## Chapter 9 Monopoly

### 9.0 Preliminaries

While pure competition is located at one end of the scale measuring an industry's "degree of competitiveness," pure monopoly is located at the other end: It is the least competitive market structure possible. We reproduce below the definition of monopoly first introduced in Chapter 7.

DEF 7.x: A pure monopoly is a market or industry with a single producer or seller of a good or a service for which there are no close substitutes and into which entry is blocked.

NOTE 9.1: In practice it is sometimes hard to identify an industry as a monopoly because of the difficulty of market definition: We consider a firm a pure monopoly only if there are no close substitutes for its product; but what constitutes a close substitute? The answer to this question is often difficult to determine.

EXAMPLE 9.1: Anheuser-Busch is the only firm which brews Budweiser beer; but (unless you are a Bud fanatic!) there are many substitutes for Budweiser, hence the company is not viewed as a monopoly.

EXAMPLE 9.2: Cimetidine is a drug that revolutionized the treatment of ulcers in the late 1970s. It was patented by the pharmaceutical giant SmithKline Beckman (now known as GlaxoSmithKline or GSK) and sold in the United States under the trade name Tagamet. Since there were no close substitutes for this drug, GSK had a patentprotected monopoly (until substitutes were invented in the mid1980s.)

EXAMPLE 9.3: Electric power is a unique commodity for which, in most uses, there are no substitutes in modern life. It is also produced under conditions which usually make it uneconomical to have more than one producer in a given region. Hence the production and especially the distribution of electric power constitute what we call a "natural" monopoly and in most countries is either a governmentowned or regulated monopoly.

COMMENT 9.1: Just as there are few industries in the Unites States and most of the rest of the world that meet the strict criteria of pure competition, there are also few industries that meet the criteria of "pure" monopoly. But there are many industries that are "near" monopolies (think Microsoft, Google or Intel) so that much of our analysis in this chapter will apply to those industries to some degree. The notion of a "near" monopoly suggests that there may be "degrees" of monopoly, sometimes called market power:

> DEF 9.1 Market power is defined as the ability of a firm to influence (or determine) the price of a good or service. The greater this ability, the higher the degree of market power (or monopoly power) possessed by the firm.

## Why Monopoly?

If a firm has a (profitable) monopoly it must mean that for some reason other firms fail to enter its market. So to ask why monopoly is equivalent to asking what are the barriers to entry that make it impossible (or unprofitable) for other firms to enter a monopolized market. There are several such more or less "insurmountable" entry barriers.

## Government Franchises

In many cases a government grants a firm what is in effect a license to act as a monopoly. Such a "monopoly grant" is called a franchise. Often there may be good reasons for granting a single firm such a monopoly franchise, especially in the case of public utilities: firms that provide electric power, water, transportation or communications services. (This is discussed further below.)

EXAMPLE 9.4 In many localities in the United States a single firm is granted the right to provide cable service. For example, a firm called Cablevision Oakland holds the cable franchise in Hackensack, NJ. It expires in March 2014.

## Patents

A patent, granted by a government (in the United States, the U.S. Patent and Trademark Office), gives an inventor the exclusive right to her/his/their invention, normally for a period of 20 years. Patent laws in effect create (temporary) monopolies. (REM: EXAMPLE 9.2.)

## Economies of Scale

Economies of scale are a major source of entry barriers.
REM 9.1: Economies of scale in the production of a good or service exist if (long-run) average costs decline as the output of the good or service increases. (Chapter 6)

We noted in Chapter 7 that economies of scale create an entry barrier because largescale producers enjoy cost advantages over smaller-scale producers and it is costly and time-consuming to achieve large-scale production. We also noted that the existence of economies of scale in an industry is usually associated with the industry's technology.

If economies of scale prevail in an industry over a range of output which coincides with (or exceeds) its entire market we say that the industry is a natural monopoly. Frequently a single firm receives a government franchise giving it the right to be the sole supplier in such a market. (See above.) Natural monopoly is discussed in Section 9.3

## "Behavioral" Entry Barriers

Behavioral entry barriers are created when a dominant firm adopts successful policies aimed at maintaining or increasing its monopoly or near-monopoly status.

## High Sunk Costs

The existence of high sunk costs in an industry often creates insurmountable entry barriers which lead to the formation of a natural monopoly. (See Section 9.3.)

## Control over a Key Input

Occasionally, a monopoly arises because for some reason a single firm owns or controls or has exclusive access to a key input required for the production of a good.

EXAMPLE 9.5: Before World War II, the Aluminum Company of America (Alcoa) was the sole producer of (primary) aluminum in North America, in part because of their control over the supply of bauxite, a key input in the manufacture of aluminum. (After the end of World War II the production of aluminum became a world-wide oligopoly, partly due to actions by the U. S government.)

NOTE 9.2 "Primary" aluminum is made by a complicated industrial process from the ore bauxite. "Secondary" aluminum is in effect recycled from aluminum scrap.

EXAMPLE 9.6: DeBeers, the global diamond "monopoly," started out by being, at one point, the sole owner of all diamond operations in South Africa, a center of diamond mining. In the $20^{\text {th }}$ century the company maintained its dominant position by convincing producers in other countries to join its cartel and by buying up diamonds in world markets in order to control the supply (as well as by other "monopolizing" means).

NOTE 9.3 We place quotation marks around the word "monopoly" in EXAMPLE 9.6 because in some ways De Beers acts more like a dominant firm in a cartel than a pure monopoly.

## Network Externalities

REM 9.2: Network externalities exist in an industry (or "network") if the utility (or disutility) individual $A$ derives from being a member of the network depends on the number of other individuals ( $B, C, D, \ldots$ ) who are also participants in this network. If individual $A$ derives utility (benefits) if more individuals are member of the network, we call them positive network externalities, etc. (Chapter 7)

EXAMPLE 9.7: Think of the early days of telephone systems. Imagine that there are two companies (the Edison Co. and Bell Corp.) and because of better service, better marketing or by sheer luck, Bell signs up more subscribers than Edison. Being part (i.e., a customer) of a telephone network is useful to the extent that other individuals, businesses, etc., are linked to it. So as a consequence of the larger (and rising) number of Bell subscribers, even more people sign up with Bell, and so on and on. A point may come when no one signs up with Edison and Bell is the only surviving telephone network (i.e., a monopoly).

NOTE 9.4 Of course, network externalities may interact with economies of scale to make this process even more powerful.

### 9.1 A Monopoly's Price and Output Decisions

FIG 9.1


Consider FIG 9.1. It shows first, the set of per unit, or average, short-run cost curves of the Belton-Dixon Co., the major New Jersey-based drug manufacturer which we first encountered in Chapter 4, EXAMPLE 4.4. The company holds the patent on the wonder drug Prosaic and is thus the monopolistic producer in this market for the life of the patent (or until a substitute is discovered!) Note that the cost curves have the same "look" or shape as the unit cost curves of a purely competitive firm: except for differences in scale (which are not unimportant!), monopolies face the same production problems as any other firm. They have to buy inputs in input markets (although they may have market power on the buyers' side in these markets) and they confront the same physical and technical laws as any other producer (EXAMPLE: the law of diminishing returns!). So in general there is no fundamental difference between the monopoly's (short-run) cost curves and those of a purely competitive firm. The crucial difference arises on the demand side: the monopoly is the market, so it confronts the market demand curve. Hence in FIG 9.1 the monopoly's demand curve is depicted as a standard downsloping line embodying the law of demand.

