After developing the theory of perfect competition, we now move to the second simplest model: monopoly. The monopolist faces a profit-maximization decision that is more complex than that of the price taker. Monopoly price and output decisions are nevertheless simpler than they are for markets of sellers of differentiated products, particularly when the number of such sellers is small.

I illustrate the contrast between monopoly and perfect competition through a simple fable about the (temporary) monopolization of a competitive market. A **cartel** exists when many producers conspire to reduce output, thereby raising price and increasing profit above the level attainable under competitive conditions. We see how a **fallacy of composition** implies that while firms would like all *other* producers to decrease output, they would prefer to increase their own output in the wake of a cartel decision. This causes instability in cartel agreements that require some means by which the cartel agreement can be enforced. We consider different imperfect alternatives: **cartel** contracts (made illegal by most anti-trust legislation), government regulation, or organized crime.

We then turn our attention to the single producer monopoly, which can be created either by government franchise or market conditions. We then distinguish the efficiency effects of charging a single monopoly price to all buyers versus different forms of **price discrimination: first degree price discrimination –** charging the highest price for each unit, **second degree price discrimination** – charging different prices to each consumer, and **third degree price discrimination** – charging different prices to large, identifiable groups of consumers.

We conclude the chapter with a discussion of remedial measures against monopoly, from deterring the formation of artificial monopolies to regulating natural monopolies.

Defining Monopoly

In theory it is easy to define monopoly, a word formed from two Greek roots: mono, meaning one (as in **one seller**) and poly, meaning many, (as in many buyers). A particular and interesting market is established when there is mono a mono – one seller (monopoly) vs. one buyer (monopsony), which we call a **bilateral monopoly**. We will investigate that market type in more detail in chapter 13. Throughout the history of economics, however, the concept of monopoly has encouraged ideological inferences. As we showed in chapter 8, perfect competition is economically efficient because it maximizes the sum of consumer and producer surplus. We will learn that monopoly is inefficient precisely because it generates excess burden: the loss to consumers exceeds the gains to producers. Those who urge a **laissez faire** attitude toward the market often deny the existence of monopoly; after all, every seller competes with every other seller for consumer expenditure, so it is erroneous to argue that a market is dominated by only one seller. Those who urge government regulation emphasize that sellers that confront negatively sloped demand curves cause economic distortions when they maximize profit. They err by *assuming* that government regulation is a costless activity whose benefits always exceed their costs. As I discussed in chapter one, whenever we pit real-world institutions against theoretical ideals, the ideal type always wins. The positive economic problem is to determine whether a theory of monopoly adds anything to our understanding of economic events; I believe it does. The normative question of whether one should prohibit monopoly behavior is much more difficult to determine objectively.

Monopoly and Competition

It is often easiest to introduce a new concept by contrasting the unfamiliar with the familiar one. We start with exploring how monopolization of a once competitive market changes the quantity produced and the price charged. Instead of immediately confronting the challenges of enforcing monopoly pricing on cartel members, we examine a simple fable in which one owner of a competitive firm be-

comes a multi-plant monopolist. Table 9-1 reprises the example of one in 100 producers of a generic product that we'll call widgets. Marginal cost is \$10 times the number of units produced; fixed costs equal \$210 per firm. The shut-down price is \$10, while the break-even price is \$70. The firm's supply equation is given by $MC = 10q = p \rightarrow q_s = 0.1p$. Multiplying by the number of firms in the market, we get the market supply curve as $Q_s = 100 \times q_s = 100 \times (0.1p) = 10p$.

Table 9-1

Output and Cost Schedules for a Hypothetical Price-Taking Firm

Output	Total Cost	Marginal Cost	Average Variable Cost	Average Total Cost
0	\$210			
1	\$220	\$10	\$10.00	\$220.00
2	\$240	\$20	\$20.00	\$120.00
3	\$270	\$30	\$30.00	\$90.00
4	\$310	\$40	\$40.00	\$77.50
5	\$360	\$50	\$42.00	\$72.00
6	\$420	\$60	\$45.00	\$70.00
7	\$490	\$70	\$48.57	\$70.00
8	\$570	\$80	\$52.50	\$71.25
9	\$660	\$90	\$56.67	\$73.33
10	\$760	\$100	\$61.00	\$76.00
11	\$870	\$110	\$65.45	\$79.09
12	\$990	\$120	\$70.00	\$82.50
13	\$1,120	\$130	\$74.62	\$86.15
14	\$1,260	\$140	\$79.29	\$90.00
15	\$1,410	\$150	\$84.00	\$94.00

Table 9-2 presents the market supply and demand schedules for the widget industry. Widgets would be priced out of the market at p = \$140, while quantity demanded increases by 10 units for each \$1 reduction in the price. Stating this relationship in equation form, $Q_d = 1{,}400 - 10p$. The equilibrium price at p = \$70 is obvious from Table 9-2. Equilibrium can also be established by setting $Q_d = Q_s = Q_e$ yields:

$$Q_d = 1{,}400 - 10p = 10p = Q_s = Q_e$$

$$1{,}400 = 20p$$

$$P_e = 1{,}400/20 = 70$$

$$Q_d = 1{,}400 - 10(70) = 700; Q_s = 10(70) = 700 = Q_e$$

Because average total cost also equals \$70 when q = 7, the typical firm earns zero economic profit when the market price is \$70; this implies that p = \$70 is the *long-run equilibrium price*. The fact that firms increase their output as price increases implies that the long-run equilibrium price of \$70 is also the profit-maximizing price. At p = \$70 total economic profit is \$0; at any other price, the 100 firms collectively experience large losses. In desperation, the 100 producers gather at a moderately sized seedy bar to discuss their options. After much complaining and even more drinking, they decide to par-

ticipate in a lottery. Each owner will sign the deed to his/her plant and place the deed in the bartender's hat. The bartender will draw one deed, whose owner will become the sole owner of the 100 plants; each loser will be retained as plant managers with a salary equal to their former income as entrepreneur.

Table 9-2: Revenue and Cost Schedules for 100-Firm Competitive Industry

Price	Quantity Supplied	Quantity Demanded	Point Price Elasticity	Industry Revenue	Industry Cost	Industry Profit
$140	1,400	0	-∞	$0	$126,000	-$126,000
$130	1,300	100	-13.00	$13,000	$112,000	-$99,000
$120	1,200	200	-6.00	$24,000	$99,000	-$75,000
$110	1,100	300	-3.67	$33,000	$87,000	-$54,000
$100	1,000	400	-2.50	$40,000	$76,000	-$36,000
$90	900	500	-1.80	$45,000	$66,000	-$21,000
$80	800	600	-1.33	$48,000	$57,000	-$9,000
$70	700	700	-1.00	$49,000	$49,000	$0
$60	600	800	-0.75	$36,000	$42,000	-$6,000
$50	500	900	-0.56	$25,000	$36,000	-$11,000
$40	400	1,000	-0.40	$16,000	$31,000	-$15,000
$30	0	1,100	-0.27	$0	$15,000	-$15,000

After a considerable delay, the bartender draws the winning ticket, Mona Polist (who else?). She surveys her options. According to Table 9-2, the equilibrium price of $70 appears to be the best available. Was the lottery nothing but a charade?

Just then an economist enters the bar and orders a whiskey with a Shirley Temple Chaser. "If my honor did not bind me to encourage economic efficiency, I could tell you how to increase your profits," he says, apparently to his own reflection in the mirror above the bar. "Of course, my fee for such advice would not be insubstantial – if my conscience would permit it, that is," he continues, still maintaining his aloof demeanor.

The monopolist designate slaps a $1000 bill onto the bar in front of the wide-eyed economist. Regaining his composure, he jots down the information in Table 9-3 on the back of a cocktail napkin.

"A group of competitors cannot raise prices because they raise their output, creating a market surplus. This is what Adam Smith meant by the invisible hand," says the economist somberly.

"What?" asks one of the losers, drowning his sorrows in beer.

"Something about somebody named Smith; an alias, no doubt. Sounds like some sort of con artist; you know, hand quicker than the eye and all that," his companion volunteers.

"As I was saying," continues the economist, glaring at the disturbance, "once all industry output is controlled by a single seller, surplus can be prevented by anticipating the tendency of consumers to buy less as the price rises. Instead of producing 8000 units at a price of $80, produce only 600. That way, although the revenue would fall by $1,000, costs would decrease by $7,000." The economist now addresses the monopolist directly; his earlier aloofness forgotten is his enthusiasm for an attentive audience.

"I see," muses Mona, "that would mean more profit by producing less. But of course! If 600 units can be sold for $80 each, selling 700 units at $70 each means that last hundred units only brings in $10 each. Since each unit costs an extra $70 to produce, that's like throwing $60 away. But let's see, if I can sell 500 units for $90 each, then I can earn $45,000 revenue. In exchange for a revenue loss of only $3,000(that's $30 per unit), I could cut costs by $6,000 ($60 per unit). Why, that reduction in output actually increases profit by $3000. You know, being a monopolist may be profitable after all.

