

Chapter 20

PRICE DISCRIMINATION

R. Preston McAfee*

This chapter sets out the rationale for price discrimination and discusses the two major forms of price discrimination. It then considers the welfare effects and antitrust implications of price discrimination.

1. Introduction

The Web site of computer manufacturer Dell asks prospective buyers to declare whether they are a home user, small business, large business or government entity. Two years ago, the price of a 512 MB memory module, part number A0193405, depended on which business segment one declared. At that time, Dell quoted \$289.99 for a large business, \$266.21 for a government agency, \$275.49 for a home, and \$246.49 for a small business.

What explains these price differences? How does Dell benefit from it? Different segments have different willingness to pay. Dell optimizes its prices, offering lower prices to relatively price-sensitive segments. An interesting aspect of Dell's attempt to charge different prices to different customers is that the customers aid Dell in its effort. According to a Dell spokesperson, each segment independently sets prices and the customer is free to buy from whichever is cheapest. Thus, the customers paying more are choosing to pay more, probably because they do not expect the prices to vary so significantly.

This chapter explores the economic rationale for price discrimination. Section 2 presents the basic theory of price discrimination and describes the conditions necessary for price discrimination to exist. Section 3 then discusses direct price discrimination, while indirect price discrimination is discussed in Section 4. Section 5 provides a discussion of the welfare effects associated with price discrimination, while Section 6 concludes this chapter with a discussion of the antitrust implications of price discrimination.

2. Basic theory of price discrimination

Price discrimination can exist when three conditions are met: consumers differ in their demands for a given good or service, a firm has market power,¹ and the firm can prevent or limit arbitrage. If consumers had identical demands for a good, then all consumers would demand the same amount of the good for each price, and the price and

* California Institute of Technology.

1. "Market power" as used in this chapter refers to the economic definition of the term, which is the ability to price above marginal cost. See STEVEN E. LANDSBURG, *PRICE THEORY AND APPLICATIONS* 329 (6th ed. 2005); Franklin M. Fisher, *Detecting Market Power*, which appears as Chapter 14 in this book.

quantity of the good would depend only on the number of consumers in the market and the ability of firms to supply the good (the supply curve). If firms have no market power, that is, no ability to affect the price of the goods they sell, the theory of perfect competition implies that all goods would be sold at one price (the law of one price). Finally, if consumers can arbitrage price differences, any attempt to charge higher prices to some group would be defeated by resale.

The basic theory of price discrimination is the theory of monopoly, applied to more than one market or group. What is a monopolist's profit-maximizing price? Consider a firm that can sell $q(p)$ units when it charges price p . The firm's profits are

$$\pi(p) = pq(p) - c(q(p)) \quad (1)$$

where c is the cost function. The function q is the demand facing the firm, that is, it gives the quantity the firm can sell. In the case of monopoly, the demand facing the firm and the market demand are the same. Assume that q is a downward-sloping demand curve. This means that the firm has some pricing power—a price increase does not send the quantity demanded from the firm to zero. This pricing power is known as monopoly power or market power. The assumption rules out perfect competition, for under perfect competition, a price increase would send the quantity demanded from any particular firm to zero.

The first-order conditions for profit maximization entail

$$0 = \pi'(p) = q(p) + (p - c'(q(p)))q'(p) \quad (2)$$

Recall that the elasticity of demand, which measures the responsiveness of demand to price, is given by

$$\varepsilon = -\frac{pq'(p)}{q(p)} \quad (3)$$

The elasticity is not necessarily constant, but depends on p . However, this dependence is suppressed for clarity in exposition. Rearranging Equation (3) slightly, the first-order condition for profit maximization can be expressed as

$$\frac{p - c'(q(p))}{p} = \frac{1}{\varepsilon} \quad (4)$$

The left-hand side of this expression is the proportion of the price which is a markup over marginal cost. It is known as the "price-cost margin." Historically, it is also known as the "Lerner Index." The price-cost margin matters because, in the standard neoclassical model, a competitive industry prices at marginal cost. Thus, the price-cost margin can be viewed as a measure of the deviation from marginal cost. A price-cost margin of zero means that price equals marginal cost, which is the competitive solution. A price-cost margin of $\frac{1}{2}$ means the marginal cost is marked up by 100 percent—half the price is markup.

The formula shows that profit maximization entails a price-cost margin of $1/\varepsilon$. If costs are not negative, the left-hand side is not greater than one, and profit maximization entails an elasticity at least as large as one. What happens when the elasticity is less

than one? At any price where the elasticity is less than one, a price increase is profitable. If demand is everywhere inelastic, the firm always wants a higher price.

The formula for the monopoly price can be rewritten to show

$$p = \frac{\varepsilon}{\varepsilon - 1} c'(q(p)) \quad (5)$$

This formula suggests that maximizing profits entails marking up marginal cost by a fixed percentage that depends on the elasticity of demand. Suppose, for example, that a pharmaceutical drug manufacturer sells its drug in two distinct markets, the United States and Mexico, and that arbitrage between these markets is quite difficult for most consumers. Suppose, in addition, that the drug is manufactured in a single plant, so that marginal cost in both nations is the same, and that the elasticity of demand is constant at 1.04 in the United States and 2 in Mexico. Then the Mexican price would be twice marginal cost, while the U.S. price would be $1.04/0.04 = 26$ times marginal cost, or 13 times as high as the Mexican price.

Along with consumers with varying preferences and firms with some degree of market power, price discrimination can only exist in markets where consumers cannot engage in arbitrage. Under arbitrage, a consumer who is offered a lower price for a good by a firm purchases an excess quantity of the good and resells the good to consumers who are denied the lower price by the firm. Under perfect arbitrage, the firm would be forced to sell all its output at the lowest price to consumers offered the lowest price, who would then resell to other consumers. Thus, arbitrage effectively turns price discrimination into offering a single price.

As a practical matter, there are a number of reasons while arbitrage may be difficult or impossible. These include

- high transportation costs,
- legal impediments to resale,
- personalized products or services,
- thin markets or matching problems,
- informational problems, and
- contracts and warranties.

Transportation costs permit geographic price discrimination, as in freight absorption, where firms charge their customers the same total price for the good including transportation costs. Under freight absorption, the net price firms receive for their goods after absorbing transportation costs, or “netback,” depends on the transportation costs to different consumers. High transportation costs mean that consumers will find it difficult to defeat price discrimination by buying in low-priced areas and reselling in high-priced areas, because transportation costs consume the profits.

Consumers may be prevented by law from engaging in resale. For example, individuals flying on commercial airlines must present personal identification documents that match the name on their ticket/boarding pass. Thus, someone buying a ticket cannot sell the ticket or a portion of the ticket to a third party for that third party’s use. Consequently, airline ticket resale between individual consumers effectively is legally banned.