NOTE 9.5: After careful consideration Belton-Dixon's management rejected the notion that the demand for Prosaic is completely inelastic and assumed instead that they faced a downsloping demand curve as shown in FIG 9.1 above.

We are almost ready to try to solve the monopoly's decision problem of finding its profit-maximizing price and output level but not quite.

NOTE 9.6: We are taking it for granted that ASSUMPTION 8.1 ("Firms seek to maximize economic profit") is true for a monopoly just like for any other firm.

NOTE 9.7: In Chapter 7 we point out that unlike a pure competitor, a monopoly is a price setter, that is, it determines the price at which it sells its output. That is, unlike a pure competitor a monopoly must have a price policy. But in the simple monopoly model we concentrate on in this chapter it turns out that the price and output decisions are simultaneous - that is, when the monopoly decides what output to produce they also decide what price to charge (and vice versa).

To determine the monopoly's profit-maximizing price and output we must apply DECISION RULE 8.3, i.e., find that output level at which MR = MC. But this introduces a complication which we discuss in the next section.

## Marginal Revenue When Firms Have Market Power

We know that for a firm in a purely competitive industry marginal revenue is simply the price. But when firms have market power (with monopoly representing the extreme case of market power) things are not so simple. It turns out that for such firms marginal revenue is less than the price.

According to DEF 9.1 a firm with market power is able to influence (or determine) the market price. Related to this is the fact that such a firm faces a downsloping demand curve or a demand schedule in which price and quantity demanded are inversely related. Now consider Table 9.1 below. Columns (1) and (2) represent an extremely simple demand schedule. We want to calculate the marginal revenue that a firm receives when it produces and sells an additional unit of its output. To do so we take an intermediate step and calculate total revenue (TR) at each output level. Then it is easy to calculate MR: it is simply the change in total revenue as output changes (increases) by one unit. For example, the marginal revenue of the third unit $\left(\mathrm{MR}_{3}\right)$ equals $\$ 3 .\left[\mathrm{MR}_{3}=\right.$ $\$ 15-\$ 12=\$ 3]$ The results of these calculations are shown in Column (4) of Table 9.1.

REM 9.3: What does the phrase "the marginal revenue of the third unit" mean? It means the additional revenue that a firm receives when it produces and sells three units instead of just two.

Table 9.1

| $(1)$ | $(2)$ | $(3)$ | $(4)$ |
| :---: | :---: | :---: | ---: |
| $\mathbf{P}$ | $\mathbf{Q}_{\mathbf{d}}$ | $\mathbf{T R}$ | $\mathbf{M R}$ |
| $\$ 7$ | 1 | $\$ 7$ | $\$ 7$ |
| 6 | 2 | 12 | 5 |
| 5 | 3 | 15 | 3 |
| 4 | 4 | 16 | 1 |
| 3 | 5 | 15 | -1 |
| 2 | 6 | 12 | -3 |
| 1 | 7 | 7 | -5 |

A careful look at Column (1) and Column (4) of Table 9.1 confirms our initial statement: except for the first unit, marginal revenue is less than the price at every output level. (In addition, MR falls more rapidly than the price as the output level rises.) For example, if the firm wants to sell 3 units it can charge a maximum of $\$ 5$ but the marginal revenue of the third unit is $\$ 3$.

QUESTION 9.1: That marginal revenue is less than price seems to be "merely" a matter of arithmetic. Is there anything beyond the arithmetic in Table 9.1 that explains this fact?

ANSWER 9.1: The answer is yes. Consider an output level of 2 units. Remembering that in pure competition price equals marginal revenue it is tempting to think that the marginal revenue of the second unit is $\$ 6$ (i.e., the price at which two units can be sold). But that is only part of the story: when the firm sells 2 units at $\$ 6$ each it "gives up the opportunity" so to speak to sell one unit for $\$ 7$. It is as though the price of the first unit had to be lowered by $\$ 1$. So $M R_{2}=\$ 6-\$ 1=\$ 5$, which is the number you find in Column (4) of Table 9.1.

QUESTION 9.2: $\quad$ Why is the $M R$ of the third unit in Table 9.1 equal to $\$ 3$ when the price at which 3 units can be sold is $\$ 5$ ?

ANSWER 9.2: The answer to this question is left as an exercise for the reader.

## Graphic Depiction of Marginal Revenue for Firms with "Market Power"

FIG 9.2


Consider FIG 9.2. It shows a linear demand curve associated with a single firm with market power. Below this curve lies a dashed line labeled "MR": it is the marginal revenue "curve" associated with the demand curve. In fact, we can think of it as the "shadow" of the demand curve: when a single seller faces a downsloping demand curve there always goes with it its "shadow," the marginal revenue curve. It embodies two facts we developed earlier: (1) For a firm with market power marginal revenue is less than price ( $M R<P$ ) and (2) Marginal revenue falls faster than the price as output level rises. (The MR curve lies below the demand curve and has a steeper absolute slope.)

QUESTION 9.3: How should you "read" FIG 9.2?
ANSWER 9.3: $\quad$ Consider point $f$ lying on the demand curve. It says that when the price is $\$ 17$ the firm is able to sell 10 units. But the marginal revenue of the $10^{\text {th }}$ unit (read off the MR curve) is $\$ 8$ ! Similar statements can be made about any point lying on the demand curve. Hence the way the MR curve is constructed (i.e., it lies below the demand curve) indicates the fact that (except for the first unit) $M R$ is less than $P$.

QUESTION 9.4: In FIG 9.2, when the price is $\$ 13$, what is the corresponding marginal revenue?

ANSWER 9.4: The answer to this question is left as an exercise for the reader.

QUESTION 9.5: How does one "construct" a marginal revenue curve?
ANSWER 9.5: As before, we discuss this question in relation to the simple case of a linear demand curve.

Draw any horizontal line from the price (vertical) axis to the demand curve parallel to the quantity (horizontal) axis. For example, in FIG 9.2 draw the line from " 13 " to point $e$. Find its midpoint. Draw any other line in the same way. For example, the line from " 17 " to point $f$. Find its mid-point. Then the straight line drawn through both mid-points represents the marginal revenue curve associated with the linear demand curve. (Note that because of this construction the demand curve and the marginal revenue curve have the same vertical intercept.)

QUESTION 9.6: Why is the MR curve constructed as described in ANSWER 9.5?
ANSWER 9.6: The explanation requires the application of a bit of elementary geometry but we will not do so here.

NOTE 9.8: An interesting fact emerges from FIG 9.2. Observe that point $e$ on the demand curve lies right above the point where the MR curve crosses the horizontal axis, i.e., where $M R=0$. To the left of that point then MR is positive and so a price decrease leads to higher total revenue. In Chapter 4 we call the demand in such an interval elastic. To the right of point $e \mathrm{MR}$ is negative and a lower price leads to lower total revenue. We call the demand in such an interval inelastic. Point $e$ is the boundary point between these two intervals, so right at that point demand must be unit elastic. We have found a way then to determine in which segment of a linear demand curve the demand is elastic, in which segment it is inelastic and also to find the boundary point where demand is unit elastic.