"We could continue with trial and error, but economists like to be precise," proclaims the economist, warming to his task. The demand equation is p = $140 - .1$Q$, so that marginal revenue is $140 - .2$Q$. [1] We know that marginal cost is given by $MC = .1Q$, so if we set marginal cost equal to marginal revenue and solve for quantity, we get: $MC = MR \rightarrow 0.1Q = 140 - 0.2Q \rightarrow Q_m = \frac{140}{3} = 466.67$.

"Plugging the profit-maximizing quantity into the demand equation, we find p = 140 - .1(466.67) = $93.33. Total revenue is (93.33)(466.67) = $43,556; total cost is $21,000 + .05Q^2 = $31,889 which maximizes economic profit at $11,667."

"So you are telling me to reduce output until marginal cost equals marginal revenue, and then set price according to the demand equation," intones the monopolist.

"Far be it for me to tell you what to do. We economists must maintain our objectivity. However, if you need financing for your enterprise, I know where you can get very good terms."

Table 9-3: Profit Opportunities for Monopolist who Restricts Output

Quantity Produced	Market Price	Total Revenue	Total Cost	Monopoly Profit	Marginal Revenue	Marginal Cost
0	$140	$0	$21,000	-$21,000	$140	
100	$130	$13,000	$21,500	-$8,500	$120	$10
200	$120	$24,000	$23,000	$1,000	$100	$20
300	$110	$33,000	$25,500	$7,500	$80	$30
400	$100	$40,000	$29,000	$11,000	$60	$40
466.67	$93.33	$43,556	$31,889	$11,667	$46.67	$46.67
500	$90	$45,000	$33,500	$11,500	$40	$50
600	$80	$48,000	$39,000	$9,000	$20	$60
700	$70	$49,000	$45,500	$3,500	$0	$70
800	$60	$48,000	$53,000	-$5,000	-$20	$80
900	$50	$45,000	$61,500	-$16,500	-$40	$90
1000	$40	$40,000	$71,000	-$31,000	-$60	$100
1100	$30	$33,000	$81,500	-$48,500	-$80	$110
1200	$20	$24,000	$93,000	-$69,000	-$100	$120
1300	$10	$13,000	$105,500	-$92,500	-$120	$130
1400	$0	$0	$119,000	-$119,000	-$140	$140

"Thanks, but my financial advisor will be handling that now that you've shown me how to increase my profits. What is your name, anyway?"

---

[1] Given $p = \$140 - .1Q \rightarrow TR = pQ = \$140Q - 0.1Q^2; MR = \frac{d(pQ)}{dQ} = \$140 - 0.2Q$.

The monopolist turns away to check her latest text message from one of her 99 suitors. When she looks up, the economist is gone. On the bar is a little silver calculator.

"Who was that fast talker, anyway?" murmurs a drunk at the bar.

"I don't rightly know his name," says his companion, "but I've heard him called the Loan Arranger."

To maximize profit a monopoly seller must realize that the price she can charge is limited by the output she produces. Because price falls as output increases, marginal revenue no longer equals price, but instead equals price for the last unit, minus the discount realized by all other buyers who would otherwise pay a higher price. In a competitive market, one receives transitory profits by being in the right place at the right time and by producing efficiently. A monopolist, on the other hand, is able to obtain more by producing less.

Monopoly Price and Marginal Revenue

The popular notion of the monopolist is of a firm whose owner can charge any price he or she wants. In a sense, that is correct; a monopolist, like any seller in a market economy, is free to charge whatever price (s)he desires. In a competitive market, a seller can charge any price; however, charging any price other than market price will have severely negative consequences. A monopolist can charge any price as well, but buyers have the freedom of whether to buy at all, and to determine how many units to buy if they so choose. Ultimately, the competitive seller is constrained by the market price; the monopoly seller is constrained by the market demand curve. If the competitive seller charges more than the market price, quantity sold falls instantly to zero; if the monopoly seller raises prices, quantity sold gradually declines until, at some price, the good is priced out of the market.

If the monopolist knows the demand for the firm's product, the exact relation between output and marginal revenue can be derived from the demand equation. In Figure 9-1, the monopolists commodity is priced out of the market when $p \geq p_{max}$, and price must decline by b dollars to sell an additional unit; hence, $p = p_{max} - bQ$. At price p_0 the monopolist is selling Q_0 units of output. The area of the rectangle, p_0AQ_00 measures total revenue at the output/price combination. If output increased to Q_1, the seller would ordinarily reduce the price on all units, resulting in a new revenue of p_1Q_1, given by rectangle p_1BQ_10. Labeling $(Q_1 - Q_0)$ as ΔQ, and $(p_0 - p_1)$ as Δp, the net change in total revenue for the increase in output is: $\Delta TR = \Delta(pQ) = p_1(\Delta Q) - Q_0(-\Delta p) = p_1(\Delta Q) + Q_0(\Delta p)$

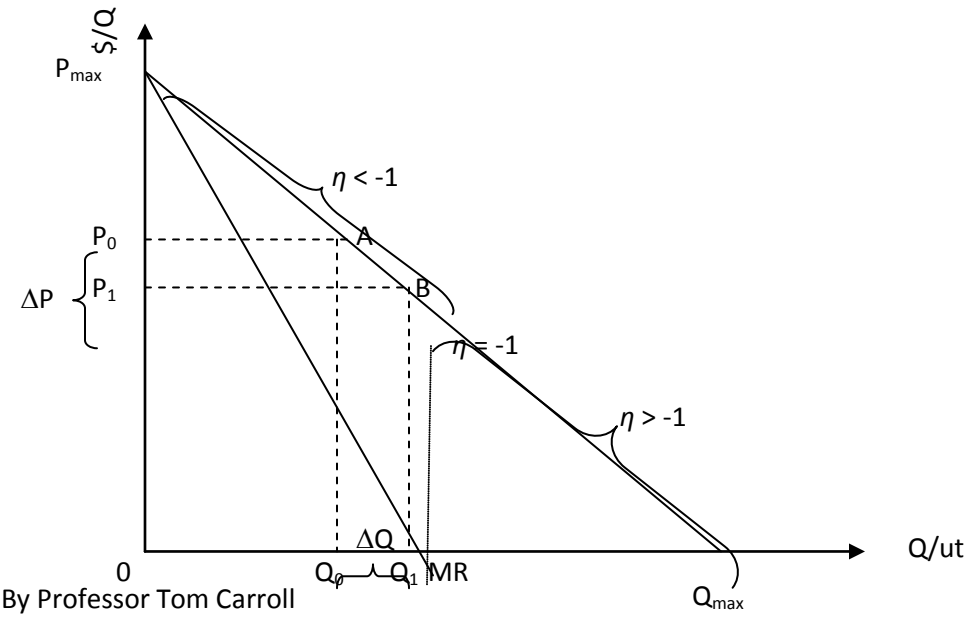


Figure 9-1

Dividing ΔTR by ΔQ yields the formula for **marginal revenue**, the change in total revenue divided by the change in output:

$$MR = \frac{\Delta TR}{\Delta Q} = p_1 \times \frac{\Delta Q}{\Delta Q} + Q_0 \times \frac{\Delta p}{\Delta Q} = p_1 + Q_0 \times \frac{\Delta p}{\Delta Q}$$

Factoring the price term out of this expression reveals a close relationship between marginal revenue (*MR*) and the own-price elasticity of demand (η) for the monopoly product:

$$MR = \frac{\Delta TR}{\Delta Q} = \frac{\Delta (pQ)}{\Delta Q} = p\left(1 + \frac{\Delta p / p}{\Delta Q / Q}\right) = p\left(1 + \frac{1}{E_{dp}}\right)$$

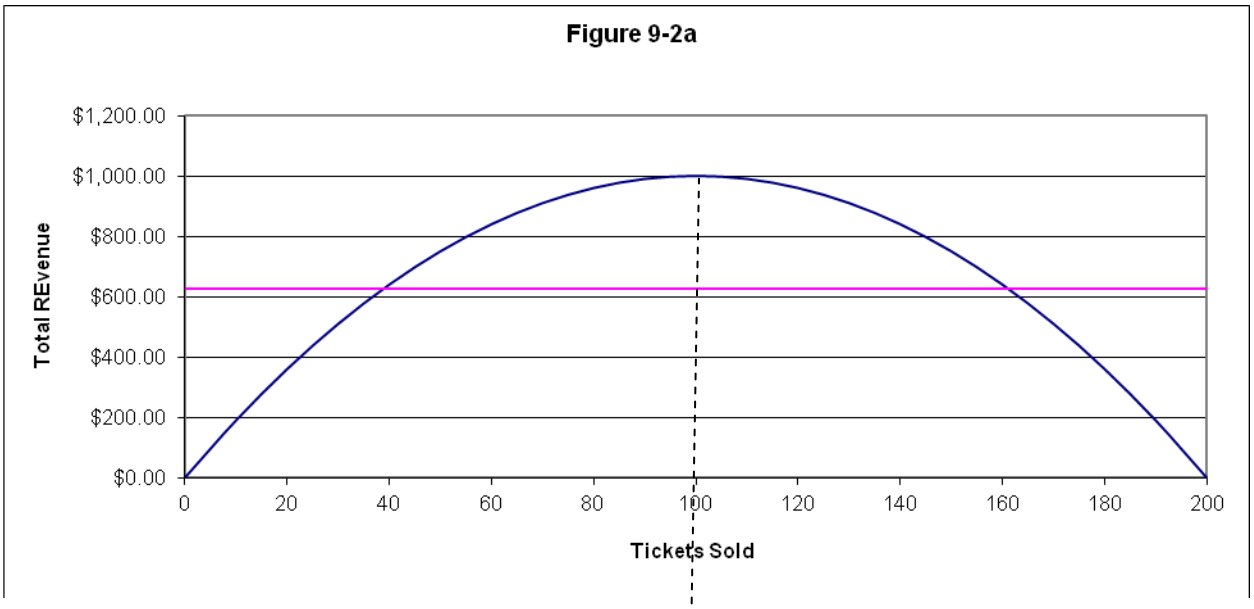
If demand is price elastic ($E_{dp} < -1$), the reciprocal of E_{dp} is a number between zero and negative one $\left(-1 < \frac{1}{E_{dp}} < 0\right)$; this makes the term $\left(1 + \frac{1}{E_{dp}}\right) > 0$, so that marginal revenue is positive for this range of output. At the midpoint of a linear demand curve, $\eta = -1$, and $\left(1 + \frac{1}{E_{dp}}\right) = 0$, implying that marginal revenue is zero. Total revenue for the monopolist is maximized at that rate of output where demand for the product is unit elastic. When demand is price inelastic, $\left(1 + \frac{1}{E_{dp}}\right) < 0$, implying that marginal revenue is negative.

