9 Imperfect Competition

9.1 Introduction

The analysis of economic efficiency in chapter 2 demonstrated the significance of the competitive assumption that no economic agent has the ability to affect market prices. Under this assumption prices reveal true economic values and act as signals that guide agents to mutually consistent decisions. As the Two Theorems of Welfare Economics showed, they do this so well that Pareto-efficiency is attained. Imperfect competition arises whenever an economic agent has the ability to influence prices. To be able to do so requires that the agent be large relative to the size of the market in which they operate. It follows from the usual application of economic rationality that those agents who can affect prices will aim to do so to their own advantage. This must be detrimental to other agents and to the economy as a whole. This basic feature of imperfect competition, and its implications for economic policy, will be explored in this chapter.

Imperfect competition can take many forms. It can arise due to monopoly in product markets and through monopsony in labor markets. Firms with monopoly power will push prices above marginal cost in order to raise their profits. This will reduce the equilibrium level of consumption below what it would have been had the market been competitive and will transfer surplus from consumers to the owners of the firm. Unions with monopoly power can ensure that the wage rate is increased above its competitive level and secure a surplus for their members. The increase in wage rate reduces employment and output. Firms (and even unions) can engage in non–price competition by choosing the quality and characteristics of their products, undertaking advertising, and blocking the entry of competitors.

Each of these forms of behavior can be interpreted as an attempt to increase market power and obtain a greater surplus. When they can occur, the assumption of price-taking behavior used to prove the Two Theorems is violated, and an economy with imperfect competition will not achieve an efficient equilibrium (with one special exception as is detailed later). It then becomes possible that policy intervention can improve on the unregulated outcome. The purpose of this chapter is to investigate how the conclusions derived in earlier chapters need to be modified and to look at some additional issues specific to imperfect competition.

The first part of the chapter focuses on imperfect competition in product markets. After categorizing types of imperfect competition, defining the market structure, and measuring the intensity of competition, the failure of efficiency is demonstrated when there is a lack of competition. This is followed by a discussion of tax incidence in competitive and imperfectly competitive markets. The effects of *specific* and *ad valorem* taxes are then distinguished, and their relative efficiency is assessed. The policies used to regulate monopoly and oligopoly in practice are also described. There is next a discussion of the recent European policy on the regulation of mergers. The final part of the chapter focuses on market power on the two sides of the labor market. Market power from the supply side (monopoly power of a labor union) is contrasted with monopsony power from the demand side. It is shown that both cases lead to inefficient underemployment with wages, respectively above and below competitive wages.

9.2 Concepts of Competition

Imperfect competition arises whenever an economic agent exploits the fact that they have the ability to influence the price of a commodity. If the influence on price can be exercised by the sellers of a product, then there is *monopoly power*. If it is exercised by the buyers, then there is *monopsony power*, and if by both buyers and sellers, there is *bilateral monopoly*. A single seller is a *monopolist* and a single buyer a *monopsonist*. *Oligopoly* arises with two or more sellers who have market power, with *duopoly* being the special case of two sellers.

An agent with market power can set either the price at which it sells, with the market choosing quantity, or the quantity it supplies, with the market determining price. When there is either monopoly or monopsony, it does not matter whether price or quantity is chosen: the equilibrium outcome will be the same. If there is more than one agent with market power, then the choice variable does make a difference. In oligopoly markets *Cournot* behavior refers to the use of quantity as the strategic variable and *Bertrand* behavior to the use of prices. Typically Bertrand behavior is more competitive in that it leads to a lower market price. Entry by new firms may either be impossible, so that an industry is composed of a fixed number of firms, and not be hindered, or incumbent firms may follow a policy of entry deterrence.

Forms of imperfect competition also vary with respect to the nature of products sold. Products may be homogeneous, so that the output of different firms is indistinguishable by the consumer, or differentiated, so that each firm offers a different variant. With homogeneous products, at equilibrium there must be a single price in the market. *Product*

differentiation can either be *vertical* (whereby products can be unambiguously ranked in terms of quality) or *horizontal* (whereby consumers differ in which specification they prefer). Equilibrium prices can vary across specifications in markets with differentiated products. The notion of product differentiation captures the idea that consumers make choices among competing products on the basis of factors other than price. The exact nature of the differentiation is very important for the market outcome. What differentiation implies is that purchases of a product do not fall off to zero when its price is raised above that of competing products. The greater the differentiation, the lower is the willingness of consumers to switch among sellers when one seller changes its price. The theory of *monopolistic competition* relates to this competition among many differentiated sellers who can enjoy some limited monopoly power if tastes differ markedly from one consumer to the next.

When products are differentiated, firms may engage in *non–price competition*. This is the use of variables other than price to gain profit. For example, firms may compete by choosing the specification of their product and the quantity of advertising used to support it. The level of investment can also be a strategic variable if this can deter entry by making credible a threat to raise output.

To limit the number of cases to be considered, this chapter will focus on Cournot behavior, so that quantity is the strategic variable, with homogeneous products. Although only one of many possible cases, this perfectly illustrates most of the significant implications of imperfect competition. It also has monopoly as a special case (when there is a single firm) and competition as another (when the number of firms tends to infinity).

9.3 Market Structure

The structure of the market describes the number and size of firms that compete within it and the intensity of this competition. To describe the structure of the market, it is first necessary to define the market.

9.3.1 Defining the Market

A market consists of the buyers and sellers whose interaction determines the price and quantity of the good that is traded. Generally, two sellers will be considered to be in the same market if their products are close substitutes. Measuring the own-price elasticity of demand for a product tells us whether there are close substitutes available, but it does not identify what those substitutes might be. To identify the close substitutes, one

must study cross-price elasticities of demand between products. When the cross-price elasticity is positive, it indicates that consumers are increasing their demand for one good as the price of the other good increases. The two products are thus close substitutes. Another approach to defining markets is to use the standard industry classification that identifies products as close competitors if they share the same product characteristics. Although products with the same classification number are often close competitors, this is not always true. For example, all drugs share the same classification number but not all drugs are close substitutes for each other.

Markets are also defined by geographic areas, since otherwise identical products will not be close substitutes if they are sold in different areas and the cost of transporting the product from one area to another is large. Given this reasoning, one would expect close competitors to locate as far as possible from each other, and it therefore may seem peculiar to see them located close to one another in some large cities. This reflects a common trade-off between market size and market share. For instance, antique stores in Brussels are located next to one another around the Place du Grand Sablon. The reason is that the bunching effect helps to attract customers in the first place (market size), even if they become closer competitors in dividing up the market (market sharing). By locating close together, Brussels' antique stores make it more convenient for shoppers to come and browse around in search of some antiques. In other words, the bunching of sellers creates a critical mass that makes it easier to attract shoppers.

9.3.2 Measuring Competition

We now proceed on the basis that the market has been defined. What does it then mean to say that there is "more" or "less competition" in this market? Three distinct dimensions are widely used and need to be clearly distinguished.

The first dimension is *contestability*, which represents the freedom of rivals to enter an industry. It depends on legal monopoly rights (patent protection, operating licenses, etc.) or other barriers to entry (economies of scale and scope, the marketing advantage of incumbents, entry-deterring strategies, etc.). Entry barriers protect the market leader from serious competition from newcomers. Contestability theory shows how the threat of entry can constrain incumbents from raising prices even if there is only one firm currently operating in the market. However, when markets are not perfectly contestable, the threat of potential competition is limited, which allows the incumbents to reap additional profits.

A second dimension is the degree of *concentration* that represents the number and distribution of rivals currently operating in the same market. As we will see, the

performance of a market depends on whether it is concentrated (having few sellers) or unconcentrated (having many sellers). A widespread measure of market concentration is the *n-firm concentration ratio*. This is defined as the consolidated market share of the n largest firms in the market. For example, the four-firm concentration ratio in the US cigarette industry is 0.92, which means that the four largest cigarette firms have a total market share of 92 percent (with the calculation of market share usually based on sales revenue). Table 9.1 shows the four-firm concentration ratios for some US industries in 1987.

The problem with the n-firm concentration ratio is that it is insensitive to the distribution of market shares among the largest firms. For example a four-firm concentration ratio does not change if the first-largest firm increases its market share at the expense of the second-largest firm. To capture the relative size of the largest firms, another commonly used measure is the *Herfindahl index*. This index is defined as the sum of the squared market shares of all the firms in the market. Letting s_i be the market share of firm i, the Herfindahl index is given by $H = \sum_i s_i^2$. Notice that the Herfindahl index in a market with two equal-size firms is $\frac{1}{2}$ and with n equal-size firms is $\frac{1}{n}$. For this reason a market with Herfindahl index of 0.20 is also said to have a number's equivalent of 5. For example, if there is one dominant firm with a market share of 44 percent and 100 identical small firms with a total market share of 56 percent, the Herfindahl is

$$H = \sum_i s_i^2 = (0.44)^2 + 100\left(\frac{0.56}{100}\right)^2 = 0.197. \tag{9.1}$$

Table 9.1
Market concentration in US manufacturing, 1987

Industry	Number of firms	Four-firm concentration ratio	Her findahl index
Cereal breakfast foods	33	0.87	0.221
Pet food	130	0.61	0.151
Book publishing	2,182	0.24	0.026
Soap and detergents	683	0.65	0.170
Petroleum refining	200	0.32	0.044
Electronic computers	914	0.43	0.069
Refrigerators/freezers	40	0.85	0.226
Laundry machines	11	0.93	0.286
Greeting cards	147	0.85	0.283

Source: Concentration ratios in manufacturing, 1992, US Bureau of the Census.

This market structure is then interpreted as being equivalent to one with 5 identical firms. Herfindahls associated to some US industries are indicated in table 9.1. These numbers show that the market for laundry firms, which has a number's equivalent less than 4, is more concentrated than the market for book publishers, which has a number's equivalent of 38.

The third dimension of the market structure is *collusiveness*. This is related to the degree of independence of firms' strategies within the market or its reciprocal, which is the possibility for sellers to agree to raise prices in unison. Collusion can either be explicit (e.g., a cartel agreement) or tacit (when it is in each firm's interest to refrain from aggressive price cutting). Explicit collusion is illegal and more easily detected than tacit collusion. However, tacit collusion is more difficult to sustain. Experience has shown that it is unusual for more than a handful of sellers to raise prices much above costs for a sustained period. One common reason is that a small firm may view the collusive bargain among larger rivals as an opportunity to steal their market shares by undercutting the collusive price, which in turn triggers a price war. The airline industry is a good example in recent years of frequent price wars. The additional problem with the airline industry is that fixed cost is high relative to variable cost. This means that once a flight is scheduled, airlines face tremendous pressure to fill their planes, and they are willing to fly passengers at prices close to marginal cost but far below average cost. Thus with such pricing practices, airlines can take large financial losses during price wars.

The three dimensions of market structure and the resulting intensity of competition may be related. The freedom to enter a market may result in a larger number of firms operating and thus a less concentrated market, which in turn may lead to the breakdown of collusive agreement to raise prices.

9.4 Welfare

Imperfect competition, along with public goods, externalities, and asymmetric information, is one of the standard forms of market failure that leads to the inefficiency of equilibrium. It is the inefficiency that provides the motivation for economic policy in relation to imperfect competition. To provide the context for the discussion of policy, this section demonstrates the source of the inefficiency and reports measures of its extent.

9.4.1 Inefficiency

The most important fact about imperfect competition is that it invariably leads to inefficiency. The cause of this inefficiency is now isolated in the profit-maximizing

behavior of firms that have an incentive to restrict output so that price is increased above the competitive level.