## A Monopoly's Profit-Maximizing Price and Output

We are finally in a position to answer our initial question: What is the monopoly's (in this case Belton-Dixon's) profit-maximizing price and output level? To answer this question we return to FIG 9.1 (page 5). To the downsloping demand curve depicted in the graph we add its "shadow," the marginal revenue curve (shown by the broken line
labeled "MR."). The rest is easy: the answer is found at the intersection of the MR and the MC curves. This happens at point $M$ in FIG 9.1 (see arrows pointing to $P_{m}$ and $Q_{m}$. The profit-maximizing quantity $\left(\mathrm{Q}_{\mathrm{m}}\right)$ is 30 and the profit-maximizing price $\left(\mathrm{P}_{\mathrm{m}}\right)$ is (approximately) \$128.


QUESTION 9.7: (REM) Jim Madison is the product manager for Belton-Dixon's patented drug, Prosaic. It took some effort but he was finally persuaded that the demand for Prosaic is not completely inelastic. Leroy Huffington, a college intern working in Madison's office glanced at FIG 9.1 and pronounced that the company's profitmaximizing price is approximately $\$ 60$. He expressed it this way: "Move up from $\mathrm{Q}_{\mathrm{m}}$ on the horizontal axis to point M and turn left (to the vertical axis)." Is this correct?

ANSWER 9.7: No! The highest price a firm can charge if it wishes to sell a particular quantity is read off its demand curve. Moving from $Q_{m}$ on the horizontal axis to the demand curve (at point $R$ ) indicates that the profit-maximizing price is approximately $\$ 128$. (See arrow at $\mathrm{P}_{\mathrm{m}}$.)

QUESTION 9.8: At its profit-maximizing output level what is the monopoly's total revenue (TR), total cost (TC) and total profit ( $\Pi$ )?

ANSWER 9.8: $\quad$ Since total revenue equals price times quantity, the monopoly's TR is given by the area of the rectangle $\mathrm{OP}_{\mathrm{m}} \mathrm{RQ}_{\mathrm{m}}$. That is: $\mathrm{TR} \approx \$ 128 \mathrm{x}$ $30 \approx \$ 3,840$.

Since total cost equals ATC times quantity and ATC $\approx \$ 95$, the monopoly's TC is given by the area of the rectangle OSTQ ${ }_{m}$. That is: $\mathrm{TC} \approx \$ 95 \times 30 \approx \$ 2,850$.

Since total profit equals TR minus TC, total profit is given by the area of the rectangle $S P_{m} R T$. That is: $\Pi \approx \$ 3,840-\$ 2,850 \approx \$ 990$

QUESTION 9.9: Is the profit earned by the monopoly a "normal" or an "abovenormal" profit?

ANSWER 9.9: Since the monopoly's price exceeds average total cost the monopoly is earning an above-normal profit.

REM 9.4 A normal profit is viewed in economics as a cost; i.e., it is the opportunity cost of the resources provided by the owners of a firm.

QUESTION 9.10: Is a monopoly guaranteed to earn an above-normal profit?
ANSWER 9.10: $\quad$ No. For all sorts of reasons having to do with the determinants of demand and costs the demand curve in FIG 9.1 could be tangent to the firm's ATC curve; then the monopoly would just earn a normal profit despite its monopoly status. The demand curve could even lie below the firm's ATC curve; then the same analysis would apply as in our discussion of the "shut-down point" in Chapter 8: If the price is less than ATC but greater than AVC the firm would operate in the short run but shut down in the long run, etc.

QUESTION 9.11: $\quad$ The firm depicted in FIG 9.1.) is earning an above normal profit. In Chapter 8 we noted that in a purely competitive industry the existence of above-normal profits attracts entry of new firms and the above-normal profits are "competed away." Will something similar happen in the case of a monopoly?

ANSWER 9.11: No. By definition, entry into a (purely) monopolistic market is blocked so the monopoly is able to hold on to its above-normal profits. So FIG 9.1 depicts both the short-run and long-run circumstances of the monopoly. We call the monopoly's abovenormal profits monopoly profits.

DEF 9.2: | Monopoly profits (sometimes called monopoly |
| :--- |
| rents) are defined as the enduring above-normal |
| profits earned by a firm because of its |
| monopoly position. |

QUESTION 9.12: Is there a "monopoly supply curve?"
ANSWER 9.12: Since we demonstrated in Chapter 8 that a pure competitor's marginal cost curve (at least that portion of it that lies above average variable cost) constitutes its supply curve, it is tempting to think that this is also true in the case of monopoly: its marginal cost curve is equivalent to the monopoly's supply curve. But this not correct. First, intuitively it should be clear that one cannot talk about a monopoly supply curve, since a supply curve in effect answers the following type of question: if the (market) price of a gadget is $\$ x$ what is the corresponding quantity a firm is willing and
able to supply? But for a monopoly such a question is meaningless, since the monopoly determines the price!

More formally, remembering that the monopoly's profitmaximizing output occurs where $M R=M C$, that is, where the $M C$ curve intersects the MR curve, it should be easy to see from FIG 9.1 that shifts in demand could occur in such a way that the intersection point results in the same output level but different (higher or lower) prices. If each level of production is not associated with a single price, we cannot speak of a "supply curve."

## The Lerner Index

Observe that in FIG 9.1 when the firm produces and sells $Q_{m}$ and charges the price $P_{m}$ the price exceeds marginal cost. This is typical of firms with market power and is an important fact we discuss here and in Section 9.2 below.

REM 9.5: $\quad$ For a pure competitor the profit maximizing output occurs where MR (= P) = MC. An output level where $\mathrm{P}=\mathrm{MC}$ implies an optimal allocation of resources.

The difference between price and marginal cost provides an important measure of the degree of monopoly (or market power) possessed by a firm. But it is not enough to look at the absolute difference between P and MC since this difference might be several thousand dollars in one industry but a few cents in another. So a relative measure is needed. This is provided by the Lerner Index (L) defined below.

DEF 9.3: $\quad$ The Lerner Index ( L ) is defined as price minus marginal cost divided by price. In symbols:

$$
L=\frac{P-M C}{P}
$$

PROBLEM 9.1: $\quad$ What is the Lerner index of the firm depicted in FIG 9.1?
SOLUTION 9.1 The difference between the monopoly's price and marginal cost is given in FIG 9.1 by the length of the line segment from $M$ to $R$. We found that $P=\$ 128$ and $M C=\$ 60$ so the firm's Lerner index is:

$$
L=\frac{128-60}{128}=0.53
$$

NOTE 9.9:
There is no "boundary value" of the index where one can say that above that value the industry is monopolistic and below it, it is not.

All one can say is that the higher a firm's Lerner index the greater its market power (or degree of monopoly).

QUESTION 9.13 What is the "range" of the Lerner Index? That is, what are its lowest and highest values?

ANSWER 9.13: The answer to this question is left as an exercise for the reader.

## A Monopoly's Profit-Maximizing Price and Output: A Tabular Approach

A monopoly's profit-maximizing price and quantity can also be found using the information contained in Table 9.2 below, which summarizes the demand and cost conditions facing the Dragon Co., a monopoly producing gadgets. Columns (1) and (2) constitute the firm's demand schedule.