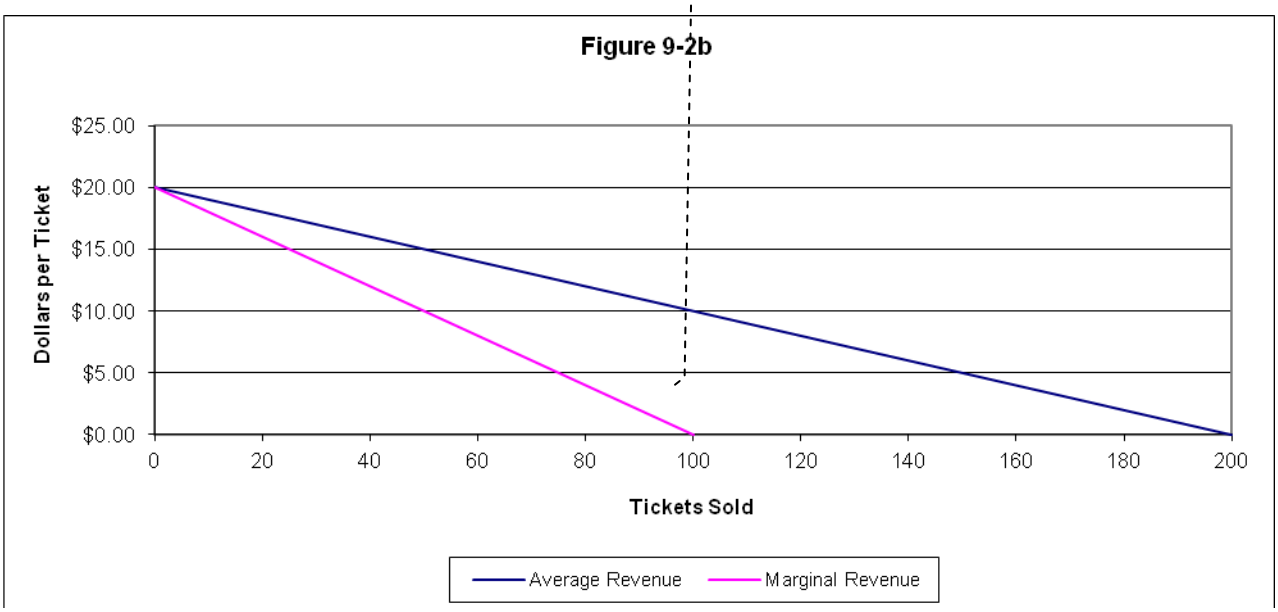
To understand the relevance of marginal revenue to the monopoly seller's output decision, we will concentrate on the special case of a monopoly with zero marginal cost. Examples include movie theaters, sports stadia[2], or concerts; once the program is staged, additional patrons can attend the event without increasing the costs to the promoter. Since all costs are fixed, profits are maximized when the positive difference between revenue and fixed cost is maximized (or the negative difference is minimized); this will occur when total revenue is maximized.[3]

In Figure 9-2 a monopoly theater operator confronts a horizontal cost function, *FC*. During any performance, costs are invariant with respect to the number of tickets sold. The inverse demand curve gives the market-clearing price for each number of tickets. The equation for the demand equation is $p = p_{\max} - bQ$, where p_{max} is the intercept, the price above which quantity demanded equals zero. The marginal revenue line has twice the slope as the inverse demand curve:

[2] My stupid spell checker does not understand that "stadia" is the plural of the Latin "stadium."
[3] We ignore the possibility that a monopolist might wish to trade off lower ticket revenue in order to increase revenue from snacks or concessions, since the latter involve positive marginal costs.

$MR = p + Q\left(\frac{\Delta p}{\Delta Q}\right) = a - bQ - bQ = a - 2bQ$. Hence, marginal revenue is positive until the price reaches the midpoint of the demand curve, at which point marginal revenue is zero and revenue is maximized. It follows that a profit-maximizing monopolist would never sell in the inelastic region of the demand curve, because over that region, $MR < 0$.[4]





If the theater could seat 100 patrons or less, the owner would set the price at capacity, since when the demand is price elastic, as the price falls, total revenue increases. If the theater could seat more than 100 patrons, the owner would maximize revenue, and hence the profit from ticket sales, by selling only 100 tickets, where demand is unit elastic. Selling additional seats would reduce the revenue

[4] The exception would be the multi-product monopolist who may sell one product (e.g., movie tickets) in the inelastic region of the demand curve to increase sales of another product (e.g., popcorn).

from the first 100 seats by more than the additional seats would receive; that is what negative marginal utility means.

Profit Maximization with Positive Marginal Costs

Typically, a monopolist's costs will vary with output; maximizing profit means stopping production before revenue is maximized. Maximizing revenue implies that marginal revenue is zero; if marginal cost is positive, the seller could increase profit by reducing output, which would cause total costs to decline by more than total revenue, until production reached the point where marginal cost equals marginal revenue, where profit would be maximized. Given positive marginal cost, the monopoly seller will maximize profit only by producing in the elastic region of the (inverse) demand curve.

Returning to the example of the 100-plant widget monopoly, Figure 9-3 reprises the total cost and revenue functions for the monopolist. If nothing were produced, the firm would lose its fixed costs of $21,000 per day. As output increases, both revenue and cost increase; losses decrease because revenue is rising faster than costs are (*MR* > *MC*). When output reaches 187.78, total revenue equals total cost and profit is zero; the seller has broken even. However, since revenue is still increasing faster than cost, the firm can increase profit by increasing output. When output reaches 466.67, marginal revenue and marginal cost are both equal to $46.67, and profit is maximized. If the monopoly were to increase output until it produced 700 units, total revenue would be maximized, and, coincidentally, profit would again have returned to zero.