Some products are so personalized that resale proves impractical. For instance, someone who buys prescription eyeglasses or contact lenses could only sell these products to another person with the same visual impairment. Similarly, individually tailored clothing can only be sold to persons sharing the same personal measurements.

In broader markets, product standards may differ enough to prevent the sharing of some products across these markets. One example is digital video discs (DVD) in the United States and Europe. Differing codes prevent DVD players in one region from playing discs defined for the other region's market. Even computer DVD drives, which are the same worldwide, respect these region codes, generally giving the owner five switches from region to region before locking in a particular region permanently and refusing to play the discs of other regions.

A related impediment to arbitrage is thin markets. A thin market is a market with few buyers and sellers, either because a product is personalized (e.g., tailored clothing) or because very few consumers demand the product (e.g., rare collectibles). Thin markets make identifying candidates for resale difficult. While one might be able to defeat price discrimination in tailored suits in principle, in practice one would be unlikely to identify someone exactly the same size who is being charged more, which is a thin markets problem. Electronic marketplaces such as eBay are gradually mitigating the role of thin markets as a barrier to arbitrage.

If there is uncertainty about the quality of the good, resale markets may not function at all, or function imperfectly. Prices for medicines intended for veterinary use often are substantially lower than prices for the identical compound for human consumption, yet few people attempt to arbitrage these differences, primarily out of fear that the veterinary use medicines will be of lower quality.

Contracts can also be used to create a legal impediment to resale. One of the more common types of contracts used to inhibit arbitrage is a nontransferable warranty. Meade Instruments Corporation, a major telescope manufacturer, offers nontransferable warranties, for example. Free textbooks given to professors come with contractual prohibitions on resale. For textbooks, these contractual provisions are usually ineffective.

3. Direct price discrimination

Historically, price discrimination has been divided into three types, using terminology introduced by Arthur Cecil Pigou (1877-1959). First degree price discrimination meant perfect price discrimination, meaning that each buyer paid 100 percent of his or her subjective value of the goods purchased, and prices were based on the buyer's identity. Third degree price discrimination meant an imperfect form of first degree. Second degree, in contrast, has come to mean offering a menu of options, like a quantity discount, and letting buyers choose what to buy. Pigou intended, however, for second degree to mean using approximations to first degree discrimination.

This nomenclature is seriously flawed. There is no sense in which second degree price discrimination is an intermediate case between first and third degree price discrimination. Instead, first and third degree price discrimination are each examples of where different groups of consumers are charged different prices for the same good,

while second degree price discrimination refers to instances where consumers in a market are presented with the same set of price and quantity options and “self-select” into different groups. A more modern and perhaps more useful delineation among the various types of price discrimination designates the old first and third degrees of price discrimination as “direct price discrimination” and the second degree as “indirect price discrimination.” While direct price discrimination may use the actual identity of the customer as a basis for price discrimination, more commonly prices are conditioned on customer characteristics, and customers with the same characteristics receive the same prices.

3.1. Common characteristics used in direct price discrimination

Assuming that the basic conditions for price discrimination are present (i.e., firms with market power, consumers with distinct levels of demand, and no arbitrage), consumers are typically divided by one or more of the following characteristics:

- geography,
- nationality,
- age,
- employer, and
- purchasing history.

Pharmaceutical pricing is a leading example of pricing based on geography and nationality. It is no secret that drug prices vary radically from country to country, even countries that are adjacent to each other. For example, drug prices in Canada and Mexico typically are lower than in the United States, often leading the U.S. elderly and other heavy drug consumers to attempt to buy their drugs across either border. Large U.S. group purchasers, including several state Medicaid programs and private health insurance plans, also have investigated buying some or all of their drugs in Mexico or Canada.

A perverse example of price discrimination exists in the market for protease inhibitors and other AIDS drugs. Counterintuitively, the prices of AIDS drugs in Norway were lower than the prices available in Uganda. A simple reading of price discrimination theory would lead one to expect that Norway would have higher prices, because its citizens’ incomes are much higher and, thus, their demand for AIDS drugs would be much more inelastic than the demand for AIDS drugs in Uganda.² The basic price discrimination equations presented in Section 2 above would then suggest that prices would be higher in Norway, where demand is more inelastic, and lower in Uganda, where demand is more price-sensitive.

2. After a certain point, AIDS patients with higher incomes would no longer demand additional drugs, because these drugs would have no additional positive marginal impact on their health (and in sufficiently high amounts, could actually harm their health). Changes in the price of AIDS drugs thus would not affect the demand for these drugs much among patients with higher incomes. In contrast, poorer patients are more likely to be price-sensitive, that is, they would purchase more of the drug if prices were lower, because these drugs would continue to have positive health benefits.

A deeper analysis explains the higher AIDS drug prices in Uganda. The number of AIDS cases in Norway is relatively low and the cases tend to be concentrated among those with the lowest incomes, where substantial elasticity exists. In Uganda, the incidence of AIDS is much higher. Most Ugandans cannot afford AIDS medication even at Norwegian prices, so the only cases that are relevant from a pricing perspective are from the wealthier Ugandans. Given the extreme dispersion of income in Uganda, these individuals may well be wealthier than average Norwegian AIDS patients. Consequently, this price discrimination was profitable, even if quite unfortunate from a social perspective. After substantial political pressure was brought to bear, prices in Africa were reduced.

An interesting example of the power of the pricing formula arises when generic drugs come into existence. It is a well-documented fact that when generic drugs begin to be sold, the price of the brand name drug rises. Why? Prior to the existence of generics, the brand name drug served all the potential buyers of the drug, and there was no means of discriminating against the inelastic customers. Generics siphon off demand from more of the price-sensitive (elastic) consumers than of the inelastic consumers, leaving relatively inelastic customers. Consequently, once generics come into existence, the demand for the brand name drug becomes more inelastic, and the price rises, even though overall demand for the name brand drug falls. The brand name drug, of course, makes a lower profit than absent competition, but more than it would have at the lower price.

Senior citizen discounts and student discounts are both examples of age-based price discrimination. Insurance policies offered to employees of particular employers discriminate on the basis of the employer. Some companies offer goods and services only to people with particular employers such as the government, military, or universities (e.g., the TIAA-CREF pension system).

Basing price offers on other purchases is an increasingly common form of direct price discrimination. In this case, "special offers" are made to those with Citibank credit cards, top airline frequent flyer status, those who purchased a computer, and the like. Credit card data are valuable for price discrimination purposes, because these data permit identifying likely value from related purchases.

A "gray market product" is a product sold to one country that appears for sale in another. Thus, Sony cameras and camcorders for Mexico appear on U.S. Internet sites at a 10 percent to 20 percent discount. These products may have manuals in Spanish, a linguistic impediment to arbitrage, and some threat that the warranty will not be honored. Similarly, European Mercedes automobiles are imported into the United States as gray market cars, although in this case, an important purpose was to evade U.S. air pollution laws rather than arbitrage the price. Gray markets, however, often involve attempts to arbitrage price differences across nations.

