Monopoly
Readings: Ch. 12

Now that we have an understanding of supply when there are lots of competitors producing the same thing, let us consider something different. In particular, let us consider what happens when there is only one firm that produces all output.

Objective: Understand what monopoly is and how it differs from perfect competition.

When a single firm produces all of the output in an industry and there are no close substitutes for that output it is called a monopoly.

Definition
Monopoly: An industry structure in which a single seller of a product with no close substitutes serves the entire market.

As with perfect competition, it is difficult to think of perfect examples of monopoly. One reason is that the U.S. tends to discourage monopolies because as we will see later they will generally not produce an efficient level of output. Another reason is that monopolies can generate profits in the long run. Therefore, there are strong incentives for others to try to imitate the monopolist's product by developing close substitutes. If successful, the monopoly industry will be transformed to some other more competitive industrial organization.

In a perfectly competitive market, all firms take price as given. Another way of saying the same thing is that the demand curve they face is perfectly elastic. Trying to increase price results in a loss of all sales, so perfect competitors can only use their quantity supplied to respond to changes in market demand.

This is not the case for a monopoly. A monopoly faces a downward sloping demand curve and has significant control over the price of the product as well as the level of output. An important implication is that a monopoly has no supply curve. This is an important fact, so I will state it again with emphasis:

A MONOPOLY HAS NO SUPPLY CURVE!

Table 1 summarizes the important differences between a perfectly competitive and monopoly industry.

Table 1

Perfect Competition	Monopoly
Many Sellers	One Seller
Many Perfect Substitutes	No Close Substitutes
Price Determined By Market Not Individual Sellers (Face Perfectly Elastic Demand)	Price Determined By The Only Seller (Face Downward Sloping Demand)

Objective: Understand reasons why monopolies exist.

There are several reasons why there might be only one seller of a product with no close substitutes.

1. Exclusive Control Over Inputs: When a product requires a special input in order to produce it (e.g. diamond rings require diamonds), then control over that input will allow a seller to control the output (e.g. DeBeers controls most of the worlds diamond mines, so DeBeers can control the sale of diamonds for jewelry and other uses).
2. Economies of Scale (Most Important): When we talked about costs, we talked about a case where the long run average cost was always decreasing. When this is the case, the larger the firm the cheaper it is to produce. One firm can supply the whole market at the lowest possible cost. This encourages the formation of monopolies. A common example is an electric utility.
3. Patents: In the U.S. and around the world, governments issue patents for new and unique products, such as cancer drugs. These patents give a seller exclusive rights to produce, sell, or license a product for a specific period of time (typically 14-20 years in the U.S.). The reason governments do this is to provide incentives for sellers to invest money into developing new products that could benefit society. The idea is that giving a 20-year monopoly to a seller provides them the opportunity to earn a fair return on research and development costs. A return they could not earn if everyone could produce and sell the product as soon as it was invented.
4. Network Economies: Products may be more valuable when they are more widely used. For example, if everyone used the same computer operating system it would be easier for everyone to share work with others. Having the majority of the world using Microsoft Windows makes it easier for everyone to exchange and share information and makes it so software developers can serve a broader market with a single product.
5. Government Licenses Or Franchises: In some instances, the government will provide an exclusive license to sell a product. More often than not these licenses are provided in response to economies of scale or a decreasing long run average cost and serve as a means for the government to control or influence monopoly behavior to the benefit of society.

Objective: Understand how price and quantity are determined in a monopoly industry.

As with perfect competition, we will assume a monopolist seeks to maximize economic profit, which if you recall is $\Pi = TR - TC$. This all looks very similar to perfect competition, but there is an important difference. The difference is that if the monopolist wants to increase the amount of output it sells, it must decrease the price it charges. Total costs can be calculated the same as we did for perfect competitors.

Recall that if demand is linear total revenue is just a hill (see Figure 1). We can now add the monopolies fixed, variable, and total cost curves to this diagram (see Figure 2).

Question: How much should the monopolist charge and how much should it produce to maximize its profits in the short run?

To answer this question we want an output where TR > TC and the difference is as big as it gets. Looking at Figure 2 suggests this is quite the challenging task. So, here is another place in economics where calculus is terribly helpful.

Recall that if we want to maximize a function we can take the first derivative and set it equal to 0: $\Pi' = TR' - TC' = 0$. Rearranging terms tells us that the profit is maximized where TR' = TC', but what are TR' and TC'? They are just marginal revenue and marginal costs. This is the same as what we said for perfect competitors, so how are things now different?

They are different because when a perfect competitor changes its output the price doesn't change. When a monopoly changes its output, the price does change. For a perfect competitor, $TR = P_0Q$, while for a monopoly, TR = P(Q)Q. Therefore, marginal revenue for a perfect competitor is $MR = P_0$, while for a monopoly it is MR = P'(Q)Q + P(Q).