Figure 9-3

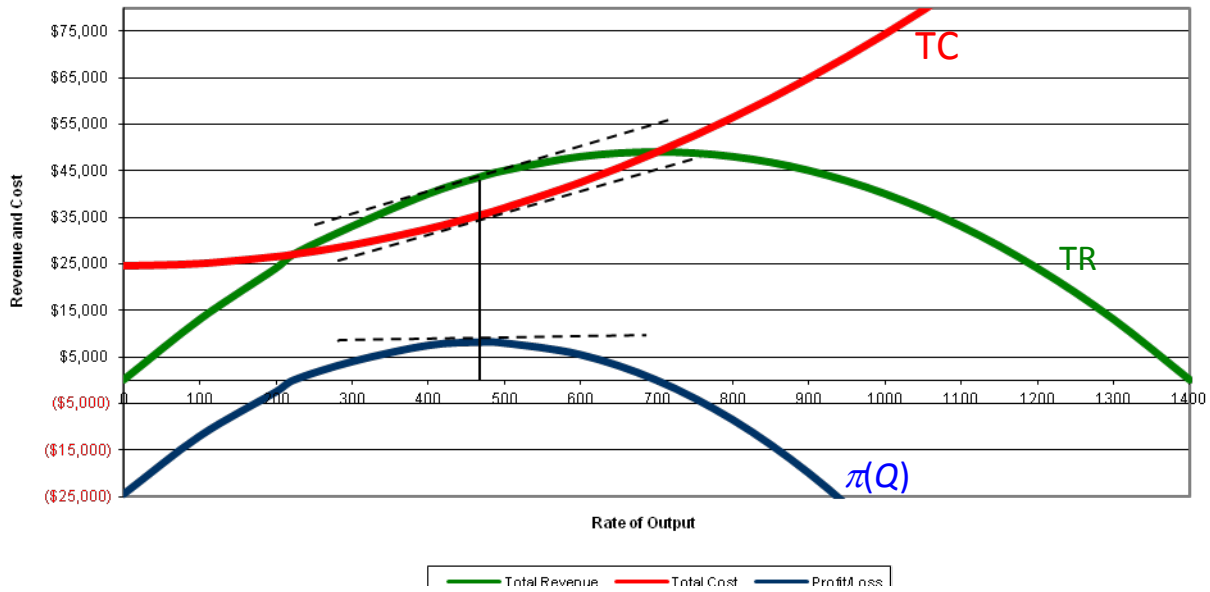
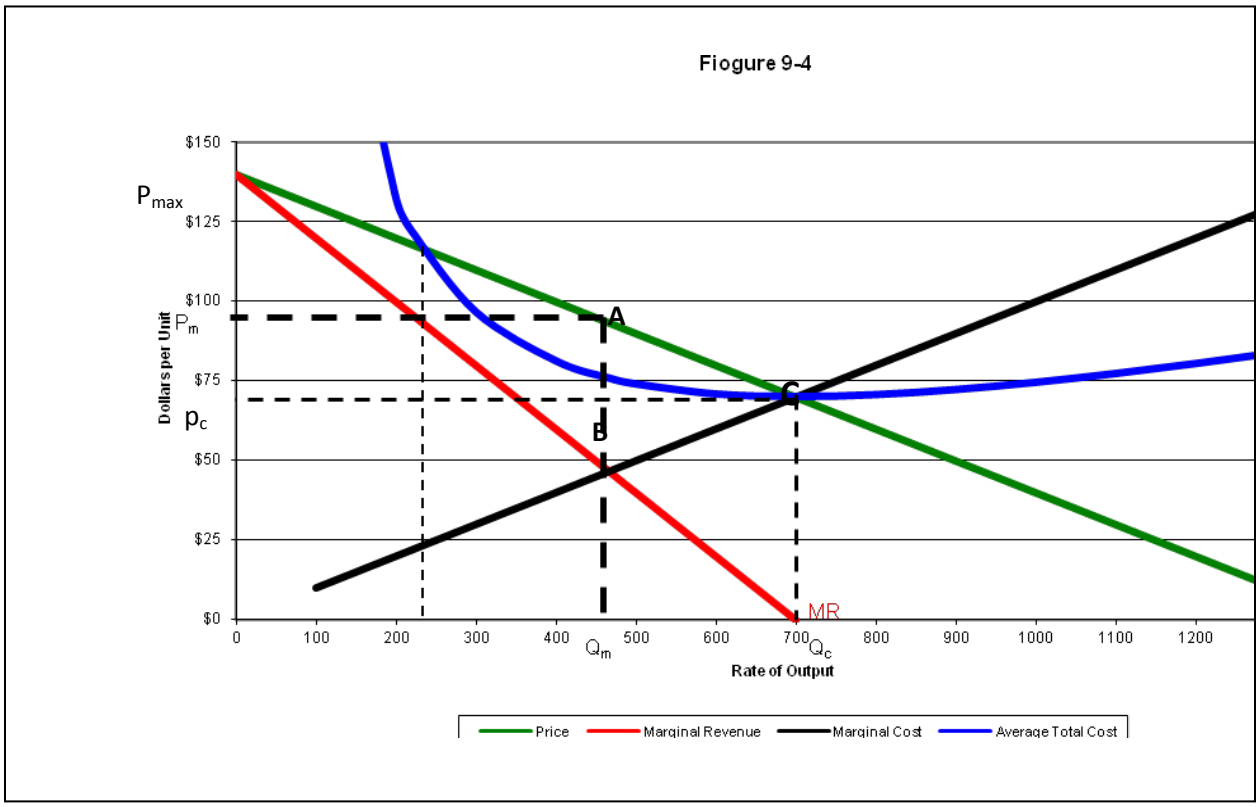


Figure 9-4 contains the same information as figure 9-3, except that information is displayed in dollars per unit. Two break-even rates of output, *Q* = 187.78 and *Q* = 700 correspond to the intersection of the inverse demand curve and the average cost curve. The positive difference between price and average cost represents profit per unit of output; total profit is the rectangle formed by multiplying profit per unit times the rate of output: $\pi = (p - ATC)Q$. The change in profit is the change in revenue minus the change in cost: $\Delta\pi = \Delta R - \Delta C$; dividing by the change in output, ΔQ, $\frac{\Delta\pi}{\Delta Q} = \frac{\Delta R}{\Delta Q} - \frac{\Delta C}{\Delta Q} = MR - MC$.

Below Q_m, $MR > MC$ and profit is increasing. Above Q_m, $MR < MC$ and profit is decreasing. It is precisely where $MR = MC$ that profit is maximized.



The Inefficiency of Monopoly

Figure 9-4 also shows the efficiency aspects of monopoly. Recall that under perfect competition, the market equilibrium price maximizes the sum of producer and consumer surplus. Furthermore, under long-run competitive equilibrium, the price is set at the lowest sustainable level, so that producer surplus exactly equals fixed costs, so that all sellers receive zero economic profit. As a result of monopoly, output decreases from Q_c to Q_m, while price increases to p_m. Under perfect competition, consumer surplus would equal $\frac{1}{2}(p_{max} - p_c)Q_c$ Under monopoly, consumer surplus decreases to $\frac{1}{2}(p_{max} - p_m)Q_m$. Part of the lost consumer surplus is transformed into producer surplus. However, because average costs at Q_m are greater than average costs at Q_c, part of the effect of the higher price is absorbed through higher average cost. In addition, the producer loses the producer surplus on the $Q_c - Q_m$ units that were never produced. The consumers also lose the consumer surplus on the units not produced. The area of triangle **ABC** in Figure 9-4 represents the **excess burden** or **dead-weight loss** caused by monopoly. The deadweight loss is the sum of consumer and producer surplus of the units that are not produced under monopoly.

In one sense, the merits of monopoly vs. competition involve issues of wealth distribution: while consumers are worse off because they lose consumer surplus, monopolists are better off because they gain producer surplus. The case against monopoly, however, is one of efficiency; since the cost to consumers exceeds the benefits to monopolists, society would be better off if output were produced under competitive circumstances. The winners of that redistribution of benefits would be able to compensate the losers, with something left over. The problem, of course, is such compensation never takes

place; the only way for an economic agent to increase producer surplus above that allowed under competitive market conditions, is to establish and hold on to monopoly power.

The distributional consequences of monopoly are complicated even further by the **capitalization of economic profit** in the market value of the monopoly firm. In Figure 9-5 we imagine a monopoly firm maximizing profit by setting output at Q_m and setting price at p_m. Taking the difference between monopoly price and average cost yields the per unit profit of (π_m/Q_m). After enjoying these profits for a number of years, the monopolist decides to retire and sell his company to the highest bidder. Whether he chooses to sell to one owner or to a group of shareholders does not concern us now, although a recall from chapter seven that the executives of the corporate entity may be less diligent in maximizing future profits (that is, perpetuating monopoly profit) than would a single proprietor. Our point is that potential buyers would bid up to the present expected value; if π_m profits per year were expected in perpetuity, potential buyers would bid up to π_m/r, where r is the market rate of interest (adjusted for the risk that the monopoly power might be lost).

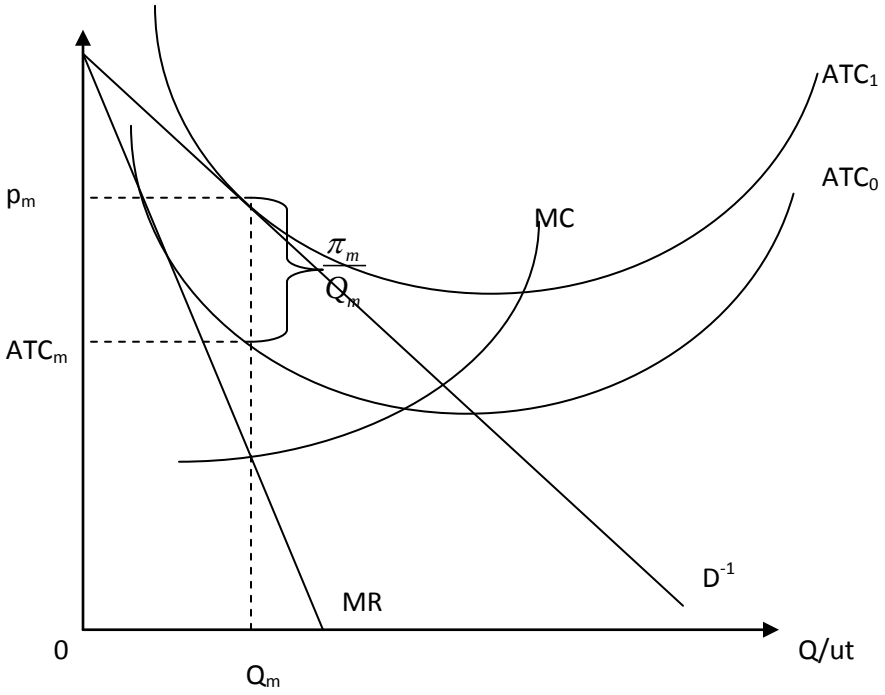


Figure 9-5: Capitalization of Monopoly Profit

If the monopoly firm is sold for a price equal to the present value of its expected future profits, the income stream which formerly constituted monopoly profit of π_m per year is now required for a "normal return" on the purchase price of the business $\left(r \times \left(\frac{\pi_m}{r} \right) = \pi_m \right)$. If the owner decides not to sell, the opportunity cost of continuing to operate the monopoly is the return she would have received if she sold the firm. Just as under competition, economic profit is a short-run phenomenon, which either is dissipated by expenditures to maintain the monopoly profit (e.g., attorneys' fees, bribes) or is absorbed into the value of the business. A nifty little sleight of hand by monopoly apologists is to point out that monopoly firms receive no higher return on investments than do competitive firms. The issue is not whether monopolies are fair, but whether they are efficient; and economists are of one mind on that score; monopolies are inefficient because of the deadweight loss (excess burden) they create.