3.2. Freight absorption

A specific example of geographically based price discrimination is known as freight absorption, in which the seller of a good does not charge for transportation. How should a seller deal with transportation costs? Historically, steel sellers in locations other than

Pittsburgh priced steel as if it were shipped from Pittsburgh. This pricing pattern likely arose because prices in Pittsburgh were competitive, and sellers outside Pittsburgh, in the process of meeting the competition, maximized profits by pricing just below the cost of steel delivered from Pittsburgh. This, in some cases, has the peculiar effect of pricing *higher* as transportation costs *fall*.

The logic is illustrated in Figure 1. Consider a single steel mill located in Gary, Indiana. This seller faces competition from competitive Pittsburgh mills, and prices from Pittsburgh involve a price at the mill door plus a transportation charge. The delivered price—price plus transportation cost—from Pittsburgh is denoted by the thick black line.

It is lowest in Pittsburgh and rising as one moves away from Pittsburgh. This represents the highest price the Gary mill can charge and still have any sales. Consequently, if there are no other relevant mills, the Gary mill charges the higher of the Pittsburgh delivered price and the Gary costs, represented by an even thicker grey line.

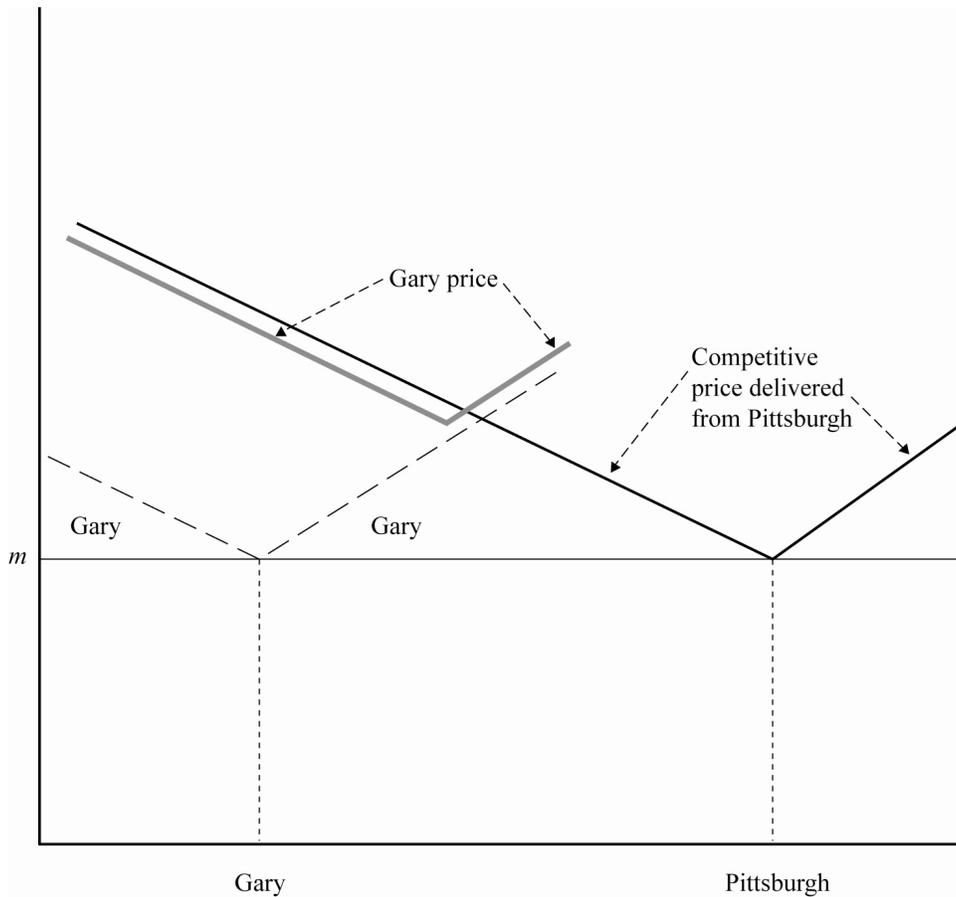


Figure 1.

A single seller in Gary can price at the competitive price of steel delivered from Pittsburgh.

Competition from Pittsburgh insures that the Gary mill charges more in Gary itself than it does in Ohio, where the competition with Pittsburgh is fiercer.

A phenomenon that is conceptually similar to freight absorption is tax absorption. A monopolist absorbs some portion of a tax, while competitive industries pass on taxes. Recall the formula

$$p = \frac{\varepsilon}{\varepsilon - 1} c'(q(p)) \quad (6)$$

which gives the markup on marginal cost by a monopolist. A per unit tax, like an excise tax on gasoline, acts like an increase in marginal cost. With constant elasticity, imposing a tax of amount t per unit increases marginal costs by t , which increases the price by $(\varepsilon / (\varepsilon - 1))t$. With constant elasticity, the seller does not absorb any of the tax, but in fact increases prices by more than the tax. In the case of linear demand, $q(p) = a - p$, a seller with marginal cost c and facing a per unit tax t maximizes profits $\pi(p) = (a - p)(p - (t + c))$, which yields the price $p = \frac{1}{2}(a + t + c)$. Thus, in the case of linear demand, a seller absorbs one-half of any tax.

3.3. Identifying price discrimination

Many common firm behaviors are inappropriately identified as price discrimination. For example, grocery stores offer advertised specials that can be quite substantial. Moreover, the size of the discounts and the items that are discounted fluctuate from week to week and from store to store. Because all customers receive the same sales price at each store, there is no price discrimination. When the availability of a low price depends on the possession of a “club card” or other identifying characteristic, the store has engaged in direct price discrimination. Although, when the cards are provided freely, as is common in grocery stores, price discrimination is indirect.

Having prices vary by location is commonly referred to as *price dispersion*, since prices vary across stores in a manner not based on costs. The only circumstance where price dispersion would rise to the definition of price discrimination is when the same company owns several stores which charge different prices for the same good. In this case, the company might be using stores or geography as a means of price discriminating and would be engaging in indirect price discrimination.

Price dispersion can have effects similar to price discrimination, even though it is a distinct phenomenon. Consumers who search across stores benefit from price dispersion and obtain lower prices on average, while consumers who do not search will tend to pay higher prices. Consequently, price dispersion creates discrimination based on the willingness and ability to search. The difference, however, is that no firm is engaging in price discrimination, but rather that discrimination is happening at the market level.