One thing all this tells us is that to find to profit maximizing monopoly output it is useful to look at the marginal revenue and cost curves rather than the total revenue and cost curves.

Question: What does the marginal revenue curve look like when demand is linear?

Consider the demand in Figure 1: Q = a – bP, which also implies $P(Q) = \frac{a}{b} - \frac{1}{b}Q$. Total revenue is then $TR = P(Q)Q = \left(\frac{a}{b} - \frac{1}{b}Q\right)Q = \frac{a}{b}Q - \frac{1}{b}Q^2$. Taking the derivative we get $MR = \frac{a}{b} - 2\frac{1}{b}Q$, which is the equation of a line. When Q = 0, MR = a/b, which is the same for the demand curve when Q = 0. When MR = 0, Q = a/2, which is ½ the quantity where demand intercepts the horizontal axis. Figure 3 shows this relationship.

Now to find the output that maximizes profit all we need to do is add marginal costs to this figure (see Figure 4) and look for the point where marginal revenue and marginal cost intersect (Q*).

Now that we know how much to produce, to find how much to charge we need to figure out how much people are willing to pay for this quantity. Using Q* and the demand curve, we see this price is P*.

Let us consider a specific example using the demand and costs from our monopoly experiment. While I did not tell you the demand curve most of you seemed to figure out it was Q = 20 – 2P. I also told you that the marginal costs was MC = 2.

To find the profit maximizing monopoly price and quantity, we need to set marginal revenue equal to marginal cost. Marginal cost is easy: MC = 2. Marginal revenue is more challenging, but we know it is the derivative of total revenue. Total revenue is TR = P(Q)Q where Q = 20 – 2P. The first problem we face is that P(Q) is the price in terms of Q, but Q = 20 – 2P is the

quantity in terms of the price. To get the price in terms of Q, we need to rearrange $Q = 20 - 2P \Rightarrow Q + 2P = 20 \Rightarrow 2P = 20 - Q \Rightarrow P = 10 - 0.5Q$. **Remember We Always Want To Make Sure The Demand Curve Is Solved For Price in Terms Of Quantity When Solving The Monopoly Problem!** Often I will do this for you, but I make no promises. Now that we have P in terms of Q we can substitute: $TR = (10 - 0.5Q)Q = 10Q - 0.5Q^2$. Taking the derivative we get marginal revenue: $MR = TR' = 10 - 2(0.5)Q^{2-1} = 10 - Q$.

Now we can set marginal revenue equal to marginal cost to find the monopoly's profit maximizing output: $MC = MR \Rightarrow 2 = 10 - Q^* \Rightarrow Q^* = 8$. To find the price the monopoly should charge, we need to substitute this quantity into the demand function: $P^* = 10 - 0.5Q^* = 10 - 0.5(8) = 6$. Therefore, a profit-maximizing monopolist should produce 8 and charge \$6. Figure 5 illustrates the solution.

Now we haven't been too careful here in that we have not made sure that we actually found a maximum and not a minimum. Taking the second derivative of the profit function and making sure it is negative will help here: $\Pi'' = TR'' - TC'' = MR' - MC' < 0$ or $MR' < MC'$. That is, the slope of the marginal revenue curve must be smaller than the slope of the marginal cost curve.

Another thing we have not been careful about is whether or not it is in the interest of the monopoly to produce anything at all. As with perfect competition, a monopolist must cover its variable cost in the short run and total costs in the long run. That is, $P > AVC$ in the short run and $P > ATC$ in the long run.

Figure 6 shows an example of a monopolist that will produce in the long run and short run. Monopoly profits in this example are equal to the area denoted by **abcd**.

Figure 7 shows an example of a monopoly that produces in the short run, but will close down in the long run if it cannot reduce its costs. In this example, the monopoly losses the area denoted by **abcd**, but it is still optimal to produce because the price it charges exceeds its average variable cost.

Figure 8 shows an example where the monopolist cannot even cover its average variable costs and will therefore shut down in the short run. Again, monopoly losses in this figure are the area denoted by **abcd**.

To summarize:
Monopoly Short Run Profit Maximizing Conditions:

1. $MR = MC$
2. $MC' > MR'$
3. $P > AVC$

Monopoly Long Run Profit Maximizing Conditions:

1. $MR = LMC$
2. $LMC' > MR'$
3. $P > LAC$

Objective: Understand how the monopoly's profit maximizing rule differs from perfect competition and why a monopoly will not produce where demand is inelastic.

Recall that with perfect competition, firms set the marginal cost equal to marginal revenue which is equal to the price: MC = P. The monopolist also sets marginal cost equal to marginal revenue, but marginal revenue is $TR' = MR = P'(Q)Q + P(Q) = P(Q)\,(1 + P'(Q)Q/P(Q)) = P(Q)(1 - 1/|\eta|)$ where if you recall η is the price elasticity of demand.

But what does all this mean?