Natural Monopoly

To this point we have analyzed a monopoly generated by the combination of all the producers in a competitive market. We will see later that such combinations are unlikely to survive because (1) they are illegal under U.S. antitrust laws, (2) monopoly profits not borne of economies of scale will encourage additional producers to enter the market, eventually eliminating economic profit, and (3) collusive agreements (cartels) are inherently unstable because of the incentive by members of the cartel to cheat.

At this point we investigate the **natural monopoly**, defined as the market type wherein a monopoly firm can produce at a profit, yet allow insufficient **residual demand** for additional firms to enter the market and dissipate that monopoly power. Figure 9-6 shows a producer with prolonged economies of scale, so that the producer encounters the market demand curve before achieving optimal plant size. The producer's long-run marginal cost curve is still decreasing when it intersects the prevailing marginal revenue curve at output Q_m. By producing with average cost curve ATC_m, the monopoly is able to cover the costs of producing the profit maximizing rate of output, Q_m and realize a monopoly profit by charging price p_m. Ironically, were the monopolist to expand production, it would be able to break even producing output Q_0 and charging price p_0 where the average cost curve intersects the demand curve. Note the huge loss of consumer surplus when natural monopolies are permitted to set a monopoly price which restricts output well below the level necessary to achieve cost-saving economies of scale.

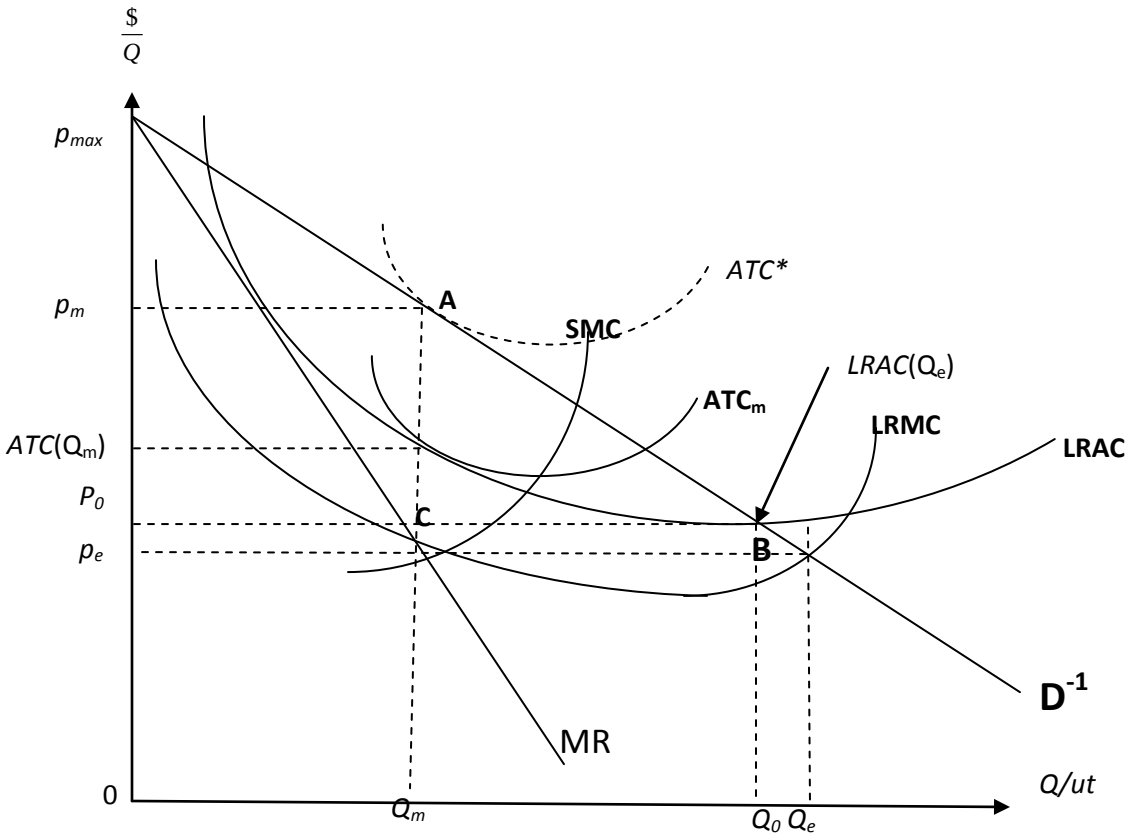


Figure 9-6: Natural Monopoly

At the monopoly price, consumer surplus equals $\frac{1}{2}(p_{max} - p_m)Q_m$ and economic profit equals $[p_m - ATC(Q_m)]Q_m$; excess burden equals the area of the triangle ***ABC*.** Because of this huge deadweight

loss from natural monopolies, many states and localities regulate the maximum price that natural monopolies can charge. We learned in chapter 2 that attempts to set ceiling prices in competitive markets are counterproductive, turning temporary profit-generating price increases into markets with perpetual shortages. Paradoxically, the imposition of a price ceiling on a monopolist often increases economic efficiency, at least in theory.

Recall that a monopoly maximizes profit by setting output where marginal revenue equals marginal cost, and then setting the price according to the demand curve. Marginal revenue is less than price because the seller must reduce the price to all buyers to sell each additional unit. But suppose a government regulator were able to prevent the monopolist from charging more than some specified price ceiling. Ideally, an economist would propose a price of p_e, where the long-run marginal cost intersects the market (inverse) demand curve. The best that the monopolist could do would be to set output at Q_e, incurring a loss equal to the difference between total revenue (p_eQ_e) and total cost ($ATC(Q_e)*Q_e$). This price-output combination would maximize gross consumer surplus. However, a subsidy of $(ATC(Q_e) - p_e)Q_e$ would be required to provide a "fair" return to the monopolist.[5] Typically, regulators set a price at p_0, where the average cost curve intersects the demand curve. Setting the price would still confront the monopolist with constant marginal revenue, whereby the monopolist's best option is to produce output Q_0 and receive revenue equal to total costs, including a "fair return on investment."

The problem with this regulation strategy is that it encourages the monopolist to seek rate increases to compensate for cost increases. At output Q_0, the monopolist is producing where the demand for its product is price inelastic. Clearly an unregulated monopolist would increase the price well into the elastic region of the demand curve, to where *MR = MC*. Regulation means that the monopolist cannot unilaterally raise price; the monopoly must propose price increases that are justified by cost increases. The monopolist no longer has an incentive to select inputs efficiently, since any cost increase could be passed on to customers (the rate payers). Indeed, by padding expense accounts, increasing executive pay, and holding stockholder meetings at plush resorts, the monopolist would attempt to push costs up to *ATC**, thereby justifying a return to the monopoly price. The only recourse for the regulatory authority is to micro-manage the monopoly's input choices, approving "necessary" charges and rejecting "rate-padding" expenses. That is, the regulatory authority cannot simply set the ceiling price; it must consistently police the decisions of the monopoly, trying to balance incompatible goals of maximizing consumer surplus and guaranteeing a "fair" return to the owners of the business.

An alternative to public regulation of natural monopolies is public ownership and operation. In the United States, public ownership of natural monopolies usually involves local governments – such as the Southern Nevada Water District or Bay-Area Rapid Transit. Examples of federal government ownership of natural monopolies include the Postal Service, the Tennessee Valley Authority, and Hoover Dam. Figure 9-7 depicts the options available for a municipally owned utility, say a local power company. As in Figure 9-6, the market demand curve intersects the long-run average cost curve in a range where the latter is still decreasing. The government has three options in Figure 9-7: maximize economic profit, maximize consumer surplus, or operate the "business" on a break-even basis.

[5] Recall that economic costs include a "normal return" on investments.