Passing on higher costs to the customer is not price discrimination. With advance purchase discounts, it is often a challenge to distinguish price discrimination from cost-based pricing. This distinction is especially challenging in a dynamic setting where both prices and marginal costs vary. Consider airline seats. If the world were perfectly predictable, the price of a seat on a plane would be a posted price and stay the same until the flight takes off. However, the world is not predictable, and an unusually large number of early sales reduces the number of available seats, increasing the value of the

remaining seats. The expected value of the remaining seats is the opportunity cost to the airline of a sale. Consequently, fluctuations in sales cause the opportunity cost to vary, in a manner that is essentially unobservable. Similarly, a matinee showing at a movie theater, or an “early bird special” at a restaurant, are not examples of price discrimination—a movie at noon is not the same product as a movie starting at 7:00 p.m.

4. Indirect price discrimination

As previously discussed, direct price discrimination is threatened by arbitrage, where consumers offered lower prices buy large quantities of a firm’s goods and sell these goods to other consumers who face higher prices. Because direct price discrimination may be difficult to maintain in the face of arbitrage, firms may also engage in indirect price discrimination practices. Generally, indirect price discrimination refers to a setting where a menu of options is offered to all and customers choose which option is best for them. By varying the quality and features of a product, the combinations of prices and quantities offered, or requiring consumers to buy more than one product, firms can offer all consumers the same set of choices, but allow consumers to sort themselves (“self-select”) into groups with differing levels of demand.

4.1. *Quantity discounts*

Quantity discounts—buy one, get the second at half price—are an important form of price discrimination. That a quantity discount is price discrimination is transparent from the fact that the units are sold at different prices.³ Quantity discounts can in principle be defeated by arbitrage. Arbitrage amounts to buying a large quantity and reselling the individual units, which is profitable if there is a quantity discount.

Quantity discounts may also have this feature of damaging the product for the purpose of cutting its price, via the creation of inconvenient bundles. Consider a company that takes 24 individually wrapped rolls of paper towels and shrink-wraps the bundle at some modest expense. Each roll regularly sells for \$2, and the company offers the “valu-pak” for \$24, or \$1 per roll. Consumers with large families buy the valu-pak, while consumers living in small apartments, driving small cars, or who use few paper towels do not. The effect of the valu-pak is that it has offered a substantial discount to large family buyers, and large family buyers tend to be more price-sensitive, if only because they have more mouths to feed. The valu-pak makes money for the seller because it offers a discount which is mostly chosen by the price-sensitive group.

4.2. *Coupons*

Coupons operate in a similar way to the valu-pak, only based on the value of time rather than the scale of demand. Individuals generally value their time at approximately their wages, so that people with low wages, who tend to be the most price-sensitive, also have the lowest value of time. Coupons, which are pieces of paper that grocery stores will redeem for fifty cents or a dollar off the price of an item, come in newspaper

3. If the manufacturer merely passes on the lower unit cost of larger containers, then no price discrimination has occurred.

bundles on hundreds of distinct products. A thrifty shopper may be able to spend an hour sorting through the coupons in the newspaper and save \$20 on a \$200 shopping expedition (the amount depends on the nature of the items the consumer purchases—buying a lot of crackers, potato chips and paper towels increase the savings). This is a good deal for a consumer who values time at less than \$20 per hour, and a bad deal for the consumer that values time in excess of \$20 per hour. Thus, relatively poor consumers choose to use coupons, which permits the seller to have a price cut that is approximately targeted at the more price-sensitive group.

There are two important points to observe about coupons. Coupons work through a correlation between the price-sensitivity and the value of time. If the people with a low value of time are the least price-sensitive for the item, coupons will not work for the seller. That is, coupons offer discounts to people with a low value of time (and a few who are obsessive). The seller would like to offer discounts to people with elastic demand. Thus, coupons are effective provided the people with a low value of time are usually the people with elastic demand. Coupons for pork rinds, for instance, are unlikely to be effective price discrimination tools because low-income people may be willing to pay more on average for pork rinds than high-income people.

A related form of indirect price discrimination is trade-in discounts. For example, a camera store may offer a \$50 discount on the purchase price of a new camera for anyone who trades in an old camera. Individuals with old cameras are likely to have more elastic demands for new cameras than consumers with no cameras to trade in, because the old camera is likely to be a close substitute for a new camera. By offering trade-in discounts, retailers provide an incentive to their customers to group themselves according to their elasticity of demand.

4.3. Bundling

Bundling of goods exploits similar logic as quantity discounts. AT&T offered a discount for long-distance service bundled with a discount for a Jiffy Lube oil change for a car. This targets the discount on long-distance service to people who either will obtain a Jiffy Lube oil change, or are close to obtaining a Jiffy Lube oil change. Does it make sense for AT&T? If the Jiffy Lube customers are, on average, more price-sensitive (elastic demanders) for long-distance telephone service than people generally, it will be a discount that targets the correct group. If, on the other hand, Jiffy Lube customers tend to be less price-sensitive, it targets the wrong group with a discount.

4.4. Performance-based discrimination

Some instances of indirect price discrimination involve offering two versions of a good, one of which has been damaged or “crimped” so as to offer reduced functionality. IBM did this with its popular LaserPrinter by adding chips that slowed down the printing to about half the speed of the regular printer. The slowed printer sold for about half the price, under the IBM LaserPrinter E name. Similarly, Sony sold two versions of its mini-discs (a form of compact disc created by Sony): a 60-minute version and a 74-minute version. The 60-minute version differs from the 74-minute version by software instructions that prevent writers from using a portion of the disc.

The IBM LaserPrinter E and the 60-minute Sony mini-disc are literally damaged products that sell for less. Package delivery company Federal Express appears to slow down the delivery of “second day air” to insure that it is usually two days, for otherwise more people would choose the second day air delivery over the more expensive next day delivery. These companies create two versions of their product, one worse than the other, for the purpose of inducing consumers to voluntarily split into two groups based on willingness to pay.⁴

4.5. Placing restrictions on purchase and use

Placing restrictions on purchase and use, if these can be enforced, can also cause consumers to sort into different groups that may be related to their elasticity of demand. Such restrictions are precisely analogous to performance-based restrictions. For example, airlines price discriminate in several distinct ways—including offering lower prices for customers who stay over on Saturday nights, advance purchase discounts, roundtrip discounts, and point of origin discounts. Airline price discrimination represents an attempt to charge the business traveler more than the leisure traveler, because business travelers typically have less elastic demand. By offering different prices through Saturday night stay-over restrictions, advance purchase discounts, roundtrip discounts, and point of origin discounts, airlines are attempting to have consumers sort themselves between business and leisure travelers.

Saturday night stay-over restrictions on discounted tickets can be substantial. The easy way to see that a requirement to stay over Saturday night is price discrimination is to consider two roundtrips. Imagine flying from Los Angeles to Dallas on Monday, and returning to Los Angeles on Friday in week one, and then again flying Los Angeles to Dallas on Monday, and returning to Los Angeles on Friday in week two. One can buy these four flight segments as two roundtrip tickets in two different ways. They can be purchased as roundtrips leaving and departing the same week, or as roundtrips leaving and departing in different weeks. These configurations are illustrated in Table 1.