QUESTION 9.14: How can the $M R=M C$ rule be applied in Table 9.2?
ANSWER 9.14: Carefully inspect Column (3) and Column (10) which show marginal revenue and marginal cost. Find the output level where MR = MC (or where they come closest while MR still exceeds MC). This happens when $\mathrm{Q}=10$. Then $\mathrm{MR}=\$ 90, \mathrm{MC}=\$ 87$ and $\mathrm{P}_{\mathrm{m}}=\$ 190$.

QUESTION 9.15: How can we be sure that we have found the right answer?
ANSWER 9.15: Ask yourself the following question: Perhaps the firm should produce 11 units instead of 10 ? But if they do then producing the additional unit would add $\$ 91$ to their costs but only $\$ 70$ to their revenues, so producing 11 units is not a good idea.

QUESTION 9.16: Should the firm produce 9 units instead of 10?
ANSWER 9.16: The answer to this question is left as an exercise for the reader.

Answering the question this way provides a good test of your understanding of the $M R=M C$ rule. But the question can also be answered in a more straightforward way: Calculate TR (= P x Q) shown in Column (4), then calculate TC (=ATC $x$ Q) shown in Column (6). Subtract Column (6) from Column (4) and obtain the profit figures shown in Column (11). The highest profit level ( $\Pi=\$ 610$ ) occurs when $\mathrm{Q}=10$ so this again confirms the $\mathrm{MR}=\mathrm{MC}$ rule.

Table 9.2

| $(1)$ | $(2)$ | $(3)$ | $(4)$ | $(5)$ | $(6)$ | $(7)$ | $(8)$ | $(9)$ | $(10)$ | $(11)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | ---: |
| $\mathbf{Q}$ | $\mathbf{P}$ | MR | TR | TFC | TC | TVC | AVC | ATC | MC | Profit |
| 0 | 280 | --- | - | 500 | 500 | 0 | --- | --- | --- | -500 |
| 1 | 270 | 270 | 270 | 500 | 551 | 51 | 551.00 | 551.00 | 51 | -281 |
| 2 | 260 | 250 | 520 | 500 | 606 | 106 | 303.00 | 303.00 | 55 | -86 |
| 3 | 250 | 230 | 750 | 500 | 665 | 165 | 221.67 | 221.67 | 59 | 85 |
| 4 | 240 | 210 | 960 | 500 | 728 | 228 | 182.00 | 182.00 | 63 | 232 |
| 5 | 230 | 190 | 1,150 | 500 | 795 | 295 | 159.00 | 159.00 | 67 | 355 |
| 6 | 220 | 170 | 1,320 | 500 | 866 | 366 | 144.33 | 144.33 | 71 | 454 |
| 7 | 210 | 150 | 1,470 | 500 | 941 | 441 | 134.43 | 134.43 | 75 | 529 |
| 8 | 200 | 130 | 1,600 | 500 | 1,020 | 520 | 127.50 | 127.50 | 79 | 580 |
| 9 | 190 | 110 | 1,710 | 500 | 1,103 | 603 | 122.56 | 122.56 | 83 | 607 |
| 10 | 180 | 90 | 1,800 | 500 | 1,190 | 690 | 119.00 | 119.00 | 87 | 610 |
| 11 | 170 | 70 | 1,870 | 500 | 1,281 | 781 | 116.45 | 116.45 | 91 | 589 |
| 12 | 160 | 50 | 1,920 | 500 | 1,376 | 876 | 114.67 | 114.67 | 95 | 544 |
| 13 | 150 | 30 | 1,950 | 500 | 1,475 | 975 | 113.46 | 113.46 | 99 | 475 |
| 14 | 140 | 10 | 1,960 | 500 | 1,578 | 1,078 | 112.71 | 112.71 | 103 | 382 |
| 15 | 130 | -10 | 1,950 | 500 | 1,685 | 1,185 | 112.33 | 112.33 | 107 | 265 |
| 16 | 120 | -30 | 1,920 | 500 | 1,796 | 1,296 | 112.25 | 112.25 | 111 | 124 |
| 17 | 110 | -50 | 1,870 | 500 | 1,911 | 1,411 | 112.41 | 112.41 | 115 | -41 |
| 18 | 100 | -70 | 1,800 | 500 | 2,030 | 1,530 | 112.78 | 112.78 | 119 | -230 |

### 9.2 Monopoly and Allocative Efficiency

We found in Chapter 8 that an industry exhibits allocative efficiency when $\mathrm{P}=\mathrm{MC}$ (the so-called marginal cost pricing rule, RULE 8.1). This condition is achieved "automatically" in purely competitive industries. But as we saw in Section 9.1, especially in our discussion of the Lerner Index, in the case of monopoly price exceeds marginal cost. So we can immediately conclude that monopoly exhibits allocative inefficiency. To examine this idea further, we would like to make as direct a comparison as possible between monopoly and pure competition. To do this we use a highly simplified version of Figure 9.1.

Figure 9.3 depicts a purely competitive industry in long-run equilibrium. It produces widgets. We assume that the average cost (AC) of production is constant and that $\mathrm{AC}=$ $\$ 80$. Since AC is constant it is also equal to MC: the cost of producing an additional unit remains the same at all output levels, i.e., $\mathrm{AC}=\mathrm{MC}=\$ 80$. Hence the horizontal line in Figure 9.3 is labeled "AC, MC." The figure also shows the industry demand curve, labeled $D$.

NOTE 9.10: We write "AC" since we are looking at the industry from a long-run perspective; hence we do not need to differentiate between variable and total costs.

FIG 9.3


In a purely competitive industry the forces of competition drive the market price down to the average cost of production, (REM: DEF 8.11 which defines a purely competitive industry's long-run equilibrium condition) so the long-run equilibrium price is $\$ 80$ and we write $\mathrm{P}_{\mathrm{c}}=\$ 80$. We can also read off the demand curve that the quantity demanded at that price is 200, i.e, in long-run equilibrium $\mathrm{Q}_{\mathrm{d}}=200$. Now imagine that a group of investors buy up the several hundred or several thousand firms making up this industry and create a single monopolistic firm.

NOTE 9.11
Many observers believe that the antitrust laws are weakly enforced in the Unites States and that competition laws in the European Union are also not enforced very energetically. Nevertheless, the event we are describing here is highly unlikely. We are using it for illustrative purposes only!

Assume further that as a result of the monopolization of the industry there is no change in production costs and no change in demand. Hence all we need to analyze the new situation is to add to the industry's demand curve its "shadow," the marginal revenue
curve. This is shown in Figure 9.3 as the broken line labeled "MR." We are now able to read off the diagram the newly-created monopoly's profit-maximizing price and output level: this occurs at the intersection of the horizontal MC curve and the newly-drawn $M R$ curve. We conclude that $P_{m}=\$ 170$ and $Q_{m}=100$.

QUESTION 9.17 Say you want to explain to an individual untrained in economics what is wrong with monopoly. How would you go about it?

ANSWER 9.17: It is easy to read off Figure 9.3 that a monopoly produces a smaller output and charges a higher price than a comparable purely competitive industry.

NOTE 9.12: $\quad$ According to Figure 9.3 the monopoly produces and sells exactly half the output the purely competitive industry produced, but this is entirely a result of the many simplifying assumptions that we made: a linear demand curve, constant AC, etc. and cannot be generalized.