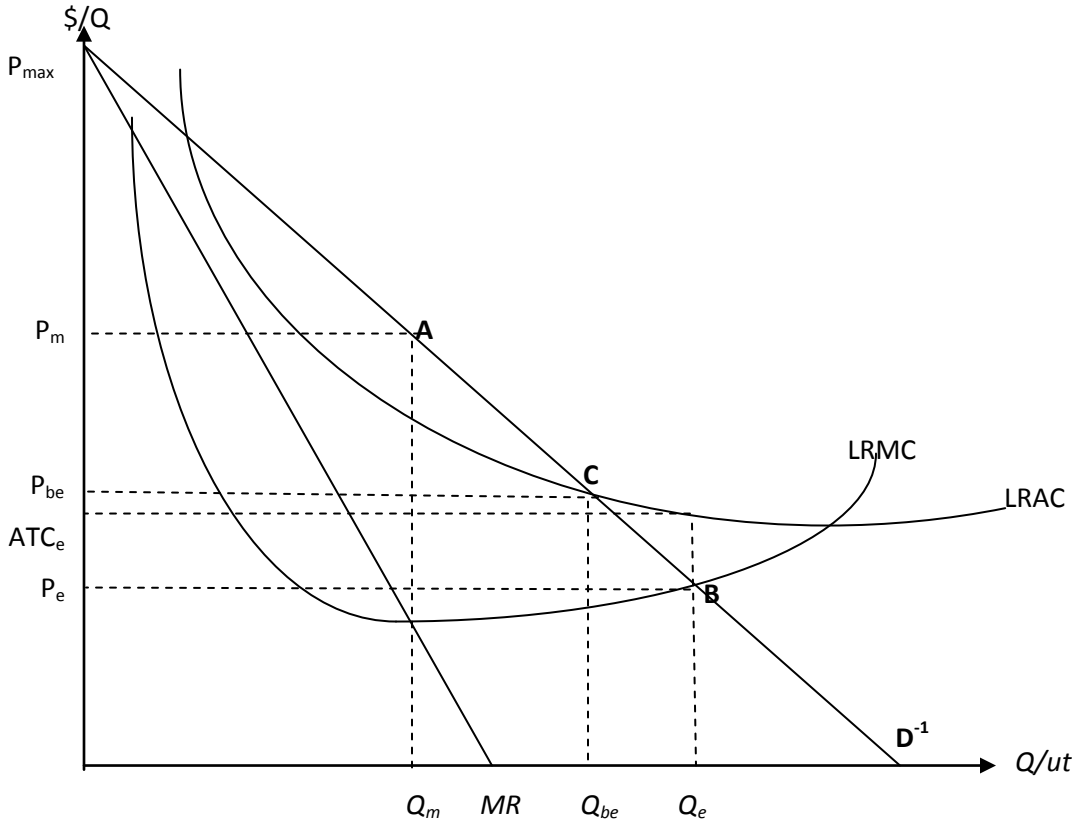


Figure 9-7: Options for Public Utility

It might seem odd for a government to operate a natural monopoly for profit, but that is precisely the behavior of the countries which participate in the OPEC cartel. Maximizing economic profit means that the government is able to cover its economic costs and generate a surplus which can be used to supplement general revenue. It should be obvious that the loss of consumer surplus, which falls from $½(p_{max} - p_{be})Q_{be}$ under break-even pricing to $½(P_{max}-P_m)Q_m$ generates the typical excess burden associated with monopoly pricing.

If governments wished to maximize consumer surplus, the appropriate price to set would be p_e, generating a net consumer surplus of $½(P_{max} - P_e)Q_e - (ATC_e - P_e)Q_e$. Most likely, government decision-makers would attempt to set a break-even price, initially at ATC_e. However, the increase in price would reduce quantity demanded, and the break-even price would eventually be set at P_{be}, resulting in a smaller net consumer surplus than setting price at P_e and subsidizing the difference between the efficient price and average cost out of general tax revenues.

Figure 9-8 illustrates folly of "break-even pricing" for a natural monopoly. The United States Constitution empowers congress to establish post offices and post roads. Until the 1970's, the post office functioned as a government entity, with the *Postmaster General,* starting with Benjamin Franklin, sitting on the President's Cabinet. As alternative means of communication evolved, the demand for postal services declined, so that, but the 1970's, the post office confronted the situation depicted in Figure 9-8; the price Americans were willing to pay for postal services was less than the average cost of providing that service.

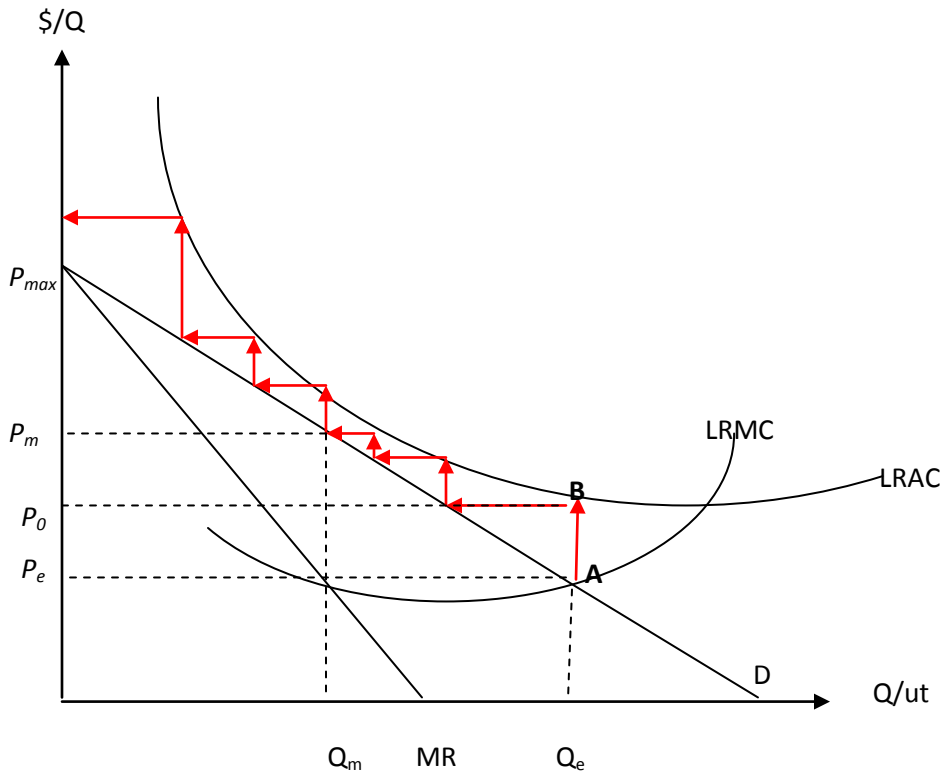


Figure 9-8: How to Kill the Postal Service

If the postal service priced stamps at P_e, postal patrons would mail Q_e letters and receive a surplus given by triangle $P_{max}P_eA$ in figure 9-8. The consumer surplus would be partly offset by the cost of covering average costs, equal to rectangle $P_eATC(Q_e)AB$, which still leaves a large net consumer surplus. In 1970, however, the US Postal Service was established as a semi-autonomous government agency charged with covering the total costs of providing postal services out of revenue. Despite well-publicized efficiency campaigns that increased the productivity of postal workers (unfortunately, the impact of which gave us the term "going postal"), the average cost curve everywhere lies above the demand curve. Hence, raising postal rates has become a nearly annual event, with the predictable consequence that each increase in postal rates reduces the volume of mail, further increasing the average cost. The obvious consequence is that the postal service will eventually price itself out of the market, since even the monopoly price of P_m does not guarantee survival.

Since the government monopoly is clearly not working as the Founding Fathers planned, why not give the private sector a chance? After all, private sector managers presumably have stronger incentives to reduce costs and increase revenue than government bureaucrats do. But Figure 9-8 implies that a monopolist, whether operating in the public sector or the private sector, would be unable to cover costs. Those who argue for privatization of failing government enterprises miss the central point: monopoly pricing generates excess burden; the most efficient allocation of resources occurs with marginal cost pricing. If the competitive market can allow firms to operate privately, leaving the activity in the private sector will eventually maximize consumer surplus. However, if a natural monopoly can generate a consumer surplus by means of a small public subsidy, such an operation could indeed be appro-

priate as a public-private partnership. Another alternative, which may help monopolies cover their costs without inefficient pricing, is through **price discrimination.**

Price Discrimination

The ultimate monopoly power is the ability to charge different prices to different customers for the same commodities. Private universities typically charge rich but academically challenged college students tuition that exceeds the cost of their education in order to provide scholarships to bright but economically challenged students. Movie theaters charge children and senior citizens a lower price than they charge teenagers and younger adults for the same movie. Business travelers typically pay higher fares to fly business class than others traveling to the same destination pay when they fly standby. The first thing that the proprietor of a body shop asks a potential client before quoting a price for an auto repair is whether the bill will be paid by an insurance company. Textbook publishers often sell an international edition of their book for about 20% cost of the price they charge American students; often the only difference is a soft cover. One of the more controversial issues of our time is whether or not Americans should be allowed to purchase prescriptions drugs from Canadian pharmacies, since Canadians pay substantially less than Americans do for the same medication.

In each of the cases mentioned, attributing the price difference to a difference in cost is rather dubious. What all these cases have in common is that the group paying the higher price has a demand that is less price elastic than is the case for the group paying the lower price. Rich parents may treat their children's college education as a form of *conspicuous consumption*, for which the bragging rights of sending one's offspring to Harvard are much more impressive than sending one's child to UNLV. Baby sitters provide an excellent substitute good for dragging children to the movie theater, so parents are more sensitive to the price of children's tickets than they are to the prices they pay. If I bill my consulting clients $300 per hour, the opportunity cost of seeing a two-hour movie is $600; whether I pay $10 or $20 for a movie ticket represents a price difference of only 1.6%. However, if babysitters charge $8.25 per hour (the current Nevada minimum wage), whether children's tickets are priced at $5 or $6 is decisive in determining whether parents of two young children buy those tickets or not. If my auto insurance has a $500 deductible, whether I fix a small ding for $200 or $400 may be very important to my decision. However, if the repair is going to cost over $500, I am indifferent whether the insurance company shells out $1,000 or $10,000.