Each configuration involves two roundtrip tickets. Configuration 1 is composed of two roundtrips originating at Los Angeles on Monday and going to Dallas, and returning Friday. Configuration 2, in contrast, has one ticket which originates at Los Angeles and returns the following week, and a second roundtrip which originates at Dallas on a Friday and returns the following Monday. The second configuration is often substantially less expensive, because each roundtrip involves a Saturday night stay-over.

Airlines bundle roundtrips, in that it is usually cheaper to buy a roundtrip than to buy two one-way fares. Sometimes it is even cheaper to buy a roundtrip and throw away the return coupon than to buy the one-way fare! Both Saturday night stay-overs and bundling of roundtrips make it possible to charge business travelers—who often want to

4. There could be a dispute about whether damaged goods are indeed the same good for the purposes of evaluating price discrimination. The fact that the company expends resources to reduce the quality of a portion of the production—so that what started as the same good is made less valuable—suggests inclusion as price discrimination. Some authors use “different markups” as evidence of price discrimination, a procedure which still leaves unresolved the question of how similar the goods must be to be treated as the same good.

Table 1.

Two different ticket configurations involving exactly the same set of four flights

<i>Configuration 1</i>			
Ticket	Week	Day	Segment
1A	1	Monday	Los Angeles to Dallas
	1	Friday	Dallas to Los Angeles
1B	2	Monday	Los Angeles to Dallas
	2	Friday	Dallas to Los Angeles
<i>Configuration 2</i>			
Ticket	Week	Day	Segment
2A	1	Monday	Los Angeles to Dallas
	2	Friday	Dallas to Los Angeles
2B	1	Monday	Dallas to Los Angeles
	2	Friday	Los Angeles to Dallas

be home with their families for the weekend and would prefer one-way fares for their flexibility—more money.

Airlines offer discounts for tickets purchased in advance. Two-week and six-week advance purchase discounts are common. It is difficult to establish that these advance purchase discounts are in fact price discrimination, because there is clearly some advantage to the airline in knowing demand earlier—the airline can save money by planning its route and plane deployment more effectively if it knows more about the state of the demand. However, the magnitude of the advance purchase discounts suggest that advance purchase discounts involve discrimination, especially because business travelers change their plans more frequently and thus are usually unable to take advantage of advance purchase discounts. One travel agency claimed that the average business traveler who purchases advance fares changes each planned itinerary an average of three times in the month preceding travel. With a \$100 change fee, that produces \$300 additional costs in change fees alone.

4.6. Knowledge-based discrimination

An interesting kind of discount is based on information. This kind of price discrimination straddles the direct and indirect categories. It is available to anyone who knows to ask for it, and in this way is like indirect price discrimination. However, only some people know to ask, and in this way it is like direct price discrimination. This chapter opens with an example of Dell's pricing that fits this description. Another example was an AT&T discount for long-distance service, in the early days of competition with Sprint and MCI, which was advertised in the *Wall Street Journal*. The

discount was available to anyone, but the users of the discount tended to be either *Wall Street Journal* readers or friends of *Wall Street Journal* readers, which put them in the upper income brackets. This sounds superficially like it targets the wrong group, but in fact higher income customers were more likely to switch to competing long-distance services than poorer customers. Consequently, the discount was targeted at the group more informed about the competition, which might in fact be the more price-sensitive group in this case.

If one calls a fine hotel and ask for a price of a room on a given night and a room is available, a price (room rate) will be quoted. If one asks whether a better rate is available, generally the answer is yes. Hotels offer discounts to people who know to ask for discounts. This is a sensible policy for the hotel because the same people who know to ask are frequent guests of hotels, more informed of their options, and as a result more responsive to a price cut. That is, basing the price on the customer's knowledge works when the knowledgeable group is also the elastic demand group.

4.7. *Nonlinear pricing*

Firms with linear pricing offer their customers the same price for each unit of a good or service, that is, the total charge for all goods sold is a straight-line function of quantity. Anything else is nonlinear pricing, including the quantity discount example discussed previously.

Another simple example of nonlinear pricing (which still generates a line, just not a line through the origin) is a "two-part tariff." To be allowed to purchase a product, consumers must first pay a fixed entry fee. After paying the entry fee, consumers then pay the same price for each additional unit, generating a straight-line relationship between quantity and price starting with a zero quantity at the fixed fee (see Figure 2).

Public utilities such as electricity, gas, water, and telephone often engage in two-part tariff pricing, where customers' bills include a fixed monthly fee and then a flat price per unit (at least up to some relatively high quantity). Two-part tariffs can be used by sellers to increase profits. In principle, two-part tariffs are sufficient to extract all of the gains from trade. The technique is to set the marginal charge equal to marginal cost, and then set the fixed fee equal to the maximum fixed fee the consumer is willing to pay, which is the value of the trade to the consumer. Setting the marginal charge equal to marginal cost makes trade efficient, maximizing the gains from trade, because the consumer maximizes the value of the good minus the cost, and minus a fixed cost, and thus chooses a quantity that maximizes the gains from trade. The fixed fee then transfers those maximal gains from trade to the seller. The sizes of these charges arise from demand and are illustrated in Figure 3.

Two-part tariffs involve quantity discounts, because the average cost of purchase is a declining function of quantity. If we represent the fixed fee by F and the marginal charge by m , the total charge is $F + mQ$, where Q is quantity. Thus, the average cost per unit is $m + F / Q$, which shrinks as quantity rises.

Buying clubs such as Sam's Club or Costco also are prominent users of two-part tariffs. Club members pay an annual fee to be allowed to shop at the club's retail outlets, where goods are sold at very low or even zero margins. The low margins can be