ANSWER 9.17 tells us that a monopoly produces a smaller output (and charges a higher price!) than a purely competitive industry. We therefore say that monopolies underallocate resources to the production of their output. In other words, they limit production so that they are able to charge higher prices and earn higher profits than a comparable purely competitive industry. But we would like to find a more precise explanation (and measure!) of the inefficiency of monopoly compared to pure competition than we get in ANSWER 9.17 above. To do this we need some additional economic concepts.

## Consumers' Surplus

In Chapter 8's discussion of allocative efficiency we pointed out that a demand curve can also be viewed as a "marginal benefit" curve. This fact is the starting point for our discussion of the concept of consumers' surplus.

Consider again the demand curve depicted in Figure 9.3. We note that according to the graph there is at least one person (individual A) who is "willing to pay" a bit less than $\$ 260$ (say $\$ 259.50$ ) for one unit of this good (one widget). We repeat a question we asked once before (in Chapter 8): why is individual A willing to pay this amount for one widget? And our answer is also the same as before: individual A expects to obtain a benefit from this unit at least equal to the amount he is willing to pay. Then there is a second individual (individual B) who is willing to pay a bit less than the first (say \$259) for one widget. We repeat the question and answer: The second individual is willing to pay this amount because it is a measure of the benefit she expects to obtain from one widget. And so it goes. We can "add up" the expected benefits of individuals A and B
and say that together they obtain a total benefit (TB) equal to $\$ 259.50+\$ 259=\$ 518.50$. We can think of each individual's benefit as represented graphically in Figure 9.3 as a thin "strip" from the demand curve to the horizontal axis. As we move to the right along the horizontal axis there are more and more individuals willing and able to buy widgets as the price falls and more and more "strips" to add up. If the industry is (a) purely competitive and (b) in long-run equilibrium, 200 widgets will be produced and sold and there will be 200 of these thin strips to add up. If we assume for simplicity that each individual buys just one widget, we can ask, what is the total benefit the 200 individuals obtain from using, consuming or owning widgets? The answer is given by the area under the demand curve from the vertical axis to the broken vertical straight line from " 200 " on the horizontal axis to point $c$, i.e., the area from the origin to " 260 " on the vertical axis to point $c$ to on the demand curve to " 200 " on the horizontal axis.

QUESTION 9.18: For the 200 individuals who buy widgets when Figure 9.3 depicts a purely competitive industry in long-run equilibrium $T B=\$ 34,000$. Can you explain why?

ANSWER 9.18: The answer requires a bit of elementary geometry and is left as an exercise for the reader.

So far we have discussed buyers' willingness to pay. But what do they actually pay? In a purely competitive industry they pay the equilibrium price, which in Figure 9.3 is $\$ 80$. So to obtain $\$ 34,000$ worth of "benefits" they must pay something. What they pay as a group is called total expenditure or total outlay (TO). In Figure 9.3 TO = \$80 x $200=$ $\$ 16,000$ and is depicted graphically as the area of the rectangle from the origin to point " 80 " to point $c$ to "200." It appears then that the buyers of a good or service as a group, obtain a kind of "net benefit" which consists of the difference between the total benefits (TB) they receive and the total outlays (TO) they must make. This net benefit is called consumers surplus (CS) and is shown geometrically in Figure 9.3 by the area of the triangle from point " 80 " to point " 260 " to point $c$.

EXAMPLE 9.8: You are planning to buy a certain brand of DVD player. After giving the matter much thought you decide that the maximum you are willing to pay for it is $\$ 200$. When you arrive at the Excellent Buys Electronics store you find that you can get the player for $\$ 130$ and you make the purchase. You therefore obtain a "consumer's surplus" of $\$ 200-\$ 130=\$ 70$.

DEF 9.4: Consumers' surplus (CS) is defined as the difference between the total benefits (TB) obtained by a group of buyers in a market from the purchase, consumption or use of a good (as measured by their willingness to pay) and their total expenditures or total outlays (TO) on that good. In symbols:

$$
\mathrm{CS}=\mathrm{TB}-\mathrm{TO}
$$

QUESTION 9.19: For the 200 individuals who buy widgets when Figure 9.3 depicts a purely competitive industry in long-run equilibrium CS $=\$ 18,000$. Can you explain why?

ANSWER 9.19: The answer requires a bit of elementary geometry and is left as an exercise for the reader.

## The "Burden" of Monopoly

We are now able to develop a precise picture of the (allocative) inefficiency that results from monopoly. Specifically, we would like to be able to answer the question, what is the cost to society of the monopolization of an industry? Consider Figure 9.3 again. We found that when the widget industry is purely competitive, consumers' surplus is given by the area of the triangle from point " 80 " to point " 260 " to point $c$ and numerically we know that CS = \$18,000.

QUESTION 9.20: What is the consumers' surplus once the industry has become monopolized?

ANSWER 9.20: When the industry becomes monopolized $Q=100$ and $P=\$ 170$. The consumers' surplus is then given by the area of the triangle from point " 170 " to point " 260 " to point $a$, or $\mathrm{CS}_{\mathrm{m}}=\$ 4,500$.

QUESTION 9.21 What is the loss of consumers' surplus that results from the monopolization of the industry?

ANSWER 9.21: The answer is given by the area from point " 80 " to point " 170 " to point $a$ to point $c$ which is equal to $\$ 18,000-\$ 4,500=\$ 13,500$ or the difference between consumers' surplus when the industry is purely competitive $(\$ 18,000)$ and when it is a pure monopoly $(\$ 4,500)$.

QUESTION 9.22: Does the loss of consumers' surplus (which we calculated to be $\$ 13,500$ for the gadget industry) represent the cost to society of the monopolization of an industry?

ANSWER 9.22: Not quite. Notice that the loss of consumers' surplus shown in Figure 9.3 can be divided into two parts: one consists of monopoly profit shown by the area of the rectangle from point " 80 " to point " 170 " to point $a$ to point $b$, or $\$ 9,000$.

NOTE 9.13: When the widget industry is monopolized its total revenue (TR) is shown by the area of the rectangle from the origin to point " 170 " to point $a$ to point " 100. ." $\left(\mathrm{TR}_{\mathrm{m}}=\$ 170 \times 100=\$ 17,000\right.$.) Total Cost (TC) is shown by the area of the rectangle from the origin to point " 80 " to point $b$ to point " 100 ", since $\mathrm{TC}=\mathrm{AC} \times \mathrm{Q}, \mathrm{TC} \mathrm{m}=\$ 80 \times 100=$ $\$ 8,000$.) Since profit $=T R-T C$, the monopoly's profit is shown by the area of the rectangle from point " 80 " to point " 170 " to point $a$ to point b. $(\Pi=\$ 17,000-\$ 8,000=\$ 9,000$.)

Monopoly profit can be thought of as a transfer from consumers (or users or buyers) of the good or service to the monopoly's owners. It does not represent a net loss to society. The net loss instead is equal to the loss of consumers' surplus minus monopoly profit. In Figure 9.3 it is shown by the area of the triangle bac (shown in red )and is equal to $\$ 4,500$.

NOTE 9.14: The economist acting as a scientific observer of market behavior is neutral if individual A gains $\$ 1$ and individual B loses $\$ 1$. In her view a net loss to society occurs only if, as result of some transaction, individual A loses \$2 but individual B gains only \$1 (or vice versa). This is exactly what happens in our example: as a result of the monopolization of the widget industry consumers lose $\$ 13,500$ but the monopoly gains only $\$ 9,000$. There is therefore a net loss to society of $\$ 13,500-9,000=\$ 4,500$ shown in Figure 9.3 by the triangle bac This net loss is sometimes called a deadweight loss or the monopoly burden, (or more generally, an excess burden).