Price discrimination is the practice of charging different prices to different customers for what are essentially the same commodities. The cost of showing a movie to a child or an adult is the same – precisely zero, once the projector is running, as long as there are empty seats. The cost of repairing a car is the same, regardless of who pays for it. But when a restaurant charges a lower price for a luncheon menu, part of the reason for the reduced price is that people are more sensitive to lunch prices than they are to dinner prices, and partly because the portions are smaller. It is often the case that college students learn more from their peers than they do from their instructors; when elite universities give scholarships to bright students, are they discriminating against higher income students, or are they augmenting the faculty?

First-Degree Price Discrimination

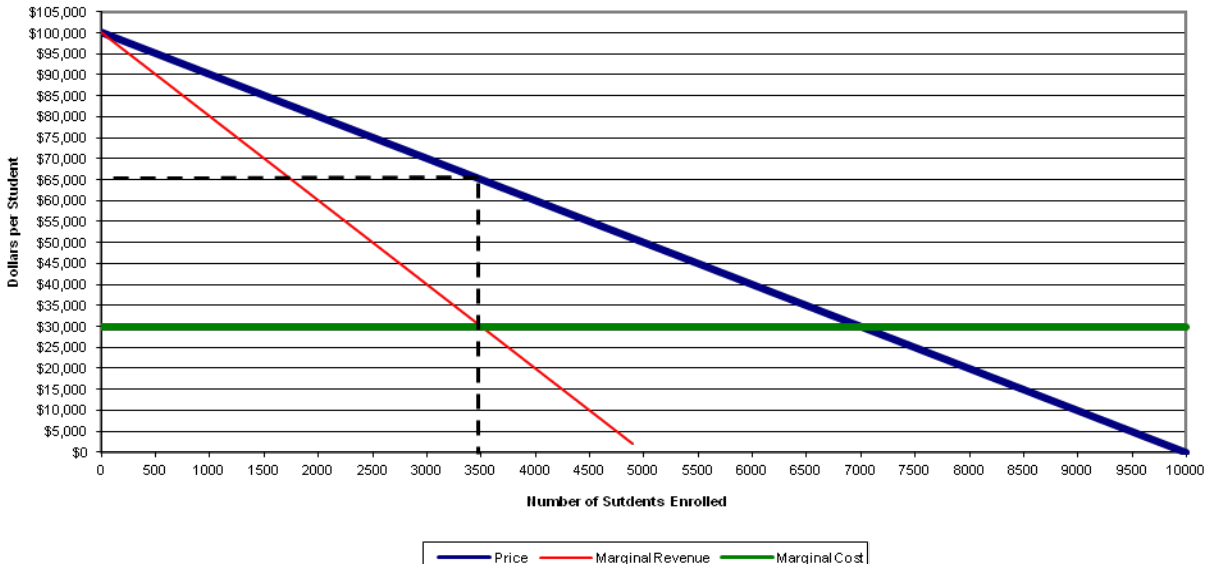
The case of differential college tuition comes closest to **first-degree price discrimination**, the practice of charging the maximum price consumers are willing to pay for each unit produced. When you applied for college, your parents completed a financial disclosure form that allows student-aid officers to determine your eligibility for student loans and need-based grants. You also took the either the SAT or the ACT (probably both) college aptitude tests which allow admissions officers to gauge your likelihood of success at their own, and other, universities. Although the information is not perfect, your fami-

ly's ability to pay and your likelihood to succeed determine the amount you or your parents would be willing to pay if you enrolled at that university.

Figure 9-9 shows the inverse demand for education at the University of Southern Nevada, a mythical privately-funded university nestled in picturesque Boulder City, Nevada. We suppose that USN has a dean of admissions who is able to glean the price each applicant's family is willing to pay to attend USN. Because education is a service, consumed at the moment of joint production between the students and their professors, it is not possible for students paying low tuition to resell the service to those who are charged higher tuition. The buying and reselling of a commodity for a financial gain is called **arbitrage**; in order for a seller to successfully discriminate among potential buyers in the prices they pay, arbitrage must be infeasible.

Suppose that the USN's dean of admissions learns that the demand for USN education is P = 100,000 – 10Q, where P is the total price USN charges students – books, tuition, fees, room and board – and Q is the number of students attending. For simplicity, we assume that it costs USN $30,000 to educate each and every student. If USN operated like a price taker, it would educate 7,000 students for $30,000 each; total revenue and total cost would both equal 7,000(30,000) = $210,000,000. Marginal-cost pricing would generate a consumer surplus of ½($100,000 – $30,000)x7,000 = $245,000,000; producer surplus would equal zero, since price equals average cost for each student.

Figure 9-9



If USN could not practice price discrimination – that is, if they had to charge each student an identical tuition, then the marginal revenue equation would be given by MR = $100,000 – 20$Q$; setting marginal revenue equal to marginal cost implies the profit-maximizing number of students:

$$\$100,000 - 20Q = \$30,000$$

$$20Q = \$70,000 \rightarrow Q_m = \frac{70,000}{20} = 3,500; P = 100,000 - 10(3,500) = \$65,000$$

As a non-discriminating monopolist, USN would educate only 3,500 students, charge tuition of $65,000, earn revenue of $227,500,000. With costs equal to $30,000 per student, profit without price discrimination would be:

$$\pi = \$227,500,000 - \$105,000,000 = \$122,500,000.$$

Monopoly would reduce the consumer surplus to ½(100,000-65,000)3,500 = $61,250,000. Added to USN's profit, this generates a total surplus of $122,500,000 + $61,250,000 = $183,750,000. The excess burden under monopoly would be the consumer surplus under perfect competition, $245,000,000 minus the total surplus under monopoly, $183,750,000, for a deadweight loss of $61,250,000. Recall that deadweight loss, or excess burden, occurs because the consumer surplus on the 3,500 students not admitted cannot be captured by the monopolist.

With price discrimination, USN's marginal revenue function becomes $MR = P = \$100,000 - 10Q$; setting this marginal revenue equal to marginal cost implies:

$$\$100,000 - 10Q = \$30,000 \rightarrow 10Q = 70,000 \rightarrow Q^* = 7,000$$

Revenue from the first student would be $100,000 – 10(1) = $99,990, revenue from the second student would be $100,000 – 10(2) = $99,980, and so forth, until USN broke even on their last student, whose tuition and fees would equal $100,000 – 10(7,000) = $30,000. This is equivalent to setting tuition equal to $100,000, and granting a $10 scholarship to the first student and a scholarship of $70,000 to the 7,000th student. Total revenue would be TR = ½(100,000 – 30,000)x (7,000) + 30,000(7,000) = $455,000,000. Total cost would be $30,000(7,000) = $210,000,000; profit (after scholarships) would be $245,000,000.

Second-Degree Price Discrimination

Second degree price discrimination typically involves charging different prices for different quantities. The best example of second degree price discrimination is **bulk pricing**. Table 9-4 shows hypothetical prices for dispensers of liquid hand soap. The smallest container, which contains six ounces, has the lowest price, but the highest price per ounce. The assumption behind second-degree price discrimination is that the quantity that people buy is a proxy for their price sensitivity. The quantity consumed is also a proxy for family size. Note that the opposite pattern is unworkable; charging higher per-unit prices for larger containers would merely encourage people to buy the multiple units of the smallest container. I remember when I went shopping for a standard 24-ounce bottle of mouthwash. I came home with six four-ounce bottles that had a lower price per ounce.

Table 9-4

Product size (ounces)	Price	Price per ounce
6	$1.71	$0.29
12	$3.24	$0.27
18	$4.59	$0.26
24	$5.76	$0.24
30	$6.75	$0.23
36	$7.56	$0.21

Third-Degree Price Discrimination

The most prevalent form of price discrimination is third-degree price discrimination, whereby monopolists charge different prices to different consumers which have identifiably different price elasticities of demand. Examples of third-degree price discrimination are: bookstore discounts to faculty, higher prices for insured patients than for uninsured patients, higher auto-body repair bills for insured motorists compared to prices paid by uninsured motorists, and movie-ticket discounts for senior-citizens and children. Table 9-5 presents hypothetical movie-ticket demand by adults and children. If the movie theater charged a uniform ticket price to all patrons, the best price to charge would be $6.00 per ticket, for which the *average* price elasticity is unitary. Uniform-pricing would generate revenue of $5,400. Since the marginal cost of serving another movie-goer is zero, maximizing revenue also maximizes profit.