In a competitive economy equilibrium will exist where the price of each commodity is equal to its marginal cost of production. This is due to the argument that firms will always move to increase supply whenever price is above marginal cost, since price is taken as given and additional supply will raise profit. As all firms raise supply, prices will fall until there is no incentive for further supply increases. This argument shows that the profit-maximizing behavior of competitive firms drives price down to marginal cost. If marginal cost is constant at value c, then competition results in a price, p, satisfying

$$p = c. \tag{9.2}$$

To see the cause of inefficiency with imperfect competition, consider first the case of monopoly. Assume that the monopolist produces with a constant marginal cost, c, and chooses its output level, y, to maximize profit. The market power of the monopolist is reflected in the fact that as its output is increased, the market price of the product will fall. This relationship is captured by the inverse demand function, $p(y)$, which determines price as a function of output. As y increases, $p(y)$ decreases. Using the inverse demand function, which the monopolist is assumed to know, we have the profit level of the firm as

$$\pi = [p(y) - c]\, y. \tag{9.3}$$

The first-order condition describing the profit-maximizing output level is

$$p + y\frac{dp}{dy} - c = 0, \tag{9.4}$$

which, since $\frac{dp}{dy} < 0$ (price falls as output increases), implies that $p > c$. The condition in (9.4) shows that the monopolist will set price above marginal cost and that the monopolist's price does not satisfy the efficiency requirement of being equal to marginal cost. The fact that the monopolist perceives that its output choice affects price (so $\frac{dp}{dy}$ is not zero) is reflected directly in the divergence of price and marginal cost.

The condition describing the choice of output can be re-arranged to provide further insight into degree of divergence between price and marginal cost. Using the elasticity of demand, $\varepsilon = \frac{dy}{dp}\frac{p}{y} < 0$, the profit-maximization condition can be written as

$$\frac{p - c}{p} = \frac{1}{|\varepsilon|}. \tag{9.5}$$

This equilibrium condition for the monopoly is called the *inverse elasticity pricing rule*. In words, the condition says that the percentage deviation between the price and the marginal cost is equal to the inverse of the elasticity of demand. The expression $\frac{p-c}{p}$ is the *Lerner index*. The Lerner index will be shown shortly to be strictly between zero and one (i.e., $|\varepsilon| > 1$). The monopoly pricing rule can also be written as

$$p = \mu c, \tag{9.6}$$

where $\mu = \frac{1}{1-(1/|\varepsilon|)} > 1$ is called the *monopoly markup* and measures the extent to which price is raised above marginal cost. This pricing rule shows that the markup above marginal cost is inversely related to the absolute value of the elasticity of demand. The higher the absolute value of the elasticity, the smaller is the monopoly markup.

In the extreme case of perfectly elastic demand, which equates to the firm having no market power, price would be equal to marginal cost. For the markup μ to be finite (i.e., price is well defined), it must be the case that $|\varepsilon| > 1$ so that the monopolist locates on the elastic part of the demand curve. If demand is inelastic, with $|\varepsilon| \leq 1$, then the monopolist makes maximum profit by selling the smallest possible quantity at an arbitrarily high price. Since the monopolist operates on the elastic part of the demand curve with $|\varepsilon| > 1$, the Lerner index, $\frac{p-c}{p} = \frac{1}{|\varepsilon|} \in (0, 1)$, provides a simple measure of market power ranging from zero for a perfectly competitive market to one for maximal market power. Therefore a firm might have a monopoly, but its market power might still be low because it is constrained by competition from substitute products outside the market. By differentiating its product, a monopolist can insulate its product from the competition of substitute products and thereby expands its market power.

This relation of the monopoly markup to the elasticity of demand can be easily extended from monopoly to oligopoly. Assume that there are m firms in the market, and denote the output of firm j by y_j. The market price is now dependent on the total output of the firms, $y = \sum_{j=1}^{m} y_j$. With output level y_j, the profit level of firm j is

$$\pi^j = [p - c]\, y_j. \tag{9.7}$$

Adopting the Cournot assumption that each firm regards its competitors' outputs as fixed when it optimizes, the choice of output for firm j satisfies

$$p + y_j \frac{dp}{dy} - c = 0. \tag{9.8}$$

Now assume that the firms are identical and each produces the same output level, $\frac{y}{m}$. The first-order condition for choice of output (9.8) can then be re-arranged to obtain the Lerner index

$$\frac{p-c}{p} = \frac{1}{m}\frac{1}{|\varepsilon|}, \tag{9.9}$$

and the oligopoly price is given by

$$p = \mu^{\circ} c, \tag{9.10}$$

where $\mu^{\circ} = \frac{m}{m-(1/|\varepsilon|)} > 1$ is the *oligopoly markup*. Thus, in the presence of several firms in the market, the Lerner index of market power is deflated according to the market share. As for monopoly, the value of the markup is related to the inverse of the elasticity of demand. The Lerner index can be used to show that an oligopoly becomes more competitive as the number of firms in the industry increases. This claim follows from the fact that $\frac{p-c}{p}$ must tend to zero as m tends to infinity. Hence, as the number of firms increases, the Cournot equilibrium becomes more competitive and price tends to marginal cost. The limiting position with an infinite number of firms can be viewed as the idealization of the competitive model.

There is one special case of monopoly for which the equilibrium is efficient. Let the firm be able to charge each consumer the maximum price that the consumer is able to pay. To do so obviously requires the firm to have considerable information about its customers. The consequence is that the firm extracts all consumer surplus and translates it into profit. It will keep supplying the good until price falls to marginal cost and there is no more surplus to extract. So total supply will be equal to that under the competition. This scenario, known as *perfect price discrimination*, results in all the potential surplus in the market being turned into monopoly profit. No surplus is lost due to the monopoly, but all surplus is transferred from the consumers to the firm. Of course, this scenario can only arise with an exceedingly well-informed monopolist.

9.4.2 Incomplete Information

Monopoly inefficiency can also arise from the firm having incomplete information, even in situations where there would be efficiency with complete information. To see this, suppose that a monopolist with constant marginal cost c faces a buyer whose willingness to pay for a unit of the firm's output is v. If there is complete information, the firm and buyer will agree to a price between c and v, and the product will be traded. The surplus from the transaction is shared between the two parties and no inefficiency arises.

The difference that imperfect information can make is that trade will sometimes not take place even though both parties would gain if they did trade. Suppose now that

the monopolist cannot observe v but knows from experience that it is drawn from a distribution $F(v)$, which is the probability that the buyer's valuation is less or equal to v. The function $(1 - F(v))$ is analogous to the expected demand when a purchaser buys at most one unit because the probability that there is a demand at price v is the probability that the buyer's valuation is higher than the price. Assume that there are potential gains from trade so $v > c$ for at least a range of v. Pareto-efficiency requires trade to occur if and only if $v \geq c$.

The monopolist's problem is to offer a price p that maximizes its expected profit (anticipating that the buyer will not accept the offer if $v < p$). This price must fall between c and v for trade to occur. The monopolist sets a price p^* that solves

$$\max_{\{p\}} \underbrace{[1 - F(p)]}_{\text{Probability of trade}} \underbrace{[p - c]}_{\text{Profit if trade}}. \tag{9.11}$$

From the assumption that there is a potential gain from trade, there must be a range of values of v higher than c, and thus it is possible for the monopolist to charge a price in excess of the marginal cost with the offer being accepted. Clearly, the price that maximizes expected profit must be $p^* > c$, so the standard conclusion of monopoly holds that price is in excess of marginal cost. When trade takes place (so a value of v occurs with $c < p^* < v$), the outcome is an efficient trade. However, when a value of v occurs with $c < v < p^*$, trade does not take place. This is inefficient because trade should occur when the benefit exceeds the cost ($v > c$). The effect of the monopolist setting price above marginal cost is to eliminate some of the potential trades.

For instance, assume that the willingness to pay v is uniformly distributed on the interval [0, 1] with the marginal cost $0 < c < 1$. Then the probability that trade takes place at price p (expected demand) is $1 - F(p) = 1 - p$, which gives expected revenue $[1 - F(p)]\, p = [1 - p]\, p$ and marginal revenue $MR = 1 - 2p$. The expected profit is $\pi = [1 - p]\,[p - c]$, and the profit-maximizing pricing satisfies the first-order condition $[1-2p]+c = 0$, which can be re-arranged to give monopoly price of $p^* = \frac{1+c}{2} > c$. The parallel between this monopoly choice under incomplete information and the standard monopoly problem is illustrated in figure 9.1.

9.4.3 Measures of Welfare Loss

It has been shown that the equilibrium of an imperfectly competitive market is not Pareto-efficient, except in the special case of perfect price discrimination. This makes it natural to consider what the degree of welfare loss may actually be. The assessment

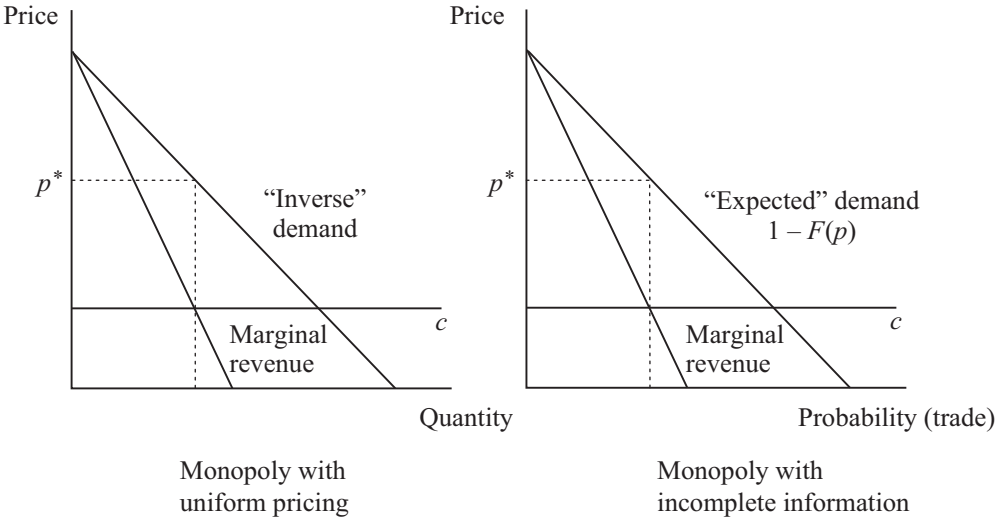

Figure 9.1
Monopoly pricing

of monopoly welfare loss has been a subject of some dispute in which calculations have provided a range of estimates from the effectively insignificant to considerable percentages of potential welfare.

The inefficiency of monopoly will be described in chapter 12 and part of that argument is now briefly provided. Figure 9.2 assumes that the marginal cost of production is constant at value c and that there are no fixed costs. The equilibrium price if the industry were competitive, p^c, would be equal to marginal cost, so $p^c = c$. This price leads to output level y^c and generates consumer surplus ADc. The inverse demand function facing the firm, $p(y)$, determines price as a function of output and is also the average revenue function for the firm. This is denoted by AR. The marginal revenue function is denoted MR. The monopolist's optimal output, y^m, occurs where marginal revenue and marginal cost are equal. At this output level the price with monopoly is p^m. Consumer surplus is ABp^m and profit is $p^m BEc$.