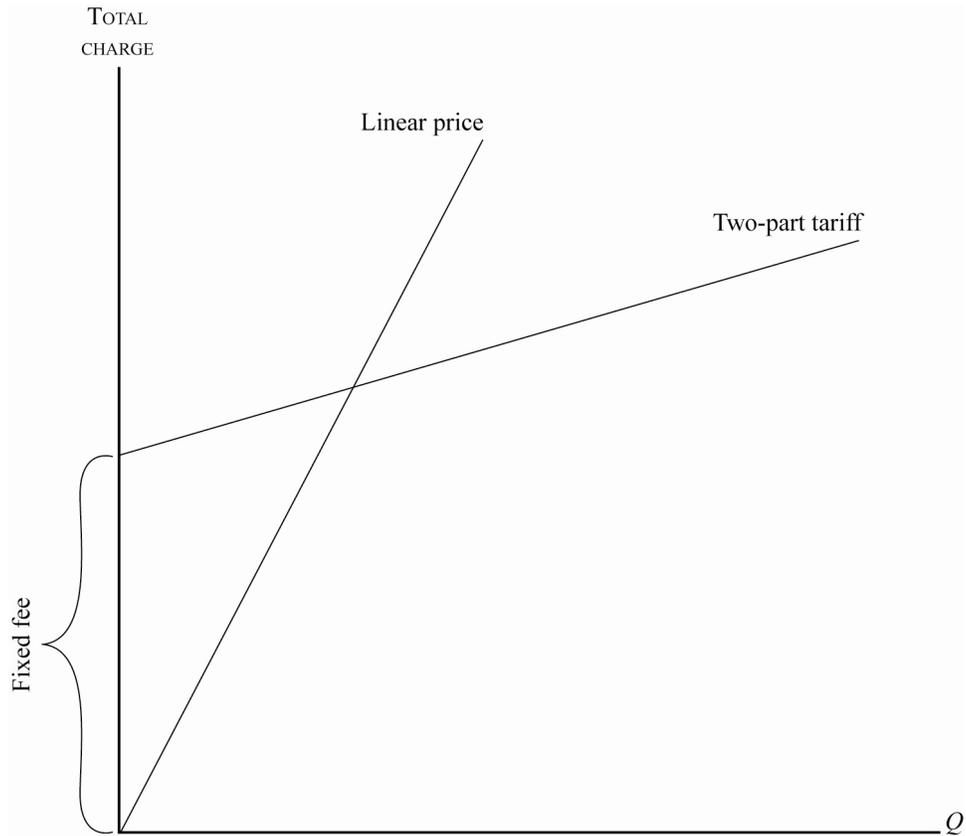


Figure 2.

A linear price and a two-part tariff.

deduced from a careful reading of these firms' annual financial reports. For example, Costco recently reported total income before taxes that was almost identical to the sum of its members' annual fees, so that Costco did not report any additional profits from its operations (its retail sales).

A set or menu of two-part tariffs can be used to generate a broader nonlinear price. Consider the two charges in Figure 2, one linear and one a two-part tariff. If both are offered, consumers with low values for large quantities will opt for the linear charge, while consumers with high values for a large quantity will opt for the two-part tariff. Generally, offering both functions and then letting the consumer choose is tantamount to offering a single charge, which is the minimum of the two functions.⁵ This nonlinear

5. However, when customers' demands vary month to month, offering the minimum is distinctly different from offering the selection. Perhaps the most prominent example involves cellular telephone pricing, which tends to involve a fixed charge up to some maximum quantity (minutes per month) and then a high price per additional unit. This pricing scheme presents customers with risk, and induces them to purchase more minutes than they expect to use. The addition of "rollover" minutes, however, mitigates the risk.

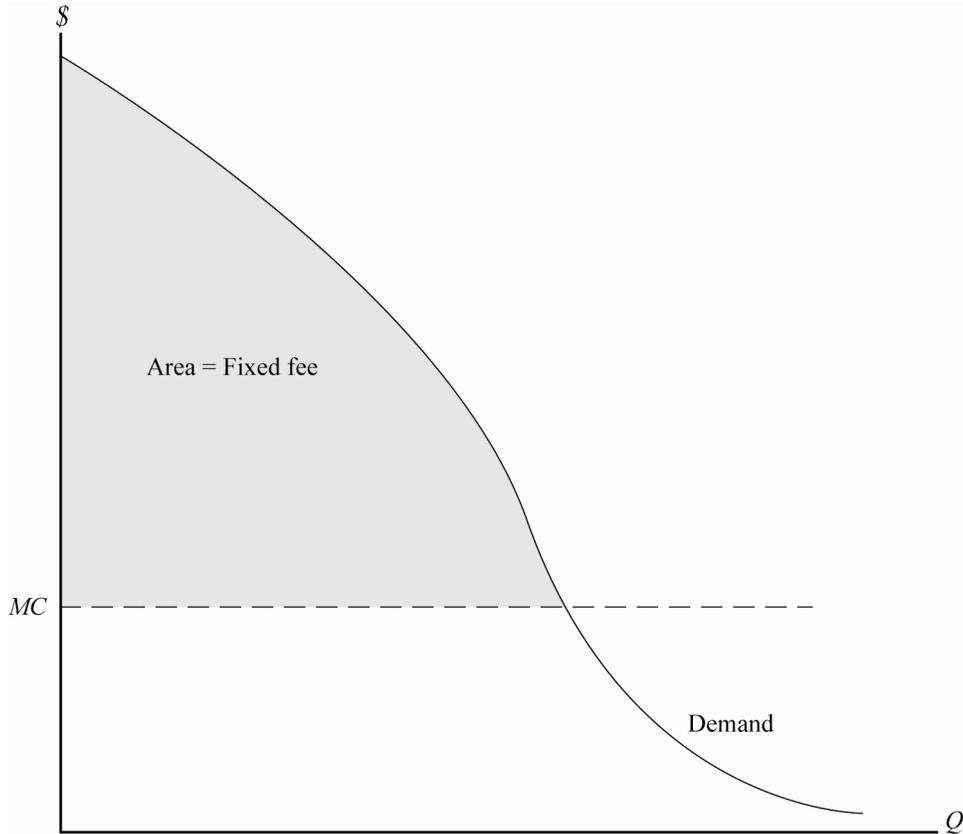


Figure 3.

The maximal fixed fee. MC refers to both marginal cost and marginal charge.

price is illustrated in Figure 4, using the same price offers from Figure 2. The minimum, which is the best the consumer can obtain in the way of charges, is denoted in dark grey.

As Figure 4 illustrates, nonlinear pricing offers discounts for larger quantities. In principle, nonlinear pricing can be used to create surcharges for larger quantities. Water pricing in some communities has this property—higher marginal charges for higher usage.

How low should the marginal charges go? Generally, they should not go below marginal cost for the purpose of raising revenue. Prices below marginal cost lose money on increased volume; replacing them with marginal cost, and appropriate changes in fixed fees, increases the gains from trade.⁶ No matter what else the seller offers, the

6. There are two kinds of circumstances where price may rationally go below marginal cost. One is when price is used to influence the behavior of rival firms. Predatory pricing is one example, but price may also go below marginal cost to increase the sale of complementary goods. A second circumstance where price may rationally go below marginal cost is when price is used to influence consumer behavior. Examples include a loss-leader inducing consumers to expend resources (such as a low price on milk to encourage consumers to visit a grocery), or an introductory offer that is used to subsidize information collection from customers.