DEF 9.5: Deadweight Loss (DWL) is defined as the net loss to society resulting from any change in market structure, adoption of a public policy, including especially tax policy, etc.

QUESTION 9.23: Have we reached a definite conclusion that monopoly suffers from economic or allocative inefficiency and is therefore undesirable from society's point of view?

ANSWER 9.23: No. To reach the conclusion about the allocative inefficiency of pure monopoly we made several important simplifying assumptions. If we considered some additional factors (such as cost savings due to economies of scale) we might reach a different conclusion. Looked at another way, we may conclude that the inefficiencies (deadweight loss) that result from the monopolization of an industry might be offset by other, compensating advantages. But both economists and noneconomists over a period of more than a century (if not longer) seem to have concluded that in general monopoly is undesirable from society's point of view and governments in many parts of the world have adopted antimonopoly legislation (called antitrust laws in the United States.) This topic will be discussed further in Chapter 13.

## 9.3 "Natural" Monopoly: An Application of Monopoly

## Analysis

We noted earlier (in Section 9.0) that economies of scale are one source of entry barriers and that if economies of scale (i.e., declining long-run average costs) prevail in an industry over a range of output which coincides with (or exceeds) its entire market we say that the industry is a natural monopoly. In everyday language as well as in the language of law and politics such industries are often referred to as public utilities and comprise businesses such as electricity generation and distribution, water supply, transportation and communications services.

DEF 9.6: A natural monopoly is an industry with continuously declining long-run average costs (over some relevant output range). As a result of this situation it may be inefficient from society's point of view to have more than one producer supplying the market.

## Why Natural Monopoly?

Natural monopoly arises from the interplay between declining long-run average costs and the size of the market.

EXAMPLE 9.9: Consider the gadget industry. Assume its technology and other characteristics are such that the lowest point on its LRAC curve (i.e., its point of minimum efficient scale, or MES) occurs when $Q=$ 20,000 . Also assume that given the demand facing the industry its total output would be approximately 20,000 units. Then the gadget industry can sustain only a single efficient producer.

This point can be illustrated using Figure 9.4 below. Note that within the output range shown in the graph (i.e., up to $Q=20,000$ at least), LRAC is continuously declining and LRMC is below LRAC.

FIG 9.4

(The latter point turns out to be important and will play a large role in our discussion of regulatory policy.) It is also true that in the graph LRMC is constant (i.e., it is shown by a horizontal straight line) but this is a result of the simplified LRAC curve we are using for illustration.

REM 9.6: Our discussion in Chapter 5 showed that if an average quantity is falling, the corresponding marginal quantity is below the average quantity while if an average quantity is rising, the corresponding marginal quantity is above the average quantity. This fact explains why in the case of natural monopoly we have LRMC < LRAC throughout.

Imagine that there is a single firm producing 20,000 gadgets. Its (long-run) average cost of production is $\$ 5$. Now assume that we are unhappy that the industry is monopolized and we break it up into two equal parts, each producing 10,000 gadgets. Then the LRAC is approximately \$6. (Note the first arrow along the vertical axis near the origin.) If we are still unhappy and try to break up the monopoly into four equal parts, each producing 5,000 gadgets, LRAC is now approximately $\$ 8$. If we go further and break the monopoly up into 10 equal parts, each producing 2,000 gadgets, (see the arrow along the horizontal axis close to the origin), the resulting LRAC is approximately $\$ 14$. Clearly, the increases in LRAC resulting from changes in the structure of the industry represent a waste of scarce economic resources and are inherently undesirable. Hence we conclude that the gadget industry can sustain only a single efficient producer.

NOTE 9.15: The discussion in the previous paragraph represents an oversimplification of the actual situation, since there is no reason to expect for example that each of the 10 firms would produce exactly $1 / 10$ th of the output of a single firm, etc.

Figure 9.4 can also be used to explore the possible historical development of a natural monopoly. Assume that there are several firms of equal size in the gadget industry. Assume further that by chance the market share of one of the firms in the industry, Company A, increases slightly. Because of the existence of economies of scale it is now able to lower its prices and gain further market share at the expense of its rivals. As a result its cost advantage increases still more and one by one its rivals have to exit the industry until Company A remains as the sole survivor.

NOTE 9.16: Whether the description above depicts the actual development of natural monopoly in the United States and elsewhere is controversial among economic historians but we will not discuss the issue further here.

QUESTION 9.24: Why does an industry have continuously declining long-run average costs? In other words, why does its LRAC curve embody economies of scale in such a way as to lead to a situation of "natural monopoly?"

ANSWER 9.24: The most common (but not the only) answer is that it results from high levels of sunk costs in an industry's production process.

REM 9.7:
Sunk costs are costs which have been incurred in the past or for some other reason are unavoidable. (See Chapter 5.)

EXAMPLE 9.10: Consider the electric power industry. To generate and distribute electricity in a given region requires a large initial investment in electric generating plants and a distribution network, including transformers, power lines, etc. Once these investments have been made they represent sunk costs. But the cost of "hooking up" and providing electricity to additional customers is relatively low. Hence the more customers an electric power company acquires (i.e., the more power it produces and sells) the lower its average costs become. To put it more simply, it would constitute an obvious waste of scarce resources (plus a nuisance!) to have more than one set of poles and wires to distribute electricity in a given town or region. Therefore the electric power industry (especially power distribution) represents a standard example of natural monopoly.

## The Regulation of Natural Monopoly

Society then faces a dilemma: On the one hand monopolies are undesirable because they restrict output and charge higher prices than would a comparable purely competitive industry, if such an industry were feasible. Technically, they "misallocate" resources to the production of the monopolized good or service. (Some people would add as an additional negative aspect that they earn monopoly profits.) On the other hand "breaking up" a natural monopoly leads to inefficiency because of the resulting sacrifice of economies of scale and the obvious cost of duplication of facilities in the case of public utilities, etc. This dilemma was resolved in late $19^{\text {th }}$ and early $20^{\text {th }}$ centuries in two major ways: through government ownership (or nationalization) or through regulation. In Europe and Asia the solution of choice was government ownership while in the United States the approach most frequently used was the regulatory approach. In effect natural monopolies were told: you may remain a monopoly (we will grant you a monopoly franchise over some good or service in a particular state or locality or even the country as a whole) but in return for this we will regulate you. Specifically, we will try to prevent you from charging monopoly prices. On the state, local and federal level regulatory commissions were established to regulate the prices (rates) natural monopolies were allowed to charge (as well as some other aspects of their business activities.) It is this approach to dealing with natural monopoly that we will discuss here.

FIG 9.5


Consider Figure 9.5 above. It represents the cost and demand situation facing a natural monopoly, the Hilton Bay Electric Power Co. (HBEP). (NOTE: The output shown along the horizontal axis is measured in "millions" of some unit)

QUESTION 9.25: How can we tell that Figure 9.5 represents a natural monopoly?
ANSWER 9.25: Because in the output range shown in the graph the industry's LRAC curve is continuously declining.

QUESTION 9.26: If Figure 9.5 represented the demand and cost situation of an unregulated monopolistic firm, what would be its price and output level?

