premium. This fixed premium,  $A$ , may go as high as  $S^c/n$  without inducing the consumers to refrain from buying anything. By offering the "affine" (but nonlinear) pricing schedule or "two-part tariff,"

$$T(q) = \begin{cases} p^c q + \frac{S^c}{n} & \text{if } q > 0 \\ 0 & \text{if } q = 0, \end{cases}$$

the monopolist realizes a profit equal to

$$\Pi = S^c + p^c q^c - C(q^c),$$



**Figure 3.1**  
( $\Pi$  is the area of the shaded section.)

which is simply the social surplus at the optimum. This is to be expected, since the marginal price charged by the monopolist is equal to his marginal cost, and consumers have zero surplus after the fixed premium  $S^c/n$  is subtracted. It is easy to see that this strategy affords the monopolist his maximum profit. (In particular, he makes more profit than with the optimal linear schedule, which yields a profit equal to  $p^m q^m - C(q^m)$ .) Unless the monopolist uses coercive means, the consumers can always guarantee themselves a zero surplus by not buying at all. Since the sum of the consumer surplus and the monopoly profit is equal to the total surplus, the maximal profit of the monopolist is equal to the maximal total surplus.

#### Exercise 3.1\*

(i) Show that an alternative way to realize the optimal profit under discrimination is to charge each consumer a tariff equal to his gross surplus:

$$T(q) = \int_0^q P(x) dx/n.$$

(ii) Generalize to the case of a monopolist facing a competitive fringe at  $p_0 > p^c$ . (By definition, a competitive fringe is willing to supply any demand at price  $p_0$  but supplies nothing for a lower price.) What is the optimal two-part tariff? To implement an alternative way of appropriating consumer surplus, how can one redefine "gross surplus"?

So far, we have assumed that consumers are identical. Now suppose that the consumers have different demand curves and assume that the monopolist knows each individual demand curve. For the monopolist, the optimal pricing scheme consists of charging for the marginal unit a price ( $p^c$ ) equal to the marginal cost and demanding a personalized fixed premium equal to the net surplus  $S^c$  from consumer  $i$  at the price  $p^c$ . A special application of this result is the case where the aggregate demand function is derived from unit demand functions (each consumer consumes one or zero unit of the good) of consumers each of whom has a different willingness to pay. Then, the two-part tariff is equivalent to a simple system of personalized prices, with each consumer paying a sum equal to his willingness to pay.

Of course, there are considerable disclosure problems. Certainly a consumer is not likely to reveal that he is among those willing to pay a higher price for the good. This may destroy the possibility of discrimination. For instance, when consumers have unit demands and the producer knows only the distribution of valuations over the population but not the individual valuations, a single price is charged—i.e., discrimination does not occur. (The optimal price is then equal to the monopoly price for the downward-sloping demand curve formed by the addition of unit demands.)<sup>4</sup>

We now turn to imperfect price discrimination.

## 2 Multimarket (Third-Degree) Price Discrimination

### 2.1 The Inverse-Elasticity Rule Again

Suppose that a monopolist produces a single product at a total cost of  $C(q)$ , and that he is able to divide the aggregate demand into  $m$  "groups" or "markets" on the basis of some "exogenous" information (e.g. age, sex, occupation, location, or first-time vs. second-time buyer). These  $m$  groups have  $m$  distinct downward-sloping demand curves for the product. The demand curves are known to the monopolist. We will assume that arbitrage cannot occur between groups but that, at the same time, the monopolist cannot discriminate (even in a second-degree sense) within a group. (These conditions will have to be checked in each application.) Hence, the monopolist charges a linear tariff for each group.

Let  $p_1, \dots, p_i, \dots, p_m$

denote the prices in the different markets, and let

$q_1 = D_1(p_1), \dots, q_i = D_i(p_i), \dots, q_m = D_m(p_m)$

denote the quantities demanded. Let

$$q = \sum_{i=1}^m D_i(p_i)$$

denote the aggregate demand. The monopolist chooses prices to maximize his profit:

$$\sum_{i=1}^m p_i D_i(p_i) - C\left(\sum_{i=1}^m D_i(p_i)\right).$$

Formally, this price-discrimination program is a special case of the multiproduct monopolist's pricing problem described in chapter 1, where demands are independent and costs are (possibly) dependent. We know from this analysis that relative price margins are given by the inverse-elasticity rule: For all  $i$ ,

$$\frac{p_i - C'(q)}{p_i} = \frac{1}{\varepsilon_i},$$

where  $\varepsilon_i = -D'_i(p_i)p_i/D_i(p_i)$  is the elasticity of demand in market  $i$ . Optimal pricing implies that the monopolist should charge more in markets with the lower elasticity of demand.

This rule explains why students and senior citizens are given discounts by private firms with no redistribution intention, why legal and medical services are priced according to the customer's income or amount of insurance coverage, why the prices of goods in different countries sometimes do not reflect transportation costs and import taxes, and why first-time subscribers to a magazine are given discounts.<sup>5</sup>

### 3.2.2 Welfare Aspects

When a monopolist's various products are actually the same physical good sold to different markets, an interesting question is: What would happen if the monopolist were forced to charge the same (uniform) price in all markets? Comparing the two situations gives a measure of the effect of third-degree price discrimination. The monopolist is better off under price discrimination, because "at worst" he can always charge the uniform price in each market. Consumers in low-elasticity markets are adversely

Discrimination could operate if the monopolist could charge a tariff function of the probability for the consumer to get the good (i.e., whether the consumer is the good or not is not determined in advance, but depends on the "toss of coin"). However, it can be shown that if both the monopolist and the consumers are "risk-neutral," such complicated mechanisms are not optimal for the monopolist.

5. In this last example, price discrimination could also be explained by matching issues—see the supplementary section of chapter 2. The idea here is more similar to that developed in the supplementary section of chapter 1; the consumers who have not yet subscribed are perceived by the monopolist as less eager to buy the magazine than those who have already subscribed.