If demand is elastic $\eta < -1$ or $|\eta| > 1$, $P > P(Q)(1 - 1/|\eta|) = MC$. That is, a monopoly will produce where the price exceeds marginal cost. Remember perfect competitors produce where price equals marginal cost. The greater the elasticity of demand, the greater the difference in the price and marginal cost, and the greater the difference in the monopoly and perfectly competitive price. If demand is inelastic $0 > \eta > -1$ or $0 < |\eta| < 1$ so $MR = P(Q)(1 - 1/|\eta|) < 0$. That is, when demand is inelastic marginal revenue is negative. Since marginal cost is always positive, a monopoly will never produce where demand is inelastic.

To summarize the important results:

1. **The difference in the monopoly and perfectly competitive price will be greater when the elasticity of demand is greater.**
2. **A monopoly will never produce where the demand curve is inelastic.**

Objective: Understand why economists say monopolies are inefficient.

Monopolies produce where the price is above marginal cost, which means they produce less than what a perfect competitor produces if costs were the same. Figure 9 illustrates. P_M* and Q_M* are the profit maximizing price and quantity for a monopoly. P_C* and Q_C* are the equilibrium price and quantity for a perfectly competitive industry with the same marginal costs as the monopoly. Notice that the monopoly produces less and charges a higher price.

Earlier in the semester we used consumer, producer, and total surplus to evaluate the welfare implications of things like taxes, subsidies, and price floors. Here we can do the same. If the monopolist behaved as a perfect competitor, consumer surplus equals the area denoted by **acf**. Producer surplus equals the area denoted by **cef**. The total surplus equals the area denoted by **ace**. For the monopoly's profit maximizing price and output, consumer surplus equals the area denoted by **abg**. Producer surplus equals the area denoted by **bdeg**. The total surplus is denoted by the area **abde**. Compared to the perfectly competitive industry, producer surplus is higher, consumer surplus is lower, and total surplus is lower. Monopolies are considered inefficient because if they set price equal to marginal cost like perfect competitors do, total consumer and producer surplus would be higher by the area denoted by **bcd**.

Objective: Understand what price discrimination is and how it affect producer and consumer surplus.

Up to now, we have assumed that a monopolist will charge all buyers the same price, but this need not be the case. In some instances, a monopoly may be able to charge some buyers a higher price than others. When this is the case, we have what is called a price discriminating monopolist. As we will see, there are good things and bad things associated with price discrimination. The good thing is that price discriminating monopolists will tend to produce a more efficient level of output. The bad thing is consumer surplus will typically be lower, while producer surplus is higher, so the result may be considered less equitable.

Definition
Price Discrimination: The practice of charging different buyers different prices for the same product.

For price discrimination to exist there can be no arbitrage opportunities.

Definition
Arbitrage: The purchase of a product for costless and risk-free resale at a higher price.

Arbitrage opportunities exist when you can buy a product and then turn right back around and sell the product for a higher price. If this is possible and a monopoly charges different prices, those who pay less for the product can turn around and resell it in competition with the monopoly, which will drive the price in both markets to be equal.

Lets consider an example where a monopoly can sell its product in the U.S. and Europe. Figure 10 illustrates. D_U and D_E are the market demand for the monopoly's product in the U.S. and Europe. MR_U and MR_E are the corresponding marginal revenues. To determine where marginal cost equals marginal revenue, the monopoly needs to add the demands ($D_U + D_E$), so it can figure out the marginal revenue from its next sale ($MR_U + MR_E$). Setting marginal revenue equal to marginal cost tells the monopoly how much to produce (Q*), but how much should it charge?

If it charges the same price in the U.S. and Europe (P*), Q_U^* will be purchased in the U.S. and Q_E* will be purchased in Europe. But could the monopoly do better?

Notice that marginal revenue for Q_U* in the U.S. (MR_U*) is greater than for Q_E* in Europe (MR_E*). What does this mean? It means that the monopoly could increase its revenues by selling one more unit in the U.S. and one less unit in Europe. Since the total quantity would not change, the monopoly's cost would not change. More revenue and no more costs mean higher profits. Therefore, as long as $MR_U^* > MR_E^*$ the monopoly can increase its profit by selling more in the U.S. and less in Europe. Alternatively, if $MR_E^* > MR_U^*$, the monopoly could increase its profit by selling more in Europe and less in the U.S. The lesson is that the monopoly will earn its highest profit by choosing a different price in each market so $MR_U = MR_E$. To find this price, we need to first find the quantities where $MR_U = MR_E = MC$, which are Q_U** and Q_E**. How much Americans and Europeans are willing to pay for these quantities is P_U** and P_E**.

Now lets look at another and even more extreme form of price discrimination that is referred to as perfect or first-degree price discrimination.

Definition
Perfect (First-Degree) Price Discrimination: Price discrimination where the price a buyer pays equals its willingness to pay.