ANSWER 9.26: The profit-maximizing price-output combination would be determined as usual by the MR = MC rule. Note the point where the MR curve intersects the (horizontal) MC line. The arrow between " 4 " and " 6 " along the horizontal axis indicates that the unregulated monopoly would produce approximately 5.25 million units and charge a little less than $\$ 12$ per unit. The firm earns monopoly profits ( $\square$ ) of approximately 20 million and $P>M C$. Its Lerner index $(\mathrm{L})$ is approximately 0.67 .

PROBLEM 9.2: Show that the numerical results in ANSWER 9.26 are correct.
SOLUTION 9.2: The solution to this problem is left as an exercise for the reader.

Assume that the public in the state of Caltex is unhappy about the monopoly prices charged by this firm (and the monopoly profits they earn). In response a commission called the Caltex Public Utilities Commission (CPUC) is set up to "regulate" the monopoly. Its members are honest and patriotic and wish to act in the public interest. But they are attorneys, ex-lobbyists and former politicians and the typical member had just one introductory economics course in college.

QUESTION 9.27: When we say "regulation" here we mean mainly (but not exclusively) price regulation. Is there a pricing rule that the members of the CPUC learned in college that deals with pricing in the public interest (i.e., so-called optimal pricing)?

ANSWER 9.27: Yes! The $P=M C$ rule (or the "marginal cost pricing rule") deals exactly with this question. As a first step then the CPUC would try to impose marginal cost pricing on the newly-regulated firm.

QUESTION 9.28: If the CPUC imposes a marginal cost pricing rule on the company, what would be the resulting price-quantity combination?

ANSWER 9.28: The (socially optimal) price-quantity combination is found at the intersection of the demand curve and the MC curve (at point $m$ ). The regulated price $\left(\mathrm{P}^{*}\right)$ is $\$ 4$ and the resulting quantity demanded $\left(Q^{*}\right)$, is shown by the third arrow from the origin on the horizontal axis, so $Q^{*}=10$ million.

QUESTION 9.29: Is this price feasible? That is, is it actually possible for the CPUC to impose such a price on HBEP?

ANSWER 9.29: The answer is no, since $\mathrm{P}^{*}$ is below long-run average cost. This is so because the LRMC line lies everywhere below the LRAC curve. (See REM 9.6.) In other words, since a continuously declining LRAC curve is typical of natural monopolies, LRMC must be below LRAC. But a price less than average cost (which is of course equivalent to TR < TC) is not feasible because a firm cannot survive under such conditions: In the long run it is not covering all of its costs (including all of its opportunity costs);

QUESTION 9.30: What is the size of the loss suffered by HBEP if marginal cost pricing is imposed on them?

ANSWER 9.30: The loss per unit is equal to ( $\mathrm{P}^{*}$ - LRAC) or the distance between point $m$ and point $d$ in FIG 9.5, which (approximately) equals $\$ 2$. The total loss equals $Q^{*} x\left(P^{*}-L R A C\right) \approx 10 \times 2 \approx \$ 20$ million.

QUESTION 9.31: Is the fact that in natural monopolies LRMC < LRAC the only reason why strict marginal cost pricing is difficult to impose on regulated firms?

ANSWER 9.31: No. A second reason is that courts in the United States have told regulatory commissions that regulated firms must be allowed to earn a "fair" return. The phrase "fair return" sounds to the economist like "the opportunity cost of capital" or a "normal profit." Hence the prices (or rates) imposed on the regulated firm must be such as to allow them to earn a normal profit.

QUESTION 9.32: What price would the CPUC impose on the monopoly so that they would earn a normal profit (or a "fair return")?

ANSWER 9.32: The price $\left(\mathrm{P}_{\mathrm{f}}\right)$ is found at the intersection of the demand curve and the LRAC curve (point $c$ ). Then $P_{f} \approx \$ 6.50$ : (Note the arrow along the vertical axis between " 4 " and " 8 "). The corresponding quantity demanded $\left(\mathrm{Q}_{\mathrm{f}}\right)$ is 8.50 (again approximately: note the arrow along the horizontal axis to the right of " 8 "). So $P_{f}=$ LRAC and In effect the regulatory commission is imposing average cost pricing on the natural monopoly.

Are we happy with the solution reached in ANSWER 9.32 above? On the surface it appears that we have made an improvement: the price is below the unregulated price and the quantity produced and sold is higher than would be the case in an unregulated monopoly. On the other hand it is still the case that $P>M C$ and we have not achieved the "ideal" allocation of resources.

QUESTION 9.33: Is there any way to achieve an optimal allocation of resources in natural monopolies which somehow "mimics" marginal cost pricing?

ANSWER 9.33: For more than half a century economists have wrestled with this question and have come up with a number of ingenious solutions. One of the most widely accepted is called a two-part tariff.

DEF 9.7: A two-part tariff is defined as a pricing scheme which consists of two parts: a fixed payment which is required to gain access to the good or service and is independent of the quantity purchased and a second part which is proportional to the quantity (or "volume") purchased.

EXAMPLE 9.11: In some parts of New Jersey individuals and business firms obtain water from a privately-owned, regulated water utility, the New Jersey-American Water Company. The prices, or rates, paid depend on many factors, including geographic location, various customer characteristics and even the size of the meter employed to monitor water usage. Currently, a customer with a 2 " meter pays a fixed fee of $\$ 84.80$ per month and $\$ 0.59405$ per 100 gallons of water used.

QUESTION 9.34: How does a two-part tariff "solve" the problem of the seeming infeasibility of marginal cost pricing?

ANSWER 9.34: The regulatory commission imposes the marginal cost pricing rule on the firm, i.e., $\mathrm{P}=\mathrm{MC}$. But we found in ANSWER 9.29 that this results in losses to the regulated firm. For the firm depicted in FIG 9.5 these losses equal $\$ 20$ million. (See ANSWER 9.30). This is then dealt with by using a fixed payment (sometimes called a "lumpsum" payment) equal to those losses. Assume for example that there are a million identical customers purchasing 10 million units of the regulated good from the firm. Then regardless of the quantity purchased by each customer, they have to pay a $\$ 20$ fixed fee to have access to this good. This step eliminates the losses resulting from adherence to the marginal cost pricing rule.

## Rate-of-Return Regulation

The reality of the regulation of natural monopolies is a bit more complicated than our discussion above indicates. In the United States, Canada and several other countries regulatory commissions have, at least until recently implemented regulatory policy by establishing an "allowed rate of return." (This approach is called "rate-of-return regulation".) In principle this so-called allowed rate of return should be equal to what a nonmonopoly with similar characteristics would earn. A simplified way to calculate rates of return is shown in equation (1).

$$
\begin{equation*}
\overline{\mathrm{r}}=\frac{\mathrm{TR}-\mathrm{TC}^{\prime}}{\mathrm{RB}} \tag{1}
\end{equation*}
$$

We write $\bar{r}$ to indicate the allowed rate of return. We write $T C^{\prime}$ to indicate that in calculating the rate of return, all costs should be included except those that constitute return on the capital invested in the firm. $R B$ stands for the rate base, i.e., the capital invested in the firm on the basis of which the allowed rate of return is calculated.The regulated firm is then permitted to charge prices (rates) so that its revenue is sufficient for it to earn the allowed rate of return.