Table 9-5: Movie Theater Example

Price	Qa	paQa	adult elasticity	Qc	pcQc	children's elasticity	Total Revenue	Average elasticity
$18.00	0	$0		0	$0		$0	
$17.00	50	$850	-17.00	0	$0		$850	-17.00
$16.00	100	$1,600	-8.00	0	$0		$1,600	-8.00
$15.00	150	$2,250	-5.00	0	$0		$2,250	-5.00
$14.00	200	$2,800	-3.50	0	$0		$2,800	-3.50
$13.00	250	$3,250	-2.60	0	$0		$3,250	-2.60
$12.00	300	$3,600	-2.00	0	$0		$3,600	-2.00
$11.00	350	$3,850	-1.57	0	$0		$3,850	-1.57
$10.00	400	$4,000	-1.25	0	$0		$4,000	-1.25
$9.00	*450*	*$4,050*	*-1.00*	0	$0		$4,050	-3.00
$8.00	500	$4,000	-0.80	100	$800	-8.00	$4,800	-2.00
$7.00	550	$3,850	-0.64	200	$1,400	-3.50	$5,250	-1.40
$6.00	600	$3,600	-0.50	300	$1,800	-2.00	$5,400	-1.00
$5.00	650	$3,250	-0.38	400	$2,000	-1.25	$5,250	-0.71
$4.50	675	$3,038	-0.33	450	$2,025	-1.00	$5,063	-0.60
$4.00	700	$2,800	-0.29	500	$2,000	-0.80	$4,800	-0.50
$3.00	750	$2,250	-0.20	600	$1,800	-0.50	$4,050	-0.33
$2.00	800	$1,600	-0.13	700	$1,400	-0.29	$3,000	-0.20
$1.00	850	$850	-0.06	800	$800	-0.13	$1,650	-0.09
$0.00	900	$0	0.00	900	$0	0.00	$0	0.00

Without price discrimination, the seller faces a dilemma. At the $6.00 price, the price elasticity of demand for adults is -0.50, so that adult revenue could be increased by raising the price. However, when p = $6.00, the children's price elasticity is -2.00; a decrease in price would raise the revenue from children. If the monopolist could charge different prices for adults and children, it would have to prevent arbitrage. This would require a ticket-seller with minimal depth perception who could distinguish between adults and children. With price discrimination, the adult ticket price would be set at $9.00, increasing adult ticket revenue from $3,600 to $4,050. On the other hand, reducing the children's price to $4.50 would increase children's revenue from $1,800 to $2,025. Price discrimination increases revenue (and hence profit) by $675.

Summary

1. A monopoly is a market in which there is only one seller. For practical purposes, this means that a monopolist can pick the most advantageous price-quantity combination from the market demand curve for the monopoly product and that there will be no retaliation by rivals if the monopolist changes price.

2. A monopolist considers buyer response to price change, so that marginal revenue is less than price by the implied price cut to buyers who would have paid a higher price. Marginal revenue equals the derivative of price times quantity with respect to quantity: $MR = \frac{\Delta(p \times q)}{\Delta q} = p\left(\frac{\Delta q}{\Delta q}\right) + q\left(\frac{\Delta p}{\Delta q}\right) = p\left(1 + \frac{1}{E_{dp}}\right)$, where E_{dp} is the price-elasticity of demand.
3. Like the competitive firm, the monopolist sets output where marginal cost equals marginal revenue. However, since marginal revenue is less than price, marginal cost is less than price.
4. Because monopolists increase price by reducing output, the monopoly gain in producer surplus is less than the loss of consumer surplus. The excess burden of monopoly equals the sum of consumer and producer surplus on the output that would have been produced under competitive market conditions.
5. As long as monopoly power can be transferred from one owner to another, the market value of the monopoly power is equal to the discounted present value of monopoly profits, meaning that the second-generation monopolist receives a normal return on their "investment."
6. A natural monopoly occurs when would-be market entrants perceive economic loss after entry; this implies that the residual demand curve confronting the would-be entrant is below its long-run average cost curve.
7. When a regulatory authority sets the ceiling price for a monopoly, the monopolist's marginal revenue is equal to the ceiling price until the market-quantity demanded at that price is produced. This can increase the quantity the natural monopolist produces, thereby mitigating the excess burden of the monopolist. However, price controls create incentives for regulated monopolies to allocate resources inefficiently to obtain rate increases.
8. The ideal price under monopoly is the same as under competition – the price associated with the intersection of the long-run marginal cost curve with the market demand curve. Attempting to force government operated monopolies – like the postal service – to cover all costs out of revenue could destroy an opportunity for a large consumer surplus.
9. Price discrimination is the practice of charging different prices to different customers based on price-elasticity of demand.
    a. First degree price discrimination charges the highest price for each unit produced, causing (marginal) price to equal marginal revenue, producing the same output as under perfect competition, but converting all consumer surplus into producer surplus.
    b. Second degree price discrimination charges prices according to the quantity purchased; the best example is charging less per ounce for "large, economy sized" packages.
    c. Third degree price discrimination is the most prevalent form of price discrimination, which involves setting different prices for different groups. Examples include lower air fare for tourists, bookstore discounts to faculty members, higher prices to the insured (who are relatively insensitive to price because the insurance company bears the burden), and senior-citizen or children's discounts.

Glossary

Arbitrage: The practice of buying a commodity at a low price and reselling it at a higher price; arbitrage interferes with a seller's ability to price discriminate.

Capitalization: The process by which the market price of a company or a franchise captures the present expected value of future monopoly profit, meaning that second and subsequent generation monopolists, along with shareholders of monopoly business receive a normal return on financial "investments."

Cartel: A monopoly composed of multiple businesses or individuals that conspire to charge a monopoly price. Because the cartel member can (at least temporarily) increase profit by undercutting the cartel price, cartels are inherently unstable.

Monopoly: A market with one seller.

Marginal revenue: The change in revenue due to selling one more unit, $MR = \frac{\Delta(p \times q)}{\Delta q} = p\left(1 + \frac{1}{E_{dp}}\right)$.

Under monopoly, $\frac{\Delta p}{\Delta Q} < 0$, so that $MR < p$.

Natural monopoly: A market in which one firm persists due to protracted economies of scale, relative to market demand, confronting would-be entrants with the specter of after-entry losses.

Price discrimination: The practice of charging different prices to consumers with different elasticities of demand for a technically homogenous commodity.

Evaluate

Indicate whether each of the following statements is true (agrees with economic theory), false (contradicts economic theory), or uncertain (could be true of false, given additional information) and *explain your answer.*

1. If a monopolist has a positively sloped marginal cost curve at the profit-maximizing rate of output, a binding price ceiling below the monopoly price would increase the firm's output but decrease its profit.

2. If a cartel wishes to maximize the total income of members it should allocate production quotas so that all producers have the same average (total) cost.

3. A monopoly movie theater discovers that the demand for snacks while the movie is playing is in the price-inelastic region of a linear demand curve. The movie theater could increase profits by reducing the price of movie tickets below the unit-elastic point on the ticket demand curve.

4. An increase in the property tax imposed on a monopolist would increase price and reduce quantity produced, *ceteris paribus.*

5. If the elasticity of demand for a monopolist's product is -3, and if the monopolist's price is three times marginal cost, the monopolist is maximizing profit.

6. As long as monopolists are free to sell their firms to the highest bidder, monopoly profits can persist in the long run.

7. An increase in demand for a monopolist's product will *always* increase the price and quantity produced.

8. Allowing monopolies to practice price discrimination enhances economic efficiency.

9. Arbitrage and price discrimination are incompatible.

10. Defining monopoly as a producer of a product that has no substitutes implies that monopolies never maximize profit, or that monopolies do not exist.

Calculate

11. A monopoly sells its output in two markets from a factory located on the boundary between market 1 and market 2. An econometrician has estimated that the inverse demand equation for market 1 is $p_1 = \$100 - 2Q_1$ and the inverse demand function in market 2 is $p_2 = \$80 - Q_2$. The monopolist's total profit function is $C = \$1{,}000 + 10Q$, regardless of whether the product is sold in market 1 or market 2.

    a. Write the formula for the marginal revenue in each market.

    b. What is the profit-maximizing price in each market, assuming that arbitrage is impossible?

    c. Suppose that the monopolist's product can be shipped between markets at a cost of \$5 per unit. How does this result change your answer to part *b* of this question?

12. A natural monopolist confronts an inverse demand curve given by $p = 100 - Q$ and has a constant marginal cost of \$10 and a fixed cost of \$1000.

    a. What is the profit-maximizing output, price, and profit for the unregulated monopolist? What are the resulting consumer and producer surpluses?

    b. If the monopolist could practice first-degree price discrimination, what would be the output, price range, producer surplus and consumer surplus?

    c. What would be the economically efficient price for the product? How much of a subsidy would be required and what would the consumer surplus be, net of the subsidy?

    d. If the price was set at a level that would allow the monopolist to receive a normal return on investment, what would be the price, the output, and the consumer and producer surpluses that would result?

Short Essay

13. A book publisher sells two books, one an advanced treatise on mathematical economics sold to Ph. D. students in economics at a price of \$200, and the other a principles of economics text that sells for \$100. Each book's marginal cost is \$20. What do these relative prices tell you about the price elasticity of demand for each book? Are these relative price elasticities reasonable? Explain.

14. Back in the era of cheap foreign oil, the price of oil produced from wells in the United States was twice the world price. This price was supported by a law that limited oil imports to only 10% of oil consumed in the USA.

    a. Did knowing that the oil import restrictions could be repealed at any time encourage oil conservation in the USA? Explain.

    b. Do you think that the USA's dependence on foreign oil would be greater or less had those import restrictions never been in place?

    c. What does this imply about foreign trade and arbitrage?

15. "Without government there could be no monopoly?" Do you agree or disagree? Explain your reasoning.