Remember that the demand curve tells us how much buyers are willing to pay for each additional unit of output. With perfect price discrimination, this willingness to pay is precisely what the monopoly charges. The net result is that marginal revenue equals demand, so a profit maximizing monopolist will end up setting P = MC just as a perfect competitor would (see Figure 11).

Figure 11 dramatically illustrates the good and the bad of price discrimination. Since the monopoly produces where price equals marginal cost, the total of producer and consumer surplus is maximized. This is the good. The bad is the distribution of producer and consumer surplus. For perfect competition, consumer surplus would equal the area denoted by **abd** and producer surplus would be equal to area **bcd**. Now what about with perfect price discrimination? Since the price the monopoly charges the buyer just equals the buyer's willingness to pay or the benefit of the sale to the buyer, there is no consumer surplus. But if there is no consumer surplus and total surplus is at a maximum, what is happening to all those surpluses? The answer is that it is all going to the monopoly. That is, producer surplus equals the area denoted by **abc**. So while perfect price discrimination is efficient, it is pretty easy to argue that it is not equitable.

The main result to remember about price discrimination is that it tends to lead to a more efficient allocation of resources (the good), but also leads to a more unequal distribution of the benefits from production (the bad).

Objective: Understand what has been done to address the inefficiency and inequities of monopoly industries.

Two concerns regarding monopoly are that they produce less than what is socially optimal (they are inefficient) and they allow the monopoly to capture a greater share of the value of production (they are inequitable). Governments have tried to address these concerns using a number of different policies.

In some instances, governments simply take over the monopoly industry. This is a common solution with natural monopolies, which if you recall have a decreasing long run average cost. When long run average cost is decreasing, long run marginal costs are always less than long run average costs: LAC < LMC. So if the government prices where P = LMC, it will loss money. Solutions to this problem include producing where P = LAC, so there is no loss or setting P = LMC while subsidizing losses with tax revenues.

Instead of owning monopolies outright, a government may choose just to regulate privately owned monopolies. This is also common for natural monopolies. As an example, the government may tell the monopolist it has to produce and sell where P = LAC. This can be a problematic policy because the government may not know what the monopoly's LAC looks like.

One way to reduce the market power of a monopoly is to increase the market power of buyers. What we have not talked about is when there is only one buyer. When there is only one buyer it is called a monopsony. Like monopolies, monopsonies can exercise power in the market, so with a monopoly seller and monopsony buyer you can actually get a more efficient result. This has led some governments to take the responsibility of representing all consumers. By negotiating with the monopoly on behalf of all buyers, the government can exercise monopsony power in an attempt to secure a more efficient and equitable result.

Governments have also used laws that simply prohibit monopoly behavior (Sherman Act 1890; Clayton Act 1914). Firms that violate this anti-trust regulation can face severe penalties.

Some governments take a Laissez-Faire Policy approach, which means they do nothing. The argument is that monopoly profits cannot be sustained indefinitely because the market will always find a way to eliminate them through competition.

Figure 1: Total Revenue When Demand Is Linear

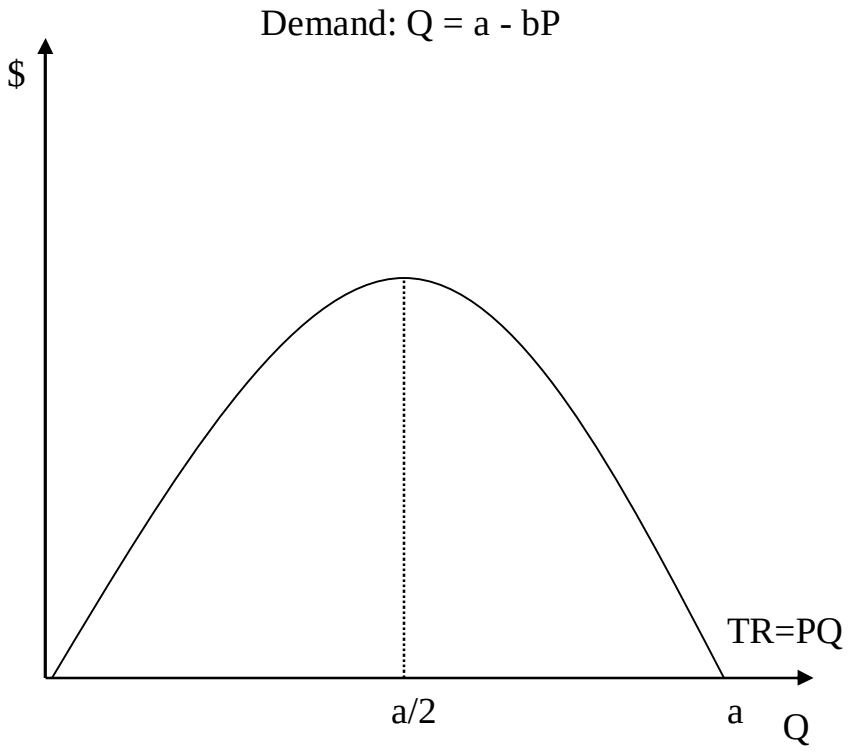


Figure 2: Total Revenue, Variable Cost, & Total Cost

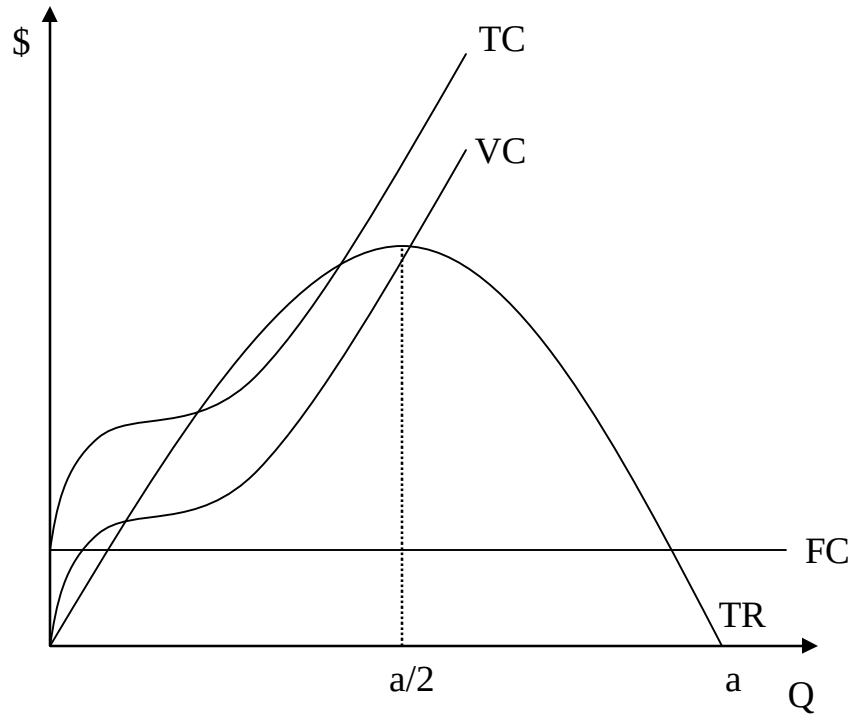


Figure 3: Linear Demand & Marginal Revenue

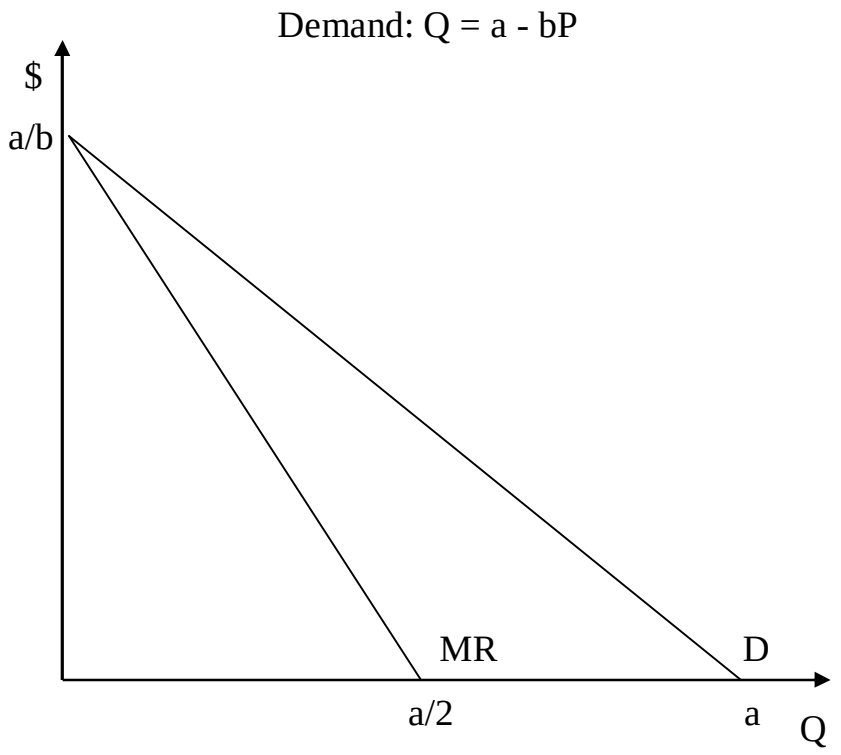


Figure 4: Marginal Revenue & Marginal Cost

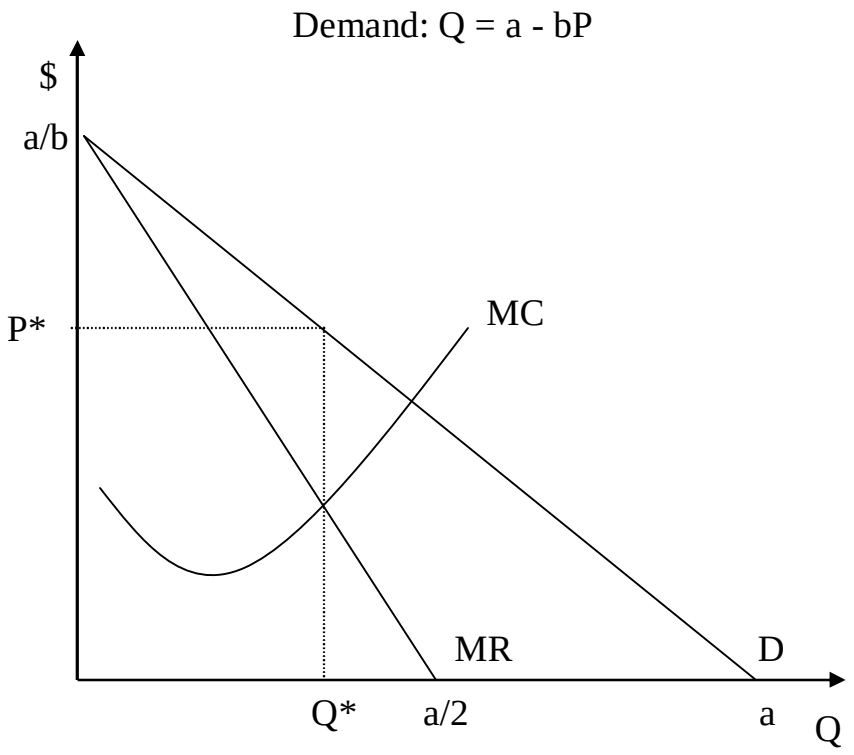


Figure 5: Monopoly Solution When Q = 20 – 2P and MC = 2

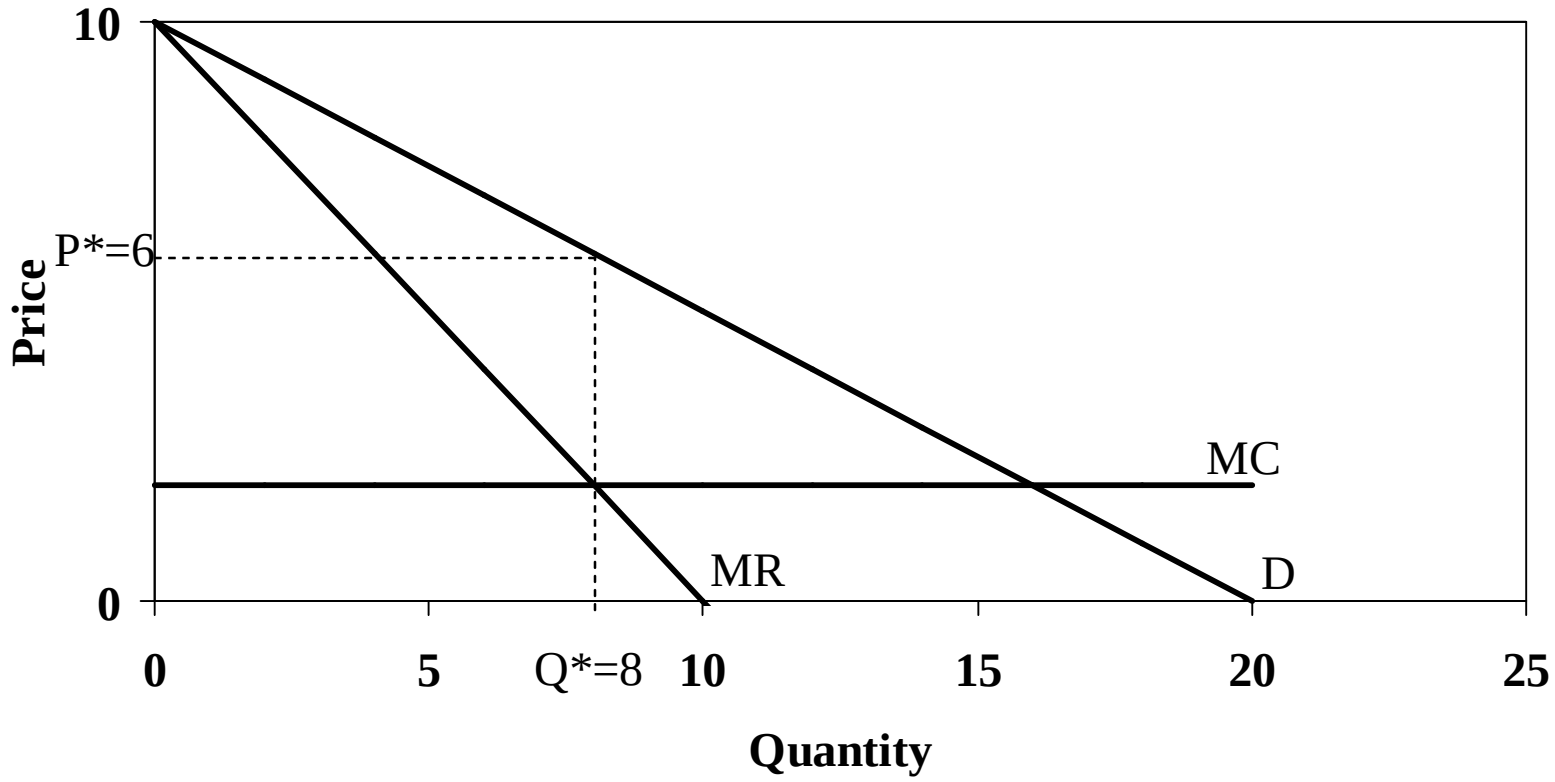


Figure 6: Example of Monopoly Profits In The Short Run

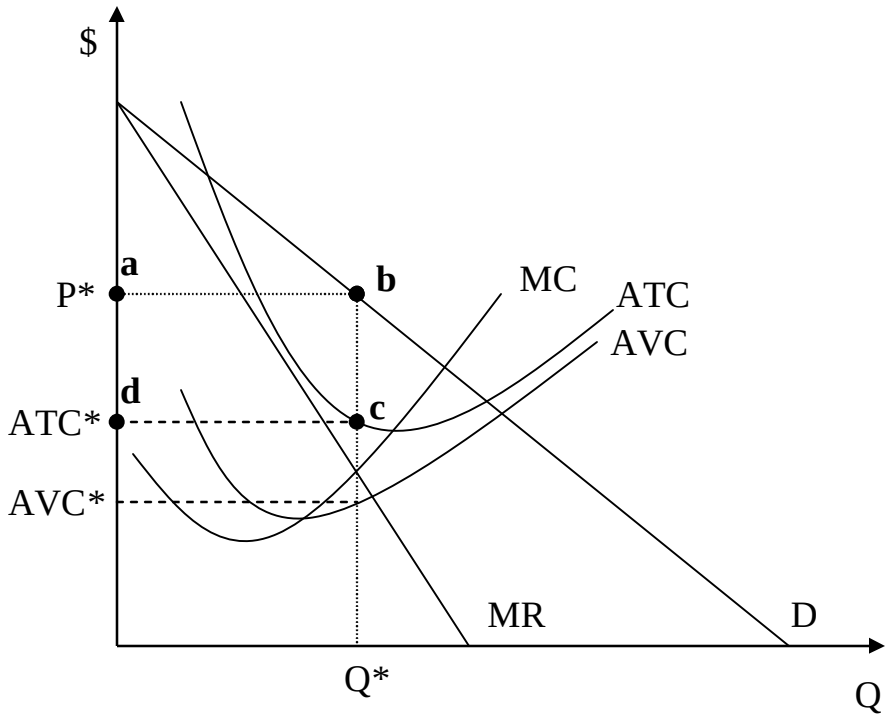


Figure 7: Example of Monopoly Losses In The Short Run

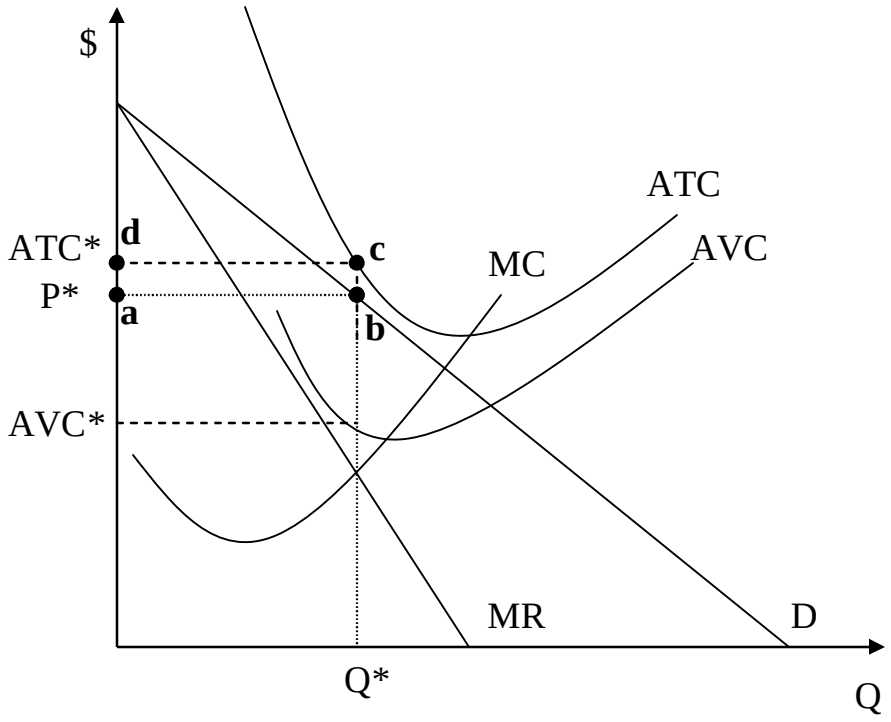


Figure 8: Example of When a Monopoly Should Shut Down in the Short Run

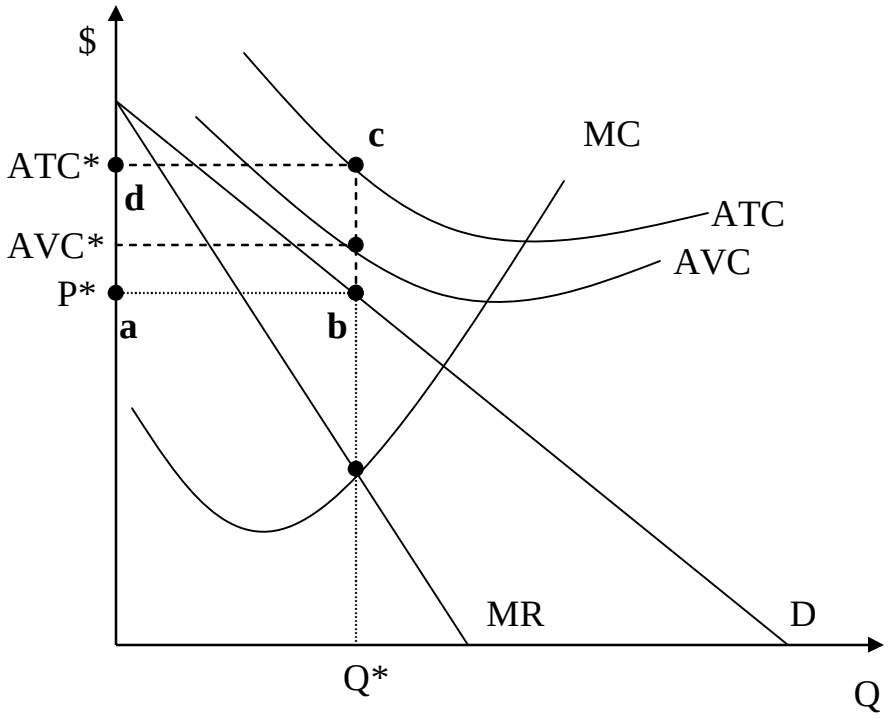


Figure 9: Inefficiency of Monopoly

Consumer (CS)
Producer (PS), and
Total (TC) Surplus

Monopoly
CS = area **abg**
PS = area **bdeg**
TS = area **abde**

Perfect Competitor
CS = area **acf**
PS = area **cef**
TS = area **ace**

Surplus Lost Due To
Monopoly = area **bcd**

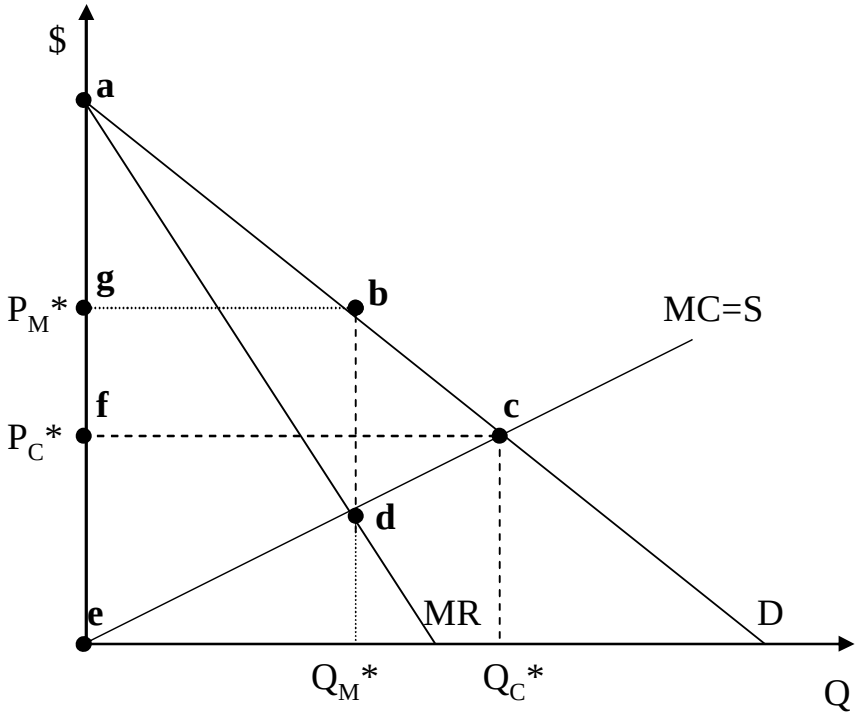