Contrasting the competitive and the monopoly outcomes shows that some of the consumer surplus under competition is transformed into profit under monopoly. This is the area $p^m BEc$, and it represents a transfer from consumers to the firm. However, some of the consumer surplus is simply lost. This loss is the area BDE, which is termed the *deadweight loss of monopoly*. Since the total social surplus under monopoly ($ABp^m + p^m BEc$) is less than that under competition (ADc), the monopoly is inefficient.

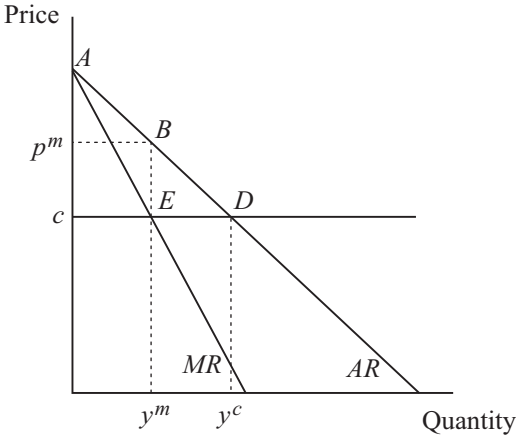

Figure 9.2
Deadweight loss with monopoly

This inefficiency is reflected in the fact that consumption is lower under monopoly than competition.

When the demand function is linear so that the *AR* curve is a straight line, the welfare loss area *BDE* is equal to half of the area $p^m BEc$. The area $p^m BEc$ is monopoly profit, which is equal to $[p^m - c]y^m$. This implies that the loss *BDE* is $\frac{1}{2}[p^m - c]\, y^m$. From the first-order condition for the choice of monopoly output, (9.5), $p^m - c = -\frac{1}{\varepsilon} p^m$. By this result it follows that a measure of the deadweight loss is

$$\text{Deadweight loss} = -\frac{p^m y^m}{2\varepsilon} = -\frac{R^m}{2\varepsilon}, \tag{9.12}$$

where R^m is the total revenue of the monopolist. This formula is especially simple to evaluate to obtain an idea of the size of the deadweight loss. For example, if the elasticity of demand is -2, then the welfare loss is 25 percent of sales revenue and is therefore quite large.

Numerous studies have been published that provide measures of the degree of monopoly welfare loss. A selection of these results is given in table 9.2. The smaller values are obtained by calculating only the deadweight loss triangle. If these were correct, then we could conclude that monopoly power is not a significant economic issue. This was the surprising conclusion of the initial study by Harberger in 1954, as it challenged the conventional wisdom that monopoly must be damaging to the economy. In contrast, the larger values of loss are obtained by including the costs of defending the

Table 9.2
Monopoly welfare loss

Author	Sector	Welfare loss (%)
Harberger	US manufacturing	0.08
Gisser	US manufacturing	0.11–1.82
Peterson and Connor	US food manufacturing	0.16–5.15
Masson and Shaanan	37 US industries	3
		16
McCorriston	UK agricultural inputs	1.6–2.5
		20–40
Cowling and Mueller	United States	4–13
	United Kingdom	3.9–7.2

monopoly position. Chapter 12 considers the arguments proposed in the rent-seeking literature for the inclusion of these additional components of welfare loss. These values reveal monopoly loss to be very substantial.

It can be appreciated from table 9.2 that a broad range of estimates of monopoly welfare loss have been produced. Some studies conclude that welfare loss is insignificant; others conclude that it is very important. What primarily distinguishes these differing estimates is whether it is only the deadweight loss that is counted or the deadweight loss plus the cost of defending the monopoly. Which one is correct is an unresolved issue that involves two competing perspectives on economic efficiency.

There is one further point that needs to be made. The calculations above have been based on a *static* analysis in which there is a single time period. The demand function, the product traded, and the costs of production are all given. The firm makes a single choice and equilibrium is attained. What this ignores are all the *dynamic* aspects of economic activity such as investment and innovation. When these factors are taken into account, as Schumpeter forcefully argued, it is even possible for a monopoly to generate dynamic welfare gains rather than losses. This claim is based on the argument that investment and innovation will only be undertaken if firms can expect to earn a sufficient return. In a competitive environment, all gains are competed away, so the incentives to innovate are eliminated. Conversely, holding a monopoly position allows gains to be realized. This creates the incentive to invest and innovate. Furthermore the incentive is strengthened by the monopoly's desire to maintain its strong hold. Dynamic gains can more than offset static losses, reinforcing the argument for allowing monopoly. We return to this issue in the discussion of regulation in section 9.7.

9.5 Tax Incidence

The study of tax incidence is about determining the changes in prices and profits that follow the imposition of a tax. The *formal* or *legal incidence* of a tax refers to who is legally responsible for paying the tax. The legal incidence can be very different from the *economic incidence*, which relates to who ultimately has to alter his behavior because of the tax.

To see this distinction, consider the following example. A tax of \$1 is levied on a commodity that costs \$10, and this tax must be paid by the retailer. The legal incidence is simple: for each unit sold the retailer must pay \$1 to the tax authority. The economic incidence is much more complex. The first question has to be: What does the price of the commodity become after the tax? It may change to \$11, but this would be an exception rather than the norm. It may, for example, rise instead only to \$10.50. If it does, \$0.50 of the tax falls on the consumer to pay. What of the other \$0.50? This depends on how the producer responds to the tax increase. The producer may lower the price of the commodity to the retailer from \$9 to \$8.75 and then bear \$0.25 of the tax. The remaining \$0.25 of the tax is paid by the retailer. The economic incidence of the tax is then very distinct from the legal incidence.

This example raises the question of what determines the economic incidence. The answer is found in the demand and supply curves for the good that is taxed. Economic incidence will first be determined for the competitive case, and then it is shown how the conclusions are modified by imperfect competition. Imperfect competition can in fact result in very interesting conclusions concerning tax incidence.

Tax incidence analysis is at its simplest when there is competition and the marginal cost of production is constant. In this case the supply curve in the absence of taxation must be horizontal at a level equal to marginal cost; see figure 9.3. This gives the before-tax price $p = c$. The introduction of a tax of amount t will raise this curve by exactly the amount of the tax. The after-tax price, q, is at the intersection of the demand curve and the new supply curve. It can be seen that $q = p + t$, so price will rise by an amount equal to the tax. Hence the tax is simply passed forward by the firms onto consumers, since price is always set equal to marginal cost plus tax.

When marginal cost is not constant and the supply curve slopes upward, the introduction of a tax still shifts the curve vertically upward by the amount equal to the tax. The extent to which price rises is then determined by the slopes of the supply and demand curve. If the demand curve is vertical, price rises by the full amount of the tax; otherwise, it will rise by less; see figure 9.4.

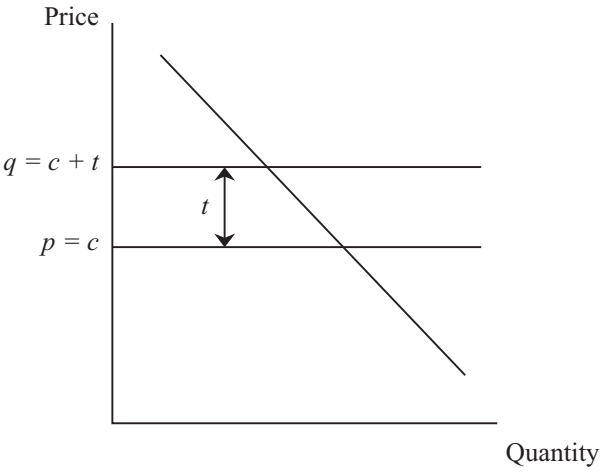

Figure 9.3
Tax incidence with perfectly elastic supply

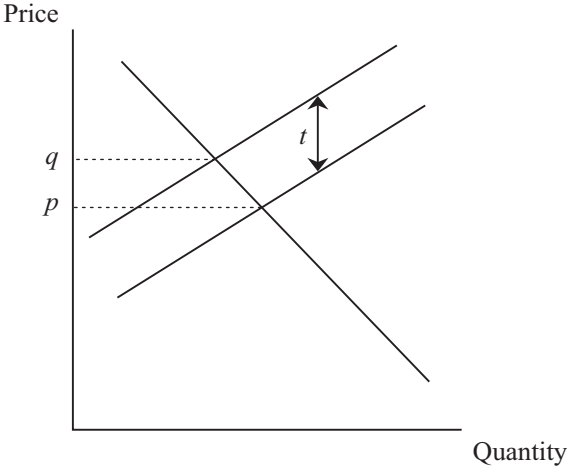

Figure 9.4
Tax incidence in the general case

In summary, if the supply curve is horizontal (so supply is infinitely elastic) or the demand curve is vertical (so demand is completely inelastic), then price will rise by exactly the amount of the tax. In all other cases it will rise by less, with the exact rise being determined by the elasticities of supply and demand. When the price increase is equal to the tax, the entire tax burden is passed by the firm onto the consumers. Otherwise, the burden of the tax is shared between firms and consumers. Consequently the extent to which the price is shifted forward from the producer onto the consumers is dependent on the elasticities of supply and demand.

There are two reasons why tax incidence with imperfect competition is distinguished from the analysis for the competitive case. First, prices on imperfectly competitive markets are set at a level above marginal cost. Second, imperfectly competitive firms may also earn nonzero profits, so taxation can affect profit. To trace the effects of taxation, it is necessary to work through the profit-maximization process of the imperfectly competitive firms. Such an exercise involves characterizing the optimal choices of the firms and then seeing how they are affected by a change in the tax rate.

The incidence of a tax on output can be demonstrated by returning to the diagram for monopoly profit maximization. A tax of value t on output changes the tax-inclusive marginal cost from c to $c + t$. In figure 9.5 this is shown to move the intersection between the marginal revenue curve and the marginal cost curve from a to b. Output falls from y^o to y^t, and price rises from p to q. In this case price rises by less than the

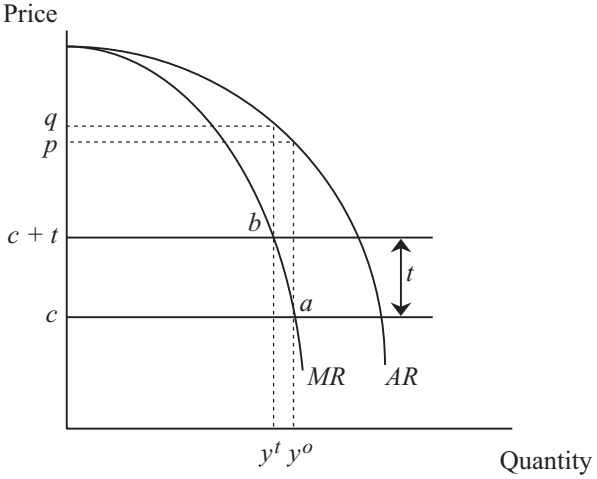

Figure 9.5
Tax undershifting

tax imposed—the difference between q and p is less than t. This is called the case of tax *undershifting*. What it means is that the monopolist is absorbing some of the tax and not passing it all on to the consumer.

With competition, the full value of the tax may be shifted to consumers but never more. With monopoly, the proportion of the tax that is shifted to consumers is determined by the shape of the *AR* curve (and hence the *MR* curve). In contrast to competition, for some shapes of *AR* curve it is possible for the imposition of a tax to be met by a price increase that exceeds the value of the tax. This is the case of tax *overshifting* and is illustrated in figure 9.6. The imposition of the tax, t, leads to a price increase from p to q. As is clear in the figure, $q - p > t$. This outcome could never happen in the competitive case.

The feature that distinguishes the cases of overshifting and undershifting is the shape of the demand function. Figure 9.5 has a demand function that is convex—it becomes increasingly steep as quantity increases. In contrast, figure 9.6 involves a concave demand function with a gradient that decreases as output increases. Either of these shapes for the demand function is entirely consistent with the existence of monopoly.

The overshifting of taxation is also a possibility with oligopoly. To illustrate this, consider the constant elasticity demand function $X = p^{\varepsilon}$, where $\varepsilon < 0$ is the elasticity of demand. Since the elasticity is constant, so must be the markup at $\mu^o = \frac{m}{m-(1/|\varepsilon|)}$. Furthermore, because $\varepsilon < 0$ it follows that $\mu^o > 1$. Applying the markup to marginal

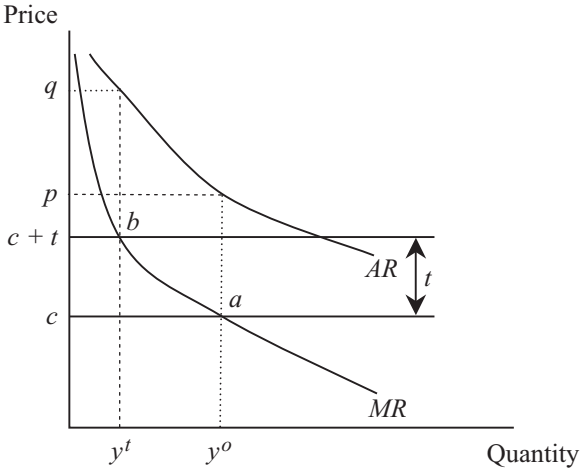

Figure 9.6
Tax overshifting

Table 9.3
Calculations of tax shifting

Baker and Brechling	UK beer 0.696	UK tobacco 0.568
Delipalla and O'Connell, tobacco	"Northern" EU 0.92	"Southern" EU 2.16
Tasarika, beer	UK 0.665	

cost plus tax obtains the equilibrium price of the oligopoly, $q = \mu^o[c + t]$. The effect of an increase in the tax is then

$$\frac{\partial q}{\partial t} = \mu^o > 1, \tag{9.13}$$

so there is always overshifting with the constant elasticity demand function. This holds for any value of $m \geq 1$, and hence applies to both monopoly ($m = 1$) and oligopoly ($m \geq 2$). In addition, as m increases and the market becomes more competitive, μ^o will tend to 1, as will $\frac{\partial q}{\partial t}$, so the competitive outcome of complete tax shifting will arise.

Some estimates of the value of the tax-shifting term are given in table 9.3 for the beer and tobacco industries. Both of these industries have a small number of dominant firms and an oligopolistic market structure. The figures show that although undershifting arises in most cases, there is evidence of overshifting in the tobacco industry.

There is an even more surprising effect that can occur with oligopoly: an increase in taxation can lead to an increase in profit. The analysis of the constant elasticity case can be extended to demonstrate this result. Since the equilibrium price is $q = \mu^o[c + t]$, we use the demand function to obtain the output of each firm as

$$x = \frac{[\mu^o]^\varepsilon\,[c + t]^\varepsilon}{m}. \tag{9.14}$$

Using these values for price and output results in a profit level for each firm of

$$\pi = \frac{[\mu^o - 1]\,[\mu^o]^\varepsilon\,[c + t]^{\varepsilon+1}}{m}. \tag{9.15}$$

The effect of an increase in the tax on the level of profit is then given by

$$\frac{\partial \pi}{\partial t} = \frac{[\mu^o - 1]\,[\mu^o]^\varepsilon\,[\varepsilon + 1]\,[c + t]^\varepsilon}{m}. \tag{9.16}$$

The possibility of the increase in tax raising profit follows by observing that if $\varepsilon > -1$, then $[\varepsilon + 1] > 0$, so $\frac{\partial \pi}{\partial t} > 0$. When the elasticity satisfies this restriction, an increase

in the tax will raise the level of profit. Put simply, the firms find the addition to their costs to be profitable.

It should be observed that such a profit increase cannot occur with monopoly because a monopolist must produce on the elastic part of the demand curve with $\varepsilon < -1$. With oligopoly the markup remains finite provided that $m - \frac{1}{|\varepsilon|} > 0$ or $\varepsilon < -\frac{1}{m}$. Therefore profit can be increased by an increase in taxation if there is oligopoly.

The mechanism that makes this outcome possible is shown in figure 9.7, which displays the determination of the Cournot equilibrium for a duopoly. The figure is constructed by first plotting the isoprofit curves. The curves denote sets of output levels for the two firms that give a constant level of profit. The profit of firm 1 is highest on the curves closest to the horizontal axis, and it reaches its maximum at the output level, m_1, which is the output firm 1 would produce if it were a monopolist. Similarly the level of profit for firm 2 is higher on the isoprofit curves closest to the vertical axis, and is maximized at its monopoly output level, m_2. The assumption of Cournot oligopoly is that each firm takes the output of the other as given when it maximizes. So for any fixed output level for firm 2, firm 1 will maximize profit on the isoprofit curve that is horizontal at the output level of firm 2. Connecting the horizontal points gives the best-reaction function for firm 1, which is labeled $r^1(y_2)$. Similarly, setting a fixed output level for firm 1, we have that firm 2 maximizes profit on the isoprofit curve that is vertical at this level of 1's output. Connecting the vertical points gives its best-reaction function $r^2(y_1)$.

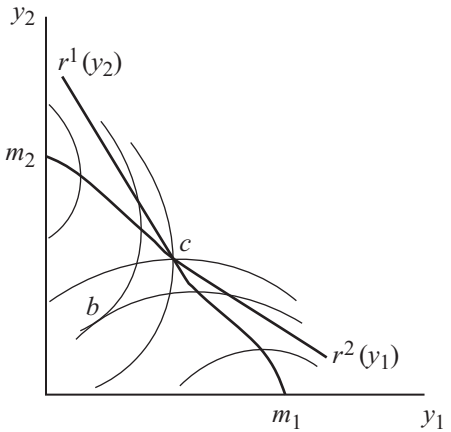

Figure 9.7
Possibility of a profit increase

The Cournot equilibrium for the duopoly is where the best-reaction functions cross, and the isoprofit curves are locally horizontal for firm 1 and vertical for firm 2. This is point c in the figure. The Cournot equilibrium is not efficient for the firms, and a simultaneous reduction in output by both firms, which would be a move from c in the direction of b, would raise both firms' profits. Further improvement in profit can be continued until the point that maximizes joint profit, $\pi_1 + \pi_2$, is reached. Joint profit maximization occurs at a point of tangency of the isoprofit curves, which is denoted by point b in figure 9.7. The firms could achieve this point if they were to collude, but such collusion would not be credible because both the firms would have an incentive to deviate from point b by increasing output.

It is this inefficiency that opens the possibility for a joint increase in profit to be obtained. Intuitively, how taxation raises profit is by shifting the isoprofit curves in such a way that the duopoly equilibrium moves closer to the point of joint profit maximization. Although total available production must fall as the tax increases, the firms secure a larger fraction of the gains from trade. Unlike collusion, the tax is binding on the firms and produces a credible reduction in output.

9.6 Specific and Ad valorem Taxation

The analysis of tax incidence has so far considered only *specific* taxation. With specific taxation, the legally responsible firm has to pay a fixed amount of tax for each unit of output. The amount that has to be paid is independent of the price of the commodity. Consequently the price the consumer pays is the producer price plus the specific tax. This is not the only way in which taxes can be levied. Commodities can alternatively be subject to *ad valorem* taxation so that the tax payment is defined as a fixed proportion of the producer price. Consequently, as price changes, so does the amount paid in tax.

The fact that tax incidence has been analyzed only for specific taxation is not a limitation when firms are competitive, since the two forms are entirely equivalent. The meaning of equivalence here is that a specific tax and an ad valorem tax that lead to the same consumer price will raise the same amount of tax revenue. Their economic incidence is therefore identical.

This equivalence can be shown as follows. Let t be the specific tax on a commodity. Then the equivalent ad valorem tax rate τ must satisfy the equation

$$q = p + t = [1 + \tau]\, p. \tag{9.17}$$

Solving this equation, we have that $\tau = \frac{t}{p}$ is the ad valorem tax rate that leads to the same consumer price as the specific tax. In terms of the incidence diagrams, both taxes would shift the supply curve for the good in exactly the same way. The demonstration of equivalence is completed by showing that the taxes raise identical levels of tax revenue. The revenue raised by the ad valorem tax is $R = \tau p X$. Using the fact that $\tau = \frac{t}{p}$, we can write this revenue level as $\frac{t}{p} p X = t X$, which is the revenue raised by the specific tax. This completes the demonstration that the specific and ad valorem taxes are equivalent.

With imperfect competition this equivalence between the two forms of taxation breaks down: specific and ad valorem taxes that generate the same consumer price generate different levels of revenue. The reason for this breakdown of equivalence, and its consequences, are now explored.

The fact that specific and ad valorem taxes have different effects can be seen very easily in the monopoly case. Assume that the firm sells at price q and that each unit of output is produced at marginal production cost, c. With a specific tax the consumer price and producer price are related by $q = p + t$. This allows the profit level with a specific tax to be written as

$$\pi = [q - t]\, x - cx = qx - [c + t]\, x. \tag{9.18}$$

The expression for this profit level shows that the specific tax acts as an addition to the marginal cost for the firm. Now consider instead the payment of an ad valorem tax at rate τ. Since an ad valorem tax is levied as a proportion of the producer price, the consumer price and producer price are related by $q = [1 + \tau]\, p$; hence the consumers pay price q and the firm receives $p = \frac{1}{1+\tau} q$. The profit level with the ad valorem tax is then

$$\pi = \frac{1}{1 + \tau} qx - cx. \tag{9.19}$$

The basic difference between the two taxes can be seen by comparing these alternative specifications of profit. From the perspective of the firm, the specific tax raises marginal production cost from c to $c + t$. In contrast, the ad valorem tax reduces the revenue received by the firm from qx to $\frac{1}{1+\tau} qx$. Hence the specific tax works via the level of costs, whereas the ad valorem tax operates via the level of revenue. With competition this difference is of no consequence. But the very basis of imperfect competition is that the firms recognize the effect their actions has on revenue—so the ad valorem tax interacts with the expression of monopoly power.

The consequence of this difference is illustrated in figure 9.8. In the left-hand panel, the effect of a specific tax is shown. In the right-hand panel, the effect of an ad valorem

tax is shown. The specific tax leads to an upward shift in the tax-inclusive marginal cost curve. This moves the optimal price from p to q. The ad valorem tax leads to a downward shift in average and marginal revenue net of tax as shown in figure 9.8. The ad valorem tax leads from price p in the absence of taxation to q with taxation. The resulting price increase is dependent on the slope of the marginal revenue curve.

What is needed to make a firm comparison between the effects of the two taxes is some common benchmark. The benchmark chosen is a given consumer price. The values of the specific and ad valorem taxes that lead to this consumer price are found. The taxes are then contrasted by determining which raises the most tax revenue. This comparison is easily conducted by returning to the definition of profit in (9.19). With the ad valorem tax, the profit level can be expressed as

$$\pi = \frac{1}{1+\tau} qx - cx = \frac{1}{1+\tau} [qx - [c + \tau c]\, x]. \tag{9.20}$$

The second term of (9.20) shows that the ad valorem tax is equivalent to the combined use of a specific tax of value τc plus a profit tax at rate $\frac{\tau}{1+\tau}$. A profit tax has no effect on the firm's choice, but it does raise revenue. Hence an ad valorem tax with its rate is set so that

$$\tau c = t \tag{9.21}$$

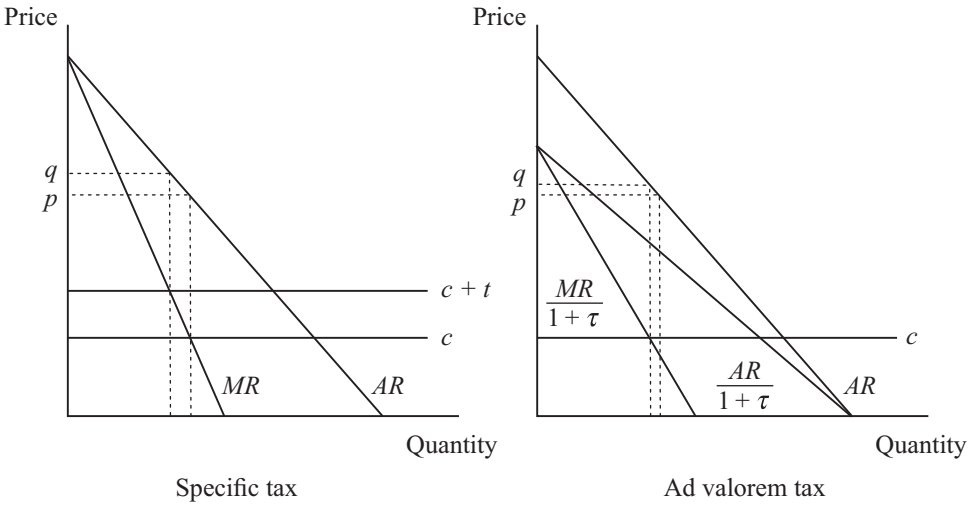

Figure 9.8
Contrasting taxes

leads to the same after-tax price as the specific tax. However, the ad valorem tax raises more revenue. This is because the component τc collects the same revenue as the specific tax t, but the ad valorem tax also collects revenue from the profit-tax component. Hence the ad valorem tax must collect more revenue for the same consumer price. This result can, alternatively, be expressed as the fact that for a given level of revenue, an ad valorem tax leads to lower consumer price than a specific tax.

In conclusion, ad valorem taxation is more effective than specific taxation when there is imperfect competition. The intuition behind this conclusion is that the ad valorem tax lowers marginal revenue, and this reduces the perceived market power of the firm. Consequently the ad valorem tax has the helpful effect of reducing monopoly power, offsetting some of the costs involved in raising revenue through commodity taxation.

9.7 Regulation of Monopoly

Up until this point the focus has been placed on the welfare loss caused by imperfect competition and on tax incidence. As we have shown, there are two competing views about the extent of the welfare loss, but even if the lower values are accepted, it is still beneficial to reduce the loss as far as possible. This raises the issue of the range of policies that are available to reduce the adverse effects of monopoly.

When faced with imperfect competition, the most natural policy response is to encourage an enhanced degree of competition. There are several ways in which this can be done. The most dramatic example is US antitrust legislation, which has been used to enforce the division of monopolies into separate competing firms. This policy was applied to the Standard Oil Company, which was declared a monopoly and broken up into competing units in 1911. More recently the Bell System telephone company was broken up in 1984. This policy of breaking up monopolists represents extreme legislation and, once enacted, leaves a major problem of how the system should be organized following the breakup. Typically the industry will require continuing regulation, a theme to which we return below.

Less dramatic than directly breaking up firms is to provides aids to competition. A *barrier to entry* is anything that allows a monopoly to sustain its position and prevent new firms from competing effectively. Barriers to entry can be legal restrictions such as the issue of a single license permitting only one firm to be active. They can also be technological in the sense of superior knowledge, the holding of patents, or the structure of the production function. Furthermore some barriers can be erected deliberately by the

incumbent monopolist specifically to deter entry. For a policy to encourage competition, it must remove or at least reduce the barriers to entry. The appropriate policy response depends on the nature of the barrier.

If a barrier to entry is created by a legal restriction, it can equally be removed by a change to the law. But here it is necessary to inquire as to why the restriction was created initially. One possible answer would take us to the concept of rent-creation, which is discussed in chapter 12. In that chapter the introduction of a restriction is seen as a way of generating rent. An interesting example of the creation of such restrictions are the activities of MITI (the Ministry of International Trade and Industries) in Japan. In 1961 MITI produced its "Concentration Plan," which aimed to concentrate the mass-production automakers into two to three groups. The intention behind this was to sustain the international competition that ensued after the liberalization of auto imports into Japan and to place the Japanese car industry in a stronger position for exporting. These intentions were never fully realized, and the plan was ultimately undermined by developments in the auto industry, especially the emergence of Honda as a major manufacturer. Despite this, the example still stands as a good illustration of a deliberate policy attempt to restrict competition.

If barriers to entry relate to technological knowledge, then it is possible for the government to insist on the sharing of this knowledge. Both the concerns over the bundling of Internet Explorer with Windows in the United States and the bundling of Media Player with Windows in Europe are pertinent examples. In the United States the outcome has been that Microsoft is obliged to provide rival software firms with information that allows them to develop competing products, and to ensure that these products work with the Windows operating system. Microsoft's rivals are pushing for a similar solution in the European Union. The existence of patents to protect the use of knowledge is also a barrier to entry. The reasoning behind patents is that they allow a reward for innovation: new discoveries are only valuable if the products in which they are embedded can be exploited without competitors immediately copying them. The production of generic drugs is one of the better-known examples of product copying. Without patents, the incentive to innovate would be much reduced and aggregate welfare would fall. The policy issue then becomes the choice of the length of a patent. It must be long enough to allow innovation to be adequately rewarded but not so long that it stifles competition. Current practice in the United States is that the term of a patent is twenty years from the date at which the application is filed.

Barriers to entry can also be erected as a deliberate part of a corporate strategy designed to deter competitors. Entry barriers can be within the law, such as sustained advertising campaigns to build brand loyalty or the building of excess capacity to deter

entry, or they can be illegal such as physical intimidation, violence, and destruction of property. Obviously the latter category can be controlled by recourse to the law if potential competitors wish to do so. Potentially limitations could be placed on advertising. The limitations on tobacco advertisements is an example of such a policy, but this has been motivated on health grounds and not competition reasons. The role of excess capacity is to provide a credible threat that the entry of a competitor will be met by an increase in output from the incumbent with a consequent reduction in market price. The reduction in price can make entry unprofitable, so sustaining the monopoly position. Although the economic reasoning is clear, it is difficult to see how litigation could ever demonstrate that excess capacity was being held as an entry deterrent, and this limits any potential policy response.

The enhancement of competition only works if it is possible for competitors to be viable. The limits of the argument that monopoly can be tackled by the encouragement of competition are confronted when the market is characterized by *natural monopoly*. The essence of natural monopoly is that there are increasing returns in production and that the level of demand is such that only a single firm can be profitable. This is illustrated in figure 9.9 where the production technology of the two firms involves a substantial fixed cost but a constant marginal cost. Consequently the average cost curve, denoted AC, is decreasing while the marginal cost curve, MC, is horizontal. When there is a monopoly, the single firm faces the demand curve AR^1. Corresponding to this average revenue curve is the marginal revenue curve MR^1. The profit-maximizing price for the

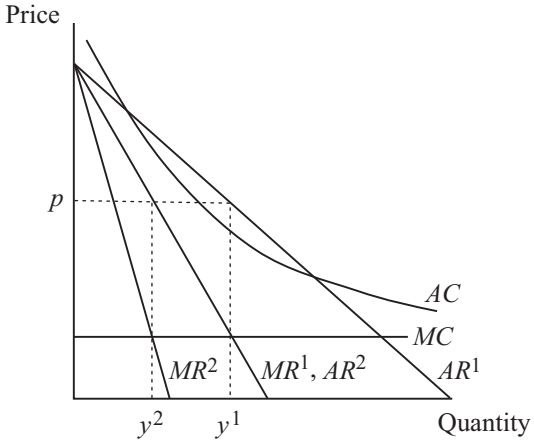

Figure 9.9
Natural monopoly

monopoly is p and output is y^1. It should be observed that the price is above the level of average cost at output y^1, so the monopolist earns a profit.

Now consider the consequence of a second firm entering the market. The cost conditions do not change, so the *AC* and *MC* curves are unaffected. Demand conditions do change since the firms have to share the market. The simplest assumption to make is that the two firms share exactly half the market each. This would hold if the total market consists of two geographical areas each of which could be served by one firm. Furthermore this is the most beneficial situation for the firms since it keeps them from competing. Any other way of sharing the market would lead to them to earning less profit. With the market shared equally, the demand facing each firm becomes AR^2 (equal to the old MR^1) and marginal revenue, MR^2. The profit-maximizing price remains at p, but now at output y^2 this is below average cost. The two firms must therefore both take a loss. Since this market sharing is the most profitable way for the two firms to behave, any other market behavior must lead to an even greater loss.

What this argument shows is that a market in which one firm can be profitable cannot support two firms. The problem is that the level of demand does not generate enough revenue to cover the fixed costs of two firms operating. The examples that are usually cited of natural monopolies involve utilities such as water supply, electricity, gas, telephone, and railways where a large infrastructure has to be in place to support the market and is very costly to replicate. If these markets do conform to the situation in the figure, then without government intervention, only a single firm could survive in the market. Furthermore any policy to encourage competition will not succeed unless the government can fundamentally alter the structure of the industry. It is not enough just to try to get another firm to operate.

The two policy responses to natural monopoly most widely employed have been public ownership and private ownership with a regulatory body controlling behavior. When the firm is run under public ownership, its price should be chosen to maximize social welfare subject to the budget constraint placed on the firm—the resulting price is termed the *Ramsey price*. The budget constraint may require the firm to break even or to generate income above production cost. Alternatively, the firm may be allowed to run a deficit that is financed from other tax revenues. Assume that all other markets in the economy are competitive. The Ramsey price for a public firm subject to a break-even constraint will then be equal to marginal cost if this satisfies the constraint. If losses arise at marginal cost, then the Ramsey price will be equal to average cost. The literature on public sector pricing has extended this reasoning to situations where marginal cost and demand vary over time such as in the supply of electricity. Doing this leads into

the theory of *peak-load pricing*. When other markets are not competitive, the Ramsey price will reflect the distortions elsewhere in the economy.

Public ownership was practiced extensively in the United Kingdom and elsewhere in Europe. All the major utilities including gas, telephone, electricity, water, and trains were taken into public ownership. This policy was eventually undermined by the problems of the lack of incentive to innovate, invest, or limit costs. Together, these produced a very poor outcome with the lack of market forces producing industries that were overmanned and inefficient. As a consequence the United Kingdom has undertaken a privatization program that has returned all these industries to the private sector.

The treatment of the various industries since the return to private ownership illustrates different responses to the regulation of natural monopoly. The water industry is broken into regional suppliers that do not compete directly but are closely regulated. With telephones, the network is owned by British Telecom, but other firms are permitted access agreements to the network. This can allow them to offer a service without the need to undertake the capital investment. In the case of the railways, the ownership of the track, which is the fixed cost, has been separated from the rights to operate trains, which generates the marginal cost. Both the track owner and the train operators remain regulated. With gas and electricity, competing suppliers are permitted to supply using the single existing network.

The most significant difference between public ownership and private ownership with regulation is that under public ownership the government is as informed as the firm about demand and cost conditions. This allows the government to determine the behavior of the firm using the best available information. Policy can only maximize the objective function in an expected sense. So, although the available information may not be complete, the best that is possible will be achieved. As an alternative to public ownership, a firm may remain under private ownership but be made subject to the control of a regulatory body. This introduces possible asymmetries in information between the firm and the regulator. Faced with limited information, one approach considered in the theoretical literature is for the regulator to design an incentive mechanism that achieves a desirable outcome. An example of such a regulatory scheme is the two-part tariff in which the payment for a commodity involves a fixed fee to permit consumption followed by a price per unit of consumption, with these values being set by the regulator. Alternatively, the regulator may impose a constraint on some observable measure of the firm's activities such as that it must not exceed a given rate of return on the capital employed. Even more simple are the regulatory schemes in the United Kingdom that involve restricting prices to rise at a slower rate than an index of the general price level.

The analysis has looked at a range of issues concerned with dealing with monopoly power and how to regulate industries. The essence of policy is to move the economy closer to the competitive outcome, but there can be distinct problems in achieving this. Monopoly can arise because of the combination of cost and demand conditions, and this can place limitations on what policies are feasible. Natural monopoly results in the need for regulation.

9.8 Regulation of Oligopoly

9.8.1 Detecting Collusion

In an oligopolistic market firms can collectively act as a monopolist and are consequently able to increase their prices. The problem for a regulatory agency is that such collusion is often tacit and so difficult to detect. However, from an economic viewpoint there is no real competition, and a high price is the prima facie evidence of collusion. The practical question for the regulator is whether a high price is the natural outcome of competition in a market where there is significant product differentiation (and so little pricing constraint from substitute products) or whether it reflects price collusion.

Nevo (2001) studied this question for the breakfast cereal industry where the four leaders Kellogg, Quaker, General Mills, and Post were accused by Congressman Chuck Schumer (March 1995) of charging "caviar prices for cornflakes quality." After estimating price elasticities of demand for each brand of cereal, Nevo (2001) used these price elasticities to calculate the Lerner index for each brand, $\frac{p-c}{p}$, that would prevail in the industry if producers were colluding and acting as a monopolist. Nevo then calculated the Lerner index for each brand if producers were really competing with each other.

Given the estimated demand elasticities, Nevo found that with collusion, the Lerner index of each brand would be on average around 65 to 75 percent. With the firms competing, the Lerner index would be on average around 40 to 44 percent. The next step was to compare these estimates of the Lerner index for the hypothetical collusive and competing industry with the actual Lerner index for the breakfast cereal industry to see which hypothesis is the most likely. According to Nevo, the actual Lerner index for the breakfast cereal market was about 45 percent in 1995. This market power index is far below the 65 to 75 percent hypothetical Lerner index that would prevail in a colluding industry and much closer to the Lerner index in the competing hypothesis. Nevo concludes that market power is significant in this industry, not because of collusion

but because of product differentiation that limits competition from substitute products (after all, what is the substitute for a "healthy" cereal breakfast?).

9.8.2 Merger Policy

In its recent reform of merger regulation, the European Commission has recognized that in oligopolistic markets a merger may harm competition and consequently increase prices. Under the original European Commission Merger Regulation (ECMR) a merger was incompatible with the common market if and only if it "creates or strengthens a dominant position as a result of which competition would be significantly impeded." The problem with this two-part cumulative test was that unless a merger was likely to create or strengthen a dominant position, the question of whether it could lessen competition did not arise and so could not be used to challenge a merger. However, one can easily think of oligopoly situations where a merger would substantially lessen competition without giving any individual firm a dominant position. Moreover the concept of dominance is not easily established especially in the presence of tacit collusion. In practice, the concept of dominance had different meanings depending on the circumstances. In particular, when there was some presumption of collusion, the European Commission could use the concept of "collective" dominance, taking as a single unit a group of sellers suspected to collude in their pricing policy. Just as Alice said in *Through the Looking Glass*, the question comes to "whether you *can* make words mean so many different things."

In the 2004 reform of merger policy the European Commission shifted the attention to the second part of the original regulation. The key article in the new ECMR says that "a concentration which would significantly impede effective competition, in the common market or in a substantial part of it, in particular as a result of the creation or strengthening of a dominant position, shall be declared incompatible with the common market" (Article 2). Thus the European Commission has recognized that reducing competition is not necessarily dominance but rather a result of how much competition is left. The fundamental idea is that in oligopolistic markets a merger of two or more rivals raises competitive concerns if the merging firms sell products that are close substitutes. By removing the competitive constraint, merging firms are able to increase their prices. This is the "unilateral effect" theory of competitive harm that has been commonly used in the US merger regulation.

Economists have developed a large number of simulation methods, mostly based on estimated demand elasticities, to determine the possible change in price resulting from a merger. Simulation models combine market data on market shares, the own-price

elasticity of demand, and the cross-price elasticities of demand with a model of firm behavior and anticipated reductions in cost from the merger to predict the likely price effects. A practical example will be useful to illustrate the method. The example is drawn from Hausman and Leonard (1997) and concerns the market for bath tissue. In 1995 the producer of the Kleenex brand acquired the producer of two competing brands (Cottonelle and ScotTissue). The market shares for these products and other brands are shown in table 9.4.

Using weekly retail scanner data that tracks household purchases in retail stores in major US cities, it was possible to estimate own-price elasticities as shown in table 9.4. The key cross-price elasticities were estimated to be 0.19 (Kleenex relative to Cottonelle), 0.18 (Kleenex relative to ScotTissue), 0.14 (Cottonelle relative to Kleenex), and 0.06 (ScotTissue relative to Kleenex). In addition it was anticipated that the acquisition would reduce the marginal cost of production for ScotTissue, Cottonelle, and Kleenex by 4, 2.4, and 2.4 percent respectively. With these estimates of demand elasticities, information about market shares, and the anticipated cost saving from the acquisition of Cottonelle and ScotTissue by the Kleenex brand, it was possible to evaluate the price effects of the merger. A simulation model based on these market estimates and other assumptions about firm and market behavior (Nash equilibrium and constant marginal costs) produced the following prices changes. The acquisition would lead to a reduction in the price of ScotTissue and Cottonelle by 2.6 and 0.3 percent respectively, and an increase in the price of Kleenex by 1.0 percent. Not surprisingly, the Antitrust did not challenge the merger.

Table 9.4
Estimating the effect of merger in the bath tissue market

Bath tissue brand	Market share	Own-price elasticity	Price change [cost change]	
Kleenex	7.5%	−3.38	+1.0%	[−2.4%]
Cottonelle	6.7	−4.52	−0.3	[−2.4]
ScotTissue	16.7	−2.94	−2.6	[−4.0]
Charmin	30.9	−2.75		
Northern	12.4	−4.21		
Angel Soft	8.8	−4.08		
Private label	7.6	−2.02		
Other	9.4	−1.98		
Market demand		−1.17		

Source: Data from tables 1 and 2 in Hausman and Leonard (1997).

9.9 Unions and Taxation

As well as monopoly on product markets, it is possible to have unions creating market power for their members on input markets. By organizing labor into a single collective organization, unions are able to raise the wage above the competitive level and generate a surplus for their members. The issue of tax incidence is also of interest when there are unions, since they can employ their market power to reduce the effect of a tax on the welfare of members.

The role of trade unions is to ensure that they secure the best deal possible for their members. In achieving this, the union faces a trade-off between the wage rate and the level of employment, since a higher wage will invariably lead to lower employment. This trade-off has to be resolved by the union's preferences.

A standard way of representing the preferences of a union is to assume that it has a fixed number, m, of members. Each employed member receives a wage $w[1-t]$, where t is the tax on wage income. The unemployed members receive a payment of b, which can represent either unemployment benefit or the payment in a nonunionized occupation. The level of employment is determined by a labor demand function $n(w)$, with higher values of w leading to lower levels of employment. If the wage rate is w, the probability of any particular member being employed and receiving $w[1-t]$ is $\frac{n(w)}{m}$. Consequently, if all members are assumed to have the same preferences, the expected utility of a typical union member is

$$U = \frac{n(w)}{m}u(w[1-t]) + \frac{m-n(w)}{m}u(b). \tag{9.22}$$

Since all union members have identical preferences, this utility function can also be taken to represent the preferences of the union.

The union chooses the wage rate to maximize utility, so that the chosen wage satisfies the first-order condition

$$n'(w)[u(w[1-t]) - u(b)] + n(w)[1-t]\,u'(w[1-t]) = 0. \tag{9.23}$$

The interpretation of this condition is that the optimal wage rate balances the marginal utility of a higher wage against the value of the marginal loss of employment. Now define the elasticity of labor demand by $\varepsilon_n = \frac{\Delta n}{\Delta w}\frac{w}{n} < 0$ and the elasticity of utility by $\varepsilon_u = \frac{\Delta u}{\Delta w[1-t]}\frac{w[1-t]}{u} > 0$. The first-order condition (9.23) can then be written as

$$u(w[1-t]) = \mu^u u(b), \tag{9.24}$$

where $\mu^u = \frac{1}{1-(\varepsilon_u/|\varepsilon_n|)} > 1$ is the union markup relating the utility of an employed member to that of an unemployed member. This markup is a measure of the unions market power. Given a value for the utility elasticity, ε_u, the markup increases the lower is the elasticity of labor demand, ε_n. As labor demand becomes perfectly elastic, as it does if the labor market is perfectly competitive, then μ^u tends to 1, and the union can achieve no advantage for its members.

The incidence of taxation can now be determined. To simplify, assume that the two elasticities—and hence the markup—are constant. Then the utility of the after-tax wage must always bear the same relation to the utility of unemployment benefit. Consequently $w[1-t]$ must be constant whatever the tax rate. This can only be achieved if the union negates any tax increase by securing an increase in the wage rate that exactly offsets the tax change. Consequently those who retain employment are left unaffected by the tax change, but since the wage has risen, employment must fall. Overall, the union members must be worse off. This argument can easily be extended to see that if the elasticities are not constant, there is the potential for the overshifting, or undershifting, of any tax increase. In this respect tax incidence with trade unions has very similar features to incidence with monopoly.

9.10 Monopsony

A monopsony market is a market consisting of a single buyer who can purchase from many sellers. The single buyer (or monopsonist) could be a firm that constitutes the only potential buyer of an input. It could also be an individual or public organization that is the only buyer of a product. For example, in many countries the government is the monopsonist in the teaching and nursing markets. In local markets with only one large employer, the local employer might literally be the only employment option in the local community (a coal mine, supermarket, government agency, etc.), so it might make sense that the local employer acts as a monopsonist in reducing the wage below the competitive level. In larger markets with more than one employer, employers association often have opportunities to coordinate their wage offers. This wage coordination allows employers to act as a "demand" cartel in the labor market and thus replicate the monopsony outcome. Just as monopoly results in supply reduction with a price or wage *above* competitive levels, monopsony will result in demand reduction with price or wage *below* competitive levels.

In a perfectly competitive market in which many firms purchase labor services, each firm takes the price of labor as given. Each firm maximizes its profits by choosing the

employment level that equates the marginal revenue product of labor with the wage rate. In contrast, in a monopsony labor market, the monopsony firm pays a wage below the competitive wage. The result is a shortage of employment relative to the competitive level. The idea is that since the marginal revenue product from additional employment exceeds the wage cost in a monopsony labor market the monopsonist employer might want to hire more people at the prevailing wage. However, it would not want to increase the wage to attract more workers because the gain from hiring additional workers (the marginal revenue product) is outweighed by the higher wage bill it would face for its existing workforce.

Figure 9.10 shows the equilibrium in a monopsony labor market. The competitive equilibrium occurs at a market-clearing wage, w^c, where the labor supply curve intersects the demand curve. Suppose now there is a single buyer on this labor market. The marginal revenue of labor is the additional revenue that the firm gets when it employs an additional unit of labor. Suppose that the firm's output as a function of its labor use is $Q(L)$ and that the firm is a price taker on the output market, so its output price p is independent of the amount of output Q. Then the marginal revenue of labor is $MR_L = p\frac{dQ}{dL}$, which is decreasing due to decreasing returns to labor. This marginal revenue is depicted in figure 9.10 as the downward-sloping labor demand curve. The supply of labor is described by the "inverse" supply curve. The inverse supply curve

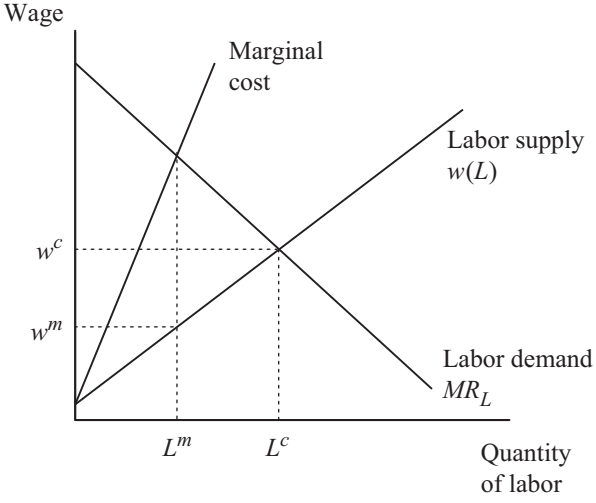

Figure 9.10
Monopsony in the labor market

$w(L)$ describes the wage required to induce any given quantity of labor to be supplied. Since the supply curve is upward sloping, $\frac{dw}{dL} > 0$. The total labor cost of the monopsonist is $Lw(L)$, and the marginal cost of labor is the extra cost that comes from hiring one more worker $MC_L = w + L\frac{dw}{dL}$. This additional cost can be decomposed into two parts: the cost from employing more workers at the existing wage, w, and the cost from raising the wage for all workers, $L\frac{dw}{dL}$. Since $\frac{dw}{dL} > 0$, the marginal labor cost curve lies everywhere above the labor supply curve, as indicated in figure 9.10. The monopsonist will maximize profit, $\pi = pQ(L) - w(L)L$, at the point where the marginal revenue of labor is equal to marginal cost, $p\frac{dQ}{dL} = w + L\frac{dw}{dL}$.

The choice that gives maximum profit occurs in figure 9.10 at the intersection between the marginal cost curve and the labor demand curve, yielding employment level L^m and wage rate w^m. Therefore in a monopsony labor market the monopsony firm pays a wage that is less than the competitive wage with employment level below the competitive level. The monopsony equilibrium condition can also be expressed as an inverse elasticity pricing rule. Indeed the elasticity of labor supply is $\varepsilon_L = \frac{dL}{dw}\frac{w}{L}$ and the profit maximization condition $MR_L = MC_L$ can be re-arranged to give

$$\frac{MR_L - w}{w} = \frac{1}{\varepsilon_L}. \tag{9.25}$$

This inverse pricing rule says that the percentage deviation from the competitive wage is inversely proportional to the elasticity of labor supply. In contrast to monopoly, the key elasticity is the *supply* elasticity. Just as monopoly results in a deadweight loss, so does monopsony leading to underemployment and underpricing of the input (in this case labor) relative to the competitive outcome.

9.11 Conclusions

This chapter has shown how imperfect competition leads to a failure to attain Pareto-efficiency. As with all such failures, this opens a potential role for government intervention to promote efficiency. Estimates of the welfare loss due to imperfect competition vary widely from the almost insignificant to considerable proportions of welfare, depending on the perspective taken upon expenditures on securing the monopoly position. These static losses have to be set against the possible dynamic gains.

Economic tax incidence relates to whoever ultimately has to change his behavior as a consequence of taxation. With competition the outcome is fairly straightforward: the

cost of a commodity tax is divided between producers and consumers, with the division depending on the elasticities of supply and demand. Imperfect competition introduces two additional factors. Taxes may be overshifted so that price rises by more than the value of the tax. In addition an increase in taxation may even raise the profits of firms. In contrast to the competitive case, specific and ad valorem taxation are not equivalent with imperfect competition. In a choice between the instruments, ad valorem taxation is more effective, since it has the effect of reducing perceived monopoly power.

To reduce the welfare loss, policy should attempt to encourage competition. In some circumstances this can work, but when there is natural monopoly, this policy has to be carefully considered. A natural monopoly could be taken into public ownership or run as a private firm with regulation. Recent policy has concentrated on the latter.

Further Reading

The measurement of welfare loss began with:

Harberger, A. C. 1954. Monopoly and resource allocation. *American Economic Review* 45: 77–87.

The other values in table 8.1 are taken from:

Cowling, K. G., and Mueller, D. C. 1978. The social costs of monopoly power. *Economic Journal* 88: 727–48.

Gisser, M. 1986. Price leadership and welfare losses in U.S. manufacturing. *American Economic Review* 76: 756–67.

Masson, R. T., and Shaanan, J. 1984. Social costs of oligopoly and the value of competition. *Economic Journal* 94: 520–35.

McCorriston, S. 1993. The welfare implications of oligopoly in agricultural input markets. *European Review of Agricultural Economics* 20: 1–17.

Peterson, E. B., and Connor, J. M. 1995. A comparison of oligopoly welfare loss estimates for U.S. food manufacturing. *American Journal of Agricultural Economics* 77: 300–308.

The basics of oligopoly theory are covered in:

Waterson, M. 1983. *Oligopoly Theory*. Cambridge: Cambridge University Press.

The analysis of tax incidence with oligopoly can be traced back to:

Seade, J. 1985. Profitable cost increases. Warwick Economic Research Paper 260. University of Warwick, Coventry.

The results in table 8.3 are compiled from:

Baker, P., and Brechling, V. 1992. The impact of excise duty changes on retail prices in the UK. *Fiscal Studies* 13: 48–65.

Delipalla, S., and O'Connell, O. 2001. Estimating tax incidence, market power and market conduct: The European cigarette industry. *International Journal of Industrial Organization* 19: 885–908.

Tasarika, E. 2001. *Aspects of International Taxation*. PhD dissertation. University of Exeter.

Results on comparison of specific and ad valorem tax are in:

Delipalla, S., and Keen, M. 1992. The comparison between ad valorem and specific taxation under imperfect competition. *Journal of Public Economics* 49: 351–67.

Myles, G. D. 1996. Imperfect competition and the optimal combination of ad valorem and specific taxation. *International Tax and Public Finance* 3: 29–44.

The example on detecting collusion is drawn from:

Nevo, A. 2001. Measuring market power in the ready-to-eat breakfast cereal industry. *Econometrica* 69: 307–42.

The merger simulation model for bath tissue is drawn from:

Hausman, J. A., and Leonard, G. K. 1997. Economic analysis of differentiated product mergers using real world data. *George Mason Law Review* 5: 321–46.

A further discussion of merger simulation analysis can be found in:

Epstein, R. J., and Rubinstein, D. L. 2002. Merger simulation: A simplified approach with new applications. *Antitrust Law Journal* 69: 883–919.

A presentation of the various concepts of competition is in:

Vickers, J. 1995. Concepts of competition. *Oxford Economic Papers* 47: 1–23.

A good perspective on the inefficiency resulting from market power with special attention on information problems is:

Vickers, J. 1996. Market power and inefficiency: A contract perspective. *Oxford Review of Economic Policy* 12: 11–26.

The basic and first paper on product differentiation is:

Hotelling, H. 1929. Stability in competition. *Economic Journal* 39: 41–47.

The other classic paper is:

d'Aspremont, C., Gabszewicz, J., and Thisse, J.-F. 1979. On Hotelling's stability in competition. *Econometrica* 17: 1145–51.

An economic analysis of regulation policies with special attention to the United Kingdom is:

Armstrong, M., Cowans, S., and Vickers, J. 1994. *Regulatory Reform: Economic Analysis and British Experience*. Cambridge: MIT Press.

Recent European merger regulation guidelines (28 January 2004) are available at: http://europa.eu.int/comm/competition/mergers/review.

A good account of antitrust law and economics is in:

Scherer, F. M. 1980. *Industrial Market Structure and Economic Performance*. Chicago: Rand McNally.

Exercises

9.1 What should be the objective of a monopoly labor union?

9.2 An industry is known to face a market price elasticity of demand $\varepsilon = -3$. Suppose that this elasticity is approximately constant as the industry moves along its demand curve. The marginal cost in this industry is \$10 per unit, and there are five firms in the industry. What would the Lerner index be at the Cournot equilibrium in this industry?

9.3 Consider a monopolist operating the underground in Europa city with a total cost curve given by $c(x) = 15 + 5x$. The monopolist sets two prices: a high price p_h and a low price p_l. Everyone is eligible for the high price, but only by taking the tube outside the peak hours is anyone eligible for the discount price. Suppose that the only off-peak travelers are those who are not willing to buy the ticket at p_h.

a. If the monopolist faces the inverse demand curve given by $p(x) = 20 - 5x$, what are the profit-maximizing values of p_h and p_l? [*Hint*: Let x_h and x_l denote the high-price and low-price quantities respectively. Then profit for the price discriminating monopolist is $p = p(x_h)x_h + p(x_h + x_l)x_l - c(x_h + x_l)$.]

b. How much economic profit does the monopolist take?

c. How much profit would be made if the same price were charged to all buyers (no price discrimination)? Discuss the difference from part b.

9.4 Demonstrate that monopoly is Pareto-inefficient. Must it always lead to a lower level of social welfare than competition?

9.5 Consider an economy with one good and a linear inverse demand $p(x) = a - bx$. Suppose that there is a single firm operating in this market and that this firm faces a linear cost function $C(x) = cx$ (with $c < a$).

a. Show that the profit maximizing output with monopoly is $x^m = \frac{a-c}{2b}$ and the resulting price is $p^m = \frac{a+c}{2}$.

b. Show that the efficient competitive output level is $x^c = \frac{a-c}{2b} = 2x^m$.

c. Calculate the monopoly profit and the monopoly deadweight loss, and show that these are respectively $\pi^m = \frac{1}{b}\left[\frac{a-c}{2}\right]^2$ and $\lambda^m = \frac{\pi^m}{2}$.

d. Consider a quantity subsidy s to the monopolist so that its cost function is $C(x) = [c-s]x$. Show that a subsidy rate of $s = a - c$ induces the monopolist to produce the efficient amount of output.

e. What is the monopolist's profit resulting from a government intervention imposing marginal cost pricing?

9.6 The inverse demand function for a product is given by $p = a - bX$. The cost function for each firm producing the product is $C(x) = F + cx$.

a. Assume that the industry is monopolized. For what value of F can the monopoly be profitable?

b. Assume that the industry is a Cournot duopoly. For what values of F can both duopolists be profitable?

c. Use the solutions to parts a and b to determine a range of F for which the industry is a natural monopoly.

d. Assume that the two duopolists decide to act collusively. Determine the values of F for which both can be profitable. Does this give higher or lower social welfare than monopoly?

e. The government can either provide a production subsidy to the monopolist (an amount s per unit of output) or subsidize the fixed costs of the Cournot duopolists. Which policy is most cost effective?