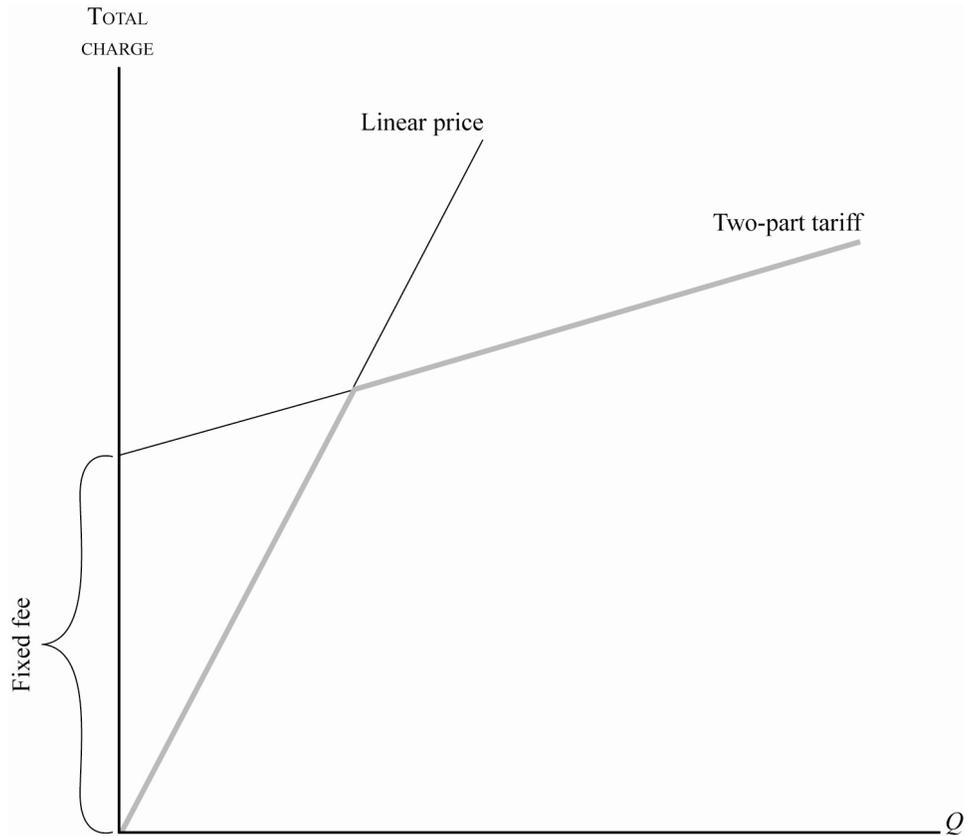


Figure 4.

A nonlinear price, in gray, constructed as the minimum of two offers.

seller profits from offering a marginal charge equal to marginal cost, with an appropriate fixed fee. To see this, suppose the buyer buying the largest quantity is paying a marginal charge in excess of marginal cost. By offering this buyer a two-part tariff with a marginal charge equal to marginal cost, and a fixed fee equal to what the buyer used to pay over marginal cost, the seller gets as much revenue as before and covers its costs, while the buyer is strictly better off. This permits the seller to charge a slightly higher fixed fee and, thus, increase profits.

5. Welfare effects

The Robinson-Patman Act prohibits price discrimination that results in a lessening of competition. A variety of state laws go further, limiting price discrimination to final consumers. Are these laws justified? Does price discrimination harm consumers overall? These questions involve assessing the welfare effects of price discrimination—who is harmed and who benefits?

5.1. Welfare effects of direct price discrimination

To model the effects of price discrimination, it generally suffices to consider two groups of buyers, with group 1 having relatively inelastic demand compared to group 2, such that at the same price, p , the elasticity of demand for group 1 is less than for group 2. Subscripts are used to denote the variables, so for example $q_1(p)$ is the quantity demanded by group 1.

The welfare effects of price discrimination can be decomposed into two separate effects. First, price discrimination changes the distribution of output. Starting from a single price, prices rise for inelastic buyers (group 1) and fall for elastic buyers. This effect reduces welfare and consumer surplus, but increases profits. Moreover, consumer surplus falls by more than profits increase, so the net effect on welfare is negative. Net welfare—the sum of profits and consumer surplus—is the gains from trade.

Price discrimination reduces the gains from trade, holding output constant, because maximizing the gains from trade requires selling at the same price in all markets. With two different prices, arbitrage would increase the gains from trade, resulting in a common price. Price discrimination represents a rearrangement of output away from high value buyers and toward buyers with lower values. The absence of price discrimination, which makes every buyer's marginal value equal, creates the distribution of output that maximizes welfare. Price discrimination, in contrast, makes group 1 have a higher price than group 2 and, thus, reduces the gains from trade. Indeed, by starting at the profit-maximizing solution, then permitting free arbitrage by the buyers, a single price would result.

The second effect of price discrimination is on the total quantity, and quantity can either rise or fall under price discrimination. From the analysis of the distributional effect, we see that *price discrimination increases overall welfare only if the quantity rises by a sufficiently large amount.*

Two examples show that the overall effect of price discrimination is ambiguous—the gains from trade may rise or fall under price discrimination when compared to the absence of price discrimination. Both examples use linear demand, with

$$q_i(p) = a_i - p \quad (7)$$

for constant a_i and marginal costs set to zero. Consistent with the elasticity assumption, assume $a_1 > a_2$. Under price discrimination, prices are set to $\frac{1}{2}a_i$, which produces gains from trade of $\frac{3}{8}(a_1^2 + a_2^2)$. With no price discrimination, there is an issue of whether the smaller market is served. It is readily verified that profit maximization implies that the smaller market is served only if $a_2 > (\sqrt{2} - 1)a_1$. That is, if price discrimination is prohibited,

$$p^* = \begin{cases} \frac{1}{2}a_1 & \text{if } a_2 < (\sqrt{2} - 1)a_1 \\ \frac{1}{4}(a_1 + a_2) & \text{if } a_2 > (\sqrt{2} - 1)a_1 \end{cases} \quad (8)$$

It is a simple computation to see that the gains from trade under a prohibition of price discrimination are

$$\begin{cases} \frac{3}{8}a_1^2 & \text{if } a_2 < (\sqrt{2}-1)a_1 \\ \frac{1}{16}(7a_1^2 + 7a_2^2 - 2a_1a_2) & \text{if } a_2 > (\sqrt{2}-1)a_1 \end{cases} \quad (9)$$

From this computation, we see that price discrimination increases social welfare when $a_2 < (\sqrt{2}-1)a_1$, and it decreases social welfare when $a_2 > (\sqrt{2}-1)a_1$. The first case arises when a prohibition on price discrimination means that the second market is not served at all, so permitting price discrimination opens up a new market. The price in the first market is the same in either case. In this case, price discrimination is a Pareto improvement, meaning that no one is harmed, and some gain, from price discrimination. Prices for group 1 remain the same, the price for group 2 falls, and profits rise.

In contrast, the second case is where a prohibition on price discrimination does not shut down the second market. In this case, it turns out that linearity of demand keeps the overall quantity constant at $\frac{1}{2}(a_1 + a_2)$, so that price discrimination has no quantity effect, just an adverse distributional effect, and welfare falls.