EXAMPLE 9.12: Caltex Electric Power (CEP) is a regulated natural monopoly. Their rate base, calculated according to rules established by the regulatory commission equals $\$ 100$ million. Its total revenue in 200X is expected to be $\$ 180$ million and its total cost (not including the opportunity cost of capital) is expected to be $\$ 166$ million. We can calculate CEP's rate of return in 200X as follows:

$$
\mathrm{r}=\frac{\$ 180 \text { million }-\$ 166 \text { million }}{\$ 100 \text { million }}=0.14
$$

If it happens that the allowed rate of return is $14 \%$, CEP will maintain its current prices. If $\bar{r}>14 \%$ and CEP is earning less than it is allowed they can go to the regulatory commission and ask for a price increase and in (rare cases) if $\bar{r}<14 \%$ and CEP is earning more than is allowed, the regulatory commission will ask for a price reduction.

## Problems Surrounding Rate-of-Return Regulation

Unfortunately rate-of-return regulation creates problems of its own. If the allowed rate of return is less than the unregulated rate of return, regulated monopolies lose their incentive to minimize costs (i.e., to be technically efficient). This is so because an increase in costs lowers the actual rate of return which the company earns; this would trigger a price increase so that the rate of return would be restored to the allowed level. Further, if there is an allowed (i.e., maximum) rate of return the firm loses its incentive to innovate (i.e., to develop new production techniques, organizational methods, etc.) since such innovation would not be rewarded by higher profits. Finally, under some circumstances the regulated firm has an incentive to "inflate the rate base", i.e., to employ more capital-intensive production methods than is technically efficient.

Over the last half century economists have come up with a large number of ingenious solutions to the problems created by natural monopolies and their regulation, but we shall not discuss them further here.

## PROBLEMS:

FIG P9.1


Table P9. 1

| $(1)$ | $(2)$ | $(3)$ | $(4)$ | $(5)$ | $(6)$ | $(7)$ | $(8)$ | $(9)$ | $(10)$ | $(11)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\mathbf{Q}$ | $\mathbf{P}$ | TR | MR | TFC | TVC | TC | AFC | AVC | ATC | MC |
| 0 | 240 | 0 | --- | 80 | 0 | 80 | --- | --- | --- | --- |
| 1 | 230 | 230 | 230 | 80 | 110 | 190 | 80.00 | 110.00 | 190.00 | 110 |
| 2 | 220 | 440 | 210 | 80 | 150 | 230 | 40.00 | 75.00 | 115.00 | 40 |
| 3 | 210 | 630 | 190 | 80 | 180 | 260 | 26.67 | 60.00 | 86.67 | 30 |
| 4 | 200 | 800 | 170 | 80 | 220 | 300 | 20.00 | 55.00 | 75.00 | 40 |
| 5 | 190 | 950 | 150 | 80 | 270 | 350 | 16.00 | 54.00 | 70.00 | 50 |
| 6 | 180 | 1,080 | 130 | 80 | 340 | 420 | 13.33 | 56.67 | 70,00 | 70 |
| 7 | 170 | 1,190 | 110 | 80 | 430 | 510 | 11.43 | 61.43 | 72.86 | 90 |
| 8 | 160 | 1,280 | 90 | 80 | 540 | 620 | 10.00 | 67.50 | 77.50 | 110 |
| 9 | 150 | 1,350 | 70 | 80 | 670 | 750 | 8.89 | 74.44 | 83.33 | 130 |
| 10 | 140 | 1,400 | 50 | 80 | 830 | 910 | 8.00 | 83.00 | 91.00 | 160 |
| 11 | 130 | 1,430 | 30 | 80 | 1,020 | 1,100 | 7.27 | 92.72 | 100.00 | 190 |

(1) Consider FIG P9.1 above, which summarizes the cost and demand situation of a monopoly.
(a) What is the monopoly's profit-maximizing price $\left(\mathrm{P}_{\mathrm{m}}\right)$ and output level $\left(\mathrm{a}_{\mathrm{m}}\right)$ ?
(b) At the $\left(P_{m}, Q_{m}\right)$ price-quantity combination, what is the monopoly's total revenue (TR), total cost (TC) and total profit ( $\Pi$ )? (Use the "tick marks" along the vertical axis to estimate the approximate $\$ / \mathrm{Q}$ figures.)
(c) Is this "normal" or "above-normal" profit? Explain.
(d) If above-normal, will it be "competed away?" Explain.
(e) What is the monopoly's Lerner Index?
(2) Consider Table P9.1 above. Columns (1) and (2) show the demand schedule facing a pure monopoly. Answer the following questions based on the information contained in the table. (Assume only "whole" units can be produced, e.g., 4 units but not 4.3 units.)
(a) What is the monopoly's profit-maximizing price ( $\mathrm{P}_{\mathrm{m}}$ ) and output level $\left(\mathrm{a}_{\mathrm{m}}\right)$ ?
(b) At $Q_{m}$, what are the monopoly's total revenue, total cost and total profit?
(c) Calculate the monopoly's total profit at any other output level and compare the results.
(d) What is the monopoly's "Lerner index"?
(e) Assume Table P.9.1 describes the monopoly's long-run cost structure. What price would/should a regulatory commission, acting in the "public interest," impose on this monopoly?
(f) Is this price feasible? Explain.
(3) Consider FIG P9.2 below. The negatively sloped line represents the demand for widgets. Initially the industry is purely competitive. Average cost of production is constant and equals $\$ 20$.
(a) What is the industry's equilibrium price and output level?
(b) What is the industry's consumers' surplus (CS)?
(c) If the industry becomes monopolized and there is no change in demand and cost conditions, what is the monopoly's price and output level?
(d) Use this diagram to explain in the simplest terms why monopoly is undesirable from the consumer's point of view.
(e) What is the CS under monopoly?
(f) What is the loss of CS that results from the monopolization of the industry?
(g) What is the monopoly's profit?
(h) What is the deadweight loss that results from the monopolization of the industry?
(i) What is the monopoly's Lerner index?
(4) Figure P9.3 below (page 31) represents the cost and demand situation facing a natural monopoly.
(a) Can you tell why?
(b) If the monopoly is unregulated, what would be its price and output level?
(c) What would be its total revenue, total cost and total profit?
(d) What would be the unregulated monopoly's Lerner index?
(e) If the monopoly were regulated what would be the "socially optimal" price ( $\mathrm{P}^{*}$ ) and the resulting quantity demanded if it were feasible?
(f) Is $\mathrm{P}^{*}$ feasible? Explain.
(g) If $\mathrm{P}^{*}$ is not feasible what price could a regulatory commission impose which would allow the company to earn a "fair return," i.e., the opportunity cost of capital?
(5) Does marginal cost pricing represent an ideal solution to the problem of natural monopoly? Why or why not? Explain.
(6) What problems are created by rate-of-return regulation that make it a less than ideal solution to the problem of natural monopoly? Explain.
(7) Assume FIG P9.3 below depicts a natural monopoly. Explain all your answers.
(a) If the firm is unregulated, what would be its price and output?
(b) What would be the socially optimal regulated price and output level?
(c) If the firm is allowed to earn a "fair return," what would be the resulting price and output?
(d) If a two-part tariff is imposed on the firm, what would be the resulting price and quantity?
(e) What would be the total "lump sum" fees customers would have to pay?

FIG P9. 3

Figure 25.P3