9.7 Assume that a monopolist can identify two distinct markets. Find the profit-maximizing prices if the demand functions for the two markets are

$$x_1 = 100 - 2p_1, \quad x_2 = 150 - 3p_2.$$

What is the level of consumer surplus in each market? If the monopolist is forced by legislation to charge a single price, what will this price be? Contrast the level of consumer surplus with and without price discrimination.

9.8 Consider two monopolists operating in separate markets with identical and constant marginal cost. Are the following statements true or false?

a. If both face different linear demand curves that are parallel, the monopolist that will have the higher markup is the one whose demand curve is farther from the origin.

b. If both face linear demand curves with identical vertical intercepts but different slopes, the monopolist with the higher markup is the one with the steeper demand curve.

c. If both face linear demand curves with identical horizontal intercepts but different slopes, the monopolist with the higher markup is the one with the steeper demand curve.

9.9 Discuss how brand promotion can increase inefficiency. Is brand proliferation good or bad?

9.10 Demand is assumed to be unit-elastic: $X(p) = \frac{1}{p}$. There are $m \geq 2$ firms operating in the market with constant marginal cost levels $c_1 \leq c_2 \leq \ldots \leq c_m$. They engage in Cournot competition.

a. Show that the equilibrium price implies Lerner indexes $\frac{p-c_i}{p} = s_i$, where s_i is the market share of firm i.

b. Using the equilibrium price, show that the profit of firm i is equal to $[s_i]^2$.

c. Show that the industry profit is equal to the Herfindahl index $H = \sum_i [s_i]^2$.

d. What is the effect of a specific tax t on equilibrium price? How does this tax affect the industry profit and the Herfindahl index?

9.11 Consider a standard Cournot oligopoly with $n = 2k$ identical firms (with $k \geq 1$), an inverse demand $P(X)$, and a cost function $C(x)$ with no fixed costs. Consider only two possible cases: $C(x)$ convex and $C(x)$ concave. Assume that there is always a unique symmetric equilibrium with per firm output x_k and profit π_k. Assume that there are k two-firm mergers.

a. List all conditions on the primitives of the model such that each firm is better off after these mergers. Explain your answer (no proof needed).

b. Can such a set of mergers be expected to take place without regulatory intervention? Explain.

c. Under what conditions can such a set of mergers increase social welfare?

9.12 Consider a standard Cournot oligopoly with $n \geq 2$ identical firms, $P(x) = a - bX$, $X \geq 0$, and $C(x) = cx^2$.

a. Find the Cournot equilibrium output and profit.

b. If m firms wish to merge, what would be their cost function, assuming that they can use all their m production plants but that they otherwise do not have any efficiency gains as a result of the merger?

c. Given the cost function from part b, when is an m-firm merger profitable to the merged entity? To the nonmerging firms?

d. Give a precise economic intuition explaining your answer relative to the usual (linear cost) case.

9.13 Consider two firms, $i = 1, 2$, producing differentiated products and engaged in Cournot competition. The inverse demand for firm i is given by $p_i = a - bq_i - dq_j$, where q_i is the amount of its own output and q_j is firm j's level of output (with $a > c$, $b > \frac{1}{2}$ and $-1 < d < 1$). Similarly the inverse demand for firm j is given by $p_j = a - bq_j - dq_i$. The goods are substitutes for $d > 0$ and complements for $d < 0$. The marginal cost of each firm is zero.

a. Given the market demands, what are the best-response functions of the two firms?

b. Draw the best-response functions both for complements ($d < 0$) and substitutes ($d > 0$).

c. Compute the Cournot equilibrium quantities and prices in this market.

d. Compare the outcome between substitutes and complements goods.

e. What are the profit-maximizing quantities and prices if firm i is a monopolist in this market? Compare with part c.

9.14 Consider a standard Cournot oligopoly with $n \geq 2$ identical firms, an inverse demand function $n \geq 2$, and cost function $C(x) = K + cx$ if $x > 0$, and 0 if $x = 0$, meaning K is a fixed cost.

a. Find the Cournot equilibrium output and profit. How many firms (as a function of K) can survive at the equilibrium?

b. When is an m-firm merger profitable to the merged entity? To the nonmerging firms?

c. Give a precise economic intuition as to why most mergers are not profitable in the usual model with $K = 0$. How is it different when $K > 0$?

9.15 Consider a homogeneous-good Cournot oligopoly with $n \geq 2$ identical firms with cost $C(x) = 0$ and inverse demand $P(X) = e^{-X}$.

a. Find a firm's best-response function, the Cournot equilibrium output, price, and profit. What type of equilibrium is this?

b. Find all the merger sizes m ($2 \leq m \leq n$) that are profitable to the merged entity. Are these mergers also profitable to the nonmerging firms?

c. Give an economic intuition, and compare it to the case of linear demand.

9.16 Consider Cournot competition with n identical firms. Suppose that the inverse demand function is linear with $P(X) = a - bX$, where X is total industry output, $a, b > 0$. Each firm has a linear cost function of the form $C(x) = cx$, where x stands for per firm output. It is assumed that $a > c$.

a. At the symmetric equilibrium, what are the industry output and price levels? What are the equilibrium per firm output and profit levels? What is the equilibrium social welfare (defined as the difference between the area under the demand function and total cost)?

b. Now let m out of n firms merge. Show that the merger is profitable for the m merged firms if and only if it involves a pre-merger market share of 80 percent.

c. Show that each of the $(n - m)$ nonmerged firms is better off after the merger.

d. Show that the m-firm merger increases industry price and also lowers consumer welfare.

9.17 What is the difference between vertical and horizontal product differentiation? Provide an example of each.

9.18 A monopolist faces the inverse demand function $P(x) = a - bx$ and produces with constant marginal cost c.

a. Determine the effect on equilibrium price of the introduction of a specific tax of value t. Is the tax overshifted?

b. Calculate the effect on profit of the tax. Show that $\frac{d\pi}{dt} = -x$, where x is the equilibrium output level. Explain this result.

c. Now replace the specific tax with an ad valorem tax at rate τ. Find a pair of taxes that lead to the same level of tax revenue. Which gives a lower price?

9.19 A Cournot oligopoly of n firms faces an inverse demand function $p = X^{-1/v}$, where X is aggregate industry output $X = \sum_{i=1}^{n} x_i$. The cost function for each firm is $C(x_i) = cx_i$ and a specific tax t is charged on each unit of output.

a. Show that the output of each firm at the symmetric Cournot equilibrium is $x = \frac{1}{n}\left[\frac{vn-1}{vn[c+t]}\right]^{v}$.

b. Show that the rate of tax shifting, $\frac{\partial p}{\partial t}$, is constant at $\frac{\partial p}{\partial t} = \frac{vn}{vn-1}$.

c. Comment on the effect of increasing the number of firms and increasing the elasticity of demand.

9.20 For the same market description used in exercise 9.19:

a. Can a tax increase raise profit if $n = 1$?

b. What conditions are required for a tax increase to raise profit?

9.21 (Mixed oligopoly) Consider a market with one public firm, denoted 0, and one private firm, denoted 1. Both firms produce a homogeneous good with identical and constant marginal c per unit of output, and face the same linear demand function $P(X) = a - bX$ with $X = x_0 + x_1$. It is assumed $a > c$. The private firm maximizes profit $\pi_1 = P(X)x_1 - cx_1$, and the public firm maximizes a combination of welfare and profit $V_0 = \theta W + [1 - \theta]\pi_0$ with welfare given by consumer surplus less cost, $W = \int_0^X P(y)\,dy - c(x_0 + x_1)$. Both firms choose output as the strategic variable.

a. Calculate the best-response functions of the public and the private firms. Use a graph of the best-response functions to illustrate what would happen if θ changed from 0 to 1.

b. Calculate the equilibrium quantities for the private and public firms. Derive the aggregate output in equilibrium as a function of θ.

c. Calculate the socially optimal output level (by using the marginal cost pricing rule), and compare with the equilibrium outcome.

d. Show that an increase in θ must increase the equilibrium industry output, and so equilibrium price must fall and welfare increase. Verify that the equilibrium outcome converges to the socially optimal outcome when $\theta = 1$.

e. Consider $\theta < 1$ and calculate the quantity subsidy s (with marginal cost after subsidy $c - s$) such that the firms will produce the socially optimal output level. What impact does a change in θ have on the optimal subsidy? Why?

9.22 Define natural monopoly. Draw the demand, marginal revenue, marginal cost, and average cost curves for a natural monopoly.

a. What does the size of a market have to do with whether an industry is a natural monopoly?

b. What are the two problems that arise when the government regulates a natural monopoly by limiting price to be equal to marginal cost?

c. Suppose that a natural monopoly is required to charge average total cost. On your diagram, label the price charged and the deadweight loss to society relative to marginal-cost pricing.

9.23 What gives the government the power to regulate mergers between firms? From the view point of the welfare to society, give a good reason and a bad reason why two firms might want to merge.

9.24 Assume that a monopolist's marginal cost is positive at all output levels. Are the following true or false?

a. When the monopolist operates on the inelastic part of the demand curve, it can increase profit by producing less.

b. When the monopolist operates on the inelastic part of the demand curve, it can increase profit by producing more.

c. The monopolist's marginal revenue can be negative for some levels of output.

9.25 (Varian) A daily dose of the AIDS drug PLC sells for \$18 in the United States and \$9 in Uganda (New York Times, September 21, 2000). Even at \$9 a dose the drug company makes a profit on additional sales. But if the drug were sold at \$9 to everyone, profits would decline. Price discrimination is not popular with consumers, especially those paying the higher price. To evaluate whether differential pricing is good or bad, the critical question from the viewpoint of economics is whether uniform price or differential pricing leads to more people getting the drug. In general, there is no easy answer. Imagine that there are only two countries involved, the United States and Uganda:

a. Imagine the US market for the PLC drug is more than five times the Ugandan market, and the drug sells respectively for \$18 and \$9. What price is likely to prevail if only one price can be charged? What would be the effect on total consumption and, especially, for drug consumers in Uganda? What would be the effect on US drug consumers?

b. Imagine an anti-malarial drug that many people in Uganda would buy at \$2 a dose and few people in the United States would buy at \$10. If the Ugandan market is more than ten times the US market, what price is likely to prevail if drug company can set only one price? What would be the effect on total consumption and for drug consumers in United States and Uganda?

c. Based on this example, discuss when price discrimination is likely to be socially useful and when it does not have much to recommend it.

9.26 A company is considering building a bridge across a river. The bridge would cost $3 million to build and nothing to maintain. The anticipated demand over the lifetime of the bridge is $x = 800 - 100p$, where x is the number of crossings (in thousands) given the price per crossing p.

a. If the company builds the bridge, what will be the profit-maximizing price?

b. Will that price lead to the efficient number of crossings? Why or why not?

c. What will be the company's profit or loss? Should it build the bridge?

d. If the government were to build the bridge, what price should it charge?

e. Should the government build the bridge? Why or why not?

9.27 The jazz singer Nora Jones has monopoly power over a scarce resource: herself on stage. She is the only person who can perform a Nora Jones concert. Does this fact imply that the government should regulate ticket prices for her concerts? Explain.