An important aspect of the welfare effects of price discrimination is that it involves lower prices for the price-sensitive group, and for most items, the price-sensitive group is also the poorer group. Thus, lower income groups tend to benefit at the expense of higher income groups under price discrimination. For example, movie discounts are aimed at students and senior citizens. The reason is that these groups are more price-sensitive with respect to movies. Historically, these were also low income groups, although that has changed over the past decades.

5.2. Welfare effects of indirect price discrimination

Generally, the welfare effects of indirect price discrimination are similar to those of direct price discrimination, and indirect price discrimination may either enhance or harm welfare. However, indirect price discrimination, and especially nonlinear pricing, often induces a larger quantity to be offered. Thus, indirect price discrimination and nonlinear pricing often are used to increase the total quantity sold, thereby increasing the gains from trade and market efficiency. Indeed, because a seller using two-part tariffs generally will offer price equal to marginal cost with an appropriate fixed fee, sales under nonlinear pricing tend to be higher than when linear prices are used.⁷ Consequently, the gains from trade and economists' standard notion of welfare are higher under nonlinear pricing than under linear pricing. Indirect price discrimination tends to increase welfare.

However, indirect price discrimination may or may not help consumers. That is, even if the gains from trade rise, consumers may lose as a group. Consumers must lose

7. A general version of this proposition is proved in Hal R. Varian's *Price Discrimination and Social Welfare*, 75 AM. ECON. REV. 870 (1985).

from perfect price discrimination, since it takes all the gains from trade away from consumers. Thus, even when the gains from trade rise under indirect price discrimination, consumers may be worse off.

As with direct price discrimination, when indirect price discrimination is used to open up markets not previously served, it tends to create a Pareto improvement: all groups benefit. This is especially likely with the introduction of multiple versions of the product, for the purposes of selling to low-value buyers previously not served.

6. Antitrust and price discrimination

Price discrimination, euphemistically known as value-based pricing, is an increasingly important phenomenon. Improvements in computing permit tracking customers much more effectively, and computing prices tailored to the customer. As a consequence, price discrimination likely will feature more prominently in antitrust analysis in the future.

Price discrimination currently is a proof of market power, and frequently is used as a basis for concluding that firms have market power in antitrust trials.⁸ However, the use of price discrimination to conclude the presence of market power in an antitrust trial requires serious care for several reasons. First, costs can vary, and passing on cost differences does not represent price discrimination. For example, a computer manufacturer may offer several distinct types of computer, and may find that purchasers of the multimedia version spend more time talking to customer service and return their units more frequently than the buyers of the standard version. Such extra demands on customer service represent an additional cost of supplying the multimedia version, and passing on those costs is not price discrimination.

Second, while price discrimination indicates the presence of market power, it is not necessarily the case that a large amount of price discrimination demonstrates a large amount of market power. For example, competition in long-distance telephone and cellular phone services appears to have increased the degree and presence of price discrimination. Essentially, a monopoly may choose to charge a single price to all, where competition forces down the prices for consumers easily able to switch across companies, leaving the prices for those unable to switch quite high. Competition pushes all prices down, but not necessarily proportionately. Therefore, while price discrimination may be used to conclude the presence of market power, the conclusion of substantial market power requires much more care.

Third, and most importantly, short-run market power leads to price discrimination, but short-run market power may not reliably entail the presence of long-run market power. Usually long-run market power is the major antitrust concern.⁹ For example,

8. See, e.g., U.S. DEP'T OF JUSTICE & FEDERAL TRADE COMM'N, HORIZONTAL MERGER GUIDELINES §§ 1.12, 1.22, 1.42 (1992) (with Apr. 8, 1997 revisions to Section 4 on efficiencies), *reprinted in* 4 Trade Reg. Rep. (CCH) ¶ 13,104.

9. See Jonathan B. Baker, *Competitive Price Discrimination: The Exercise of Market Power Without Anticompetitive Effects* (Comment on Klein and Wiley), 70 ANTITRUST L.J. 643 (2003); William J. Baumol & Daniel G. Swanson, *The New Economy and Ubiquitous Competitive Price Discrimination: Identifying Defensible Criteria of Market Power*, 70 ANTITRUST L.J. 661 (2003); Gloria J. Hurdle & Henry B. McFarland, *Criteria for Identifying Market Power: A Comment on Baumol and Swanson*, 70

U.S. airlines price discriminate in an extraordinary fashion, offering substantial discounts on bundles and the already discussed Saturday night stay-over. In some cases, one passenger may pay five times what the other pays for the identical seat at the identical time. Nevertheless, airline profits have historically been quite low, and viewed from the perspective of profits, the exercise of market power has extracted very little money from the average consumer. Consequently, while U.S. airlines seem to have substantial short-run market power, they do not seem to have much long-run market power.

Identifying price discrimination can be challenging. First, because distinct customers are involved, costs may vary. Such cost differences can be subtle. It is often alleged that dry cleaners discriminate against women because the price for cleaning a woman's shirt is higher than the price for cleaning a man's shirt. However, if the cost of cleaning the average woman's shirt is higher than the cost of cleaning the average man's shirt, perhaps because of differences in materials or the nature of the stains, some or all of the price differences may be due to cost differences, even if the stated product—shirt cleaning—is the same.

The problem of identifying price discrimination is exacerbated by dynamical considerations. A given airline sells the same product—a seat on a flight from Los Angeles to New York on January 19, 2008—at many different times. As demand changes, the value of the seat may rise or fall quite dramatically. Similarly, bottlenecks in transportation permit prices to vary geographically, a fact made quite evident by the gasoline price spike in Chicago in 2001, which did not affect cities in the Northeast.

7. Conclusion

Price discrimination presents an antitrust conundrum. Price discrimination demonstrates the presence of market power and, if the level of discrimination is significant and nontransitory, suggests that antitrust geographic or product markets should be narrowly defined. At the same time, the presence of long-run market power cannot be established by the presence or level of price discrimination. Moreover, geographic markets based on price discrimination often are convoluted due to the presence of some limited degree of arbitrage or substitution.

The prevalence of price discrimination likely will grow as technology—the computing power to offer personalized prices—permits more subtle and effective strategies. Effective antitrust analysis requires effective understanding of firm behavior. As price discrimination grows more pervasive, it grows more important for antitrust analysis.

ANTITRUST L.J. 687 (2003); Benjamin Klein & John Shepard Wiley, Jr., *Market Power in Economics and in Antitrust: Reply to Baker*, 70 ANTITRUST L.J. 655 (2003); Benjamin Klein & John Shepard Wiley, Jr., *Competitive Price Discrimination As an Antitrust Justification for Intellectual Property*, 70 ANTITRUST L.J. 599 (2003); Margaret A. Ward, *Symposium on Competitive Price Discrimination: Editor's Note*, 70 ANTITRUST L.J. 593 (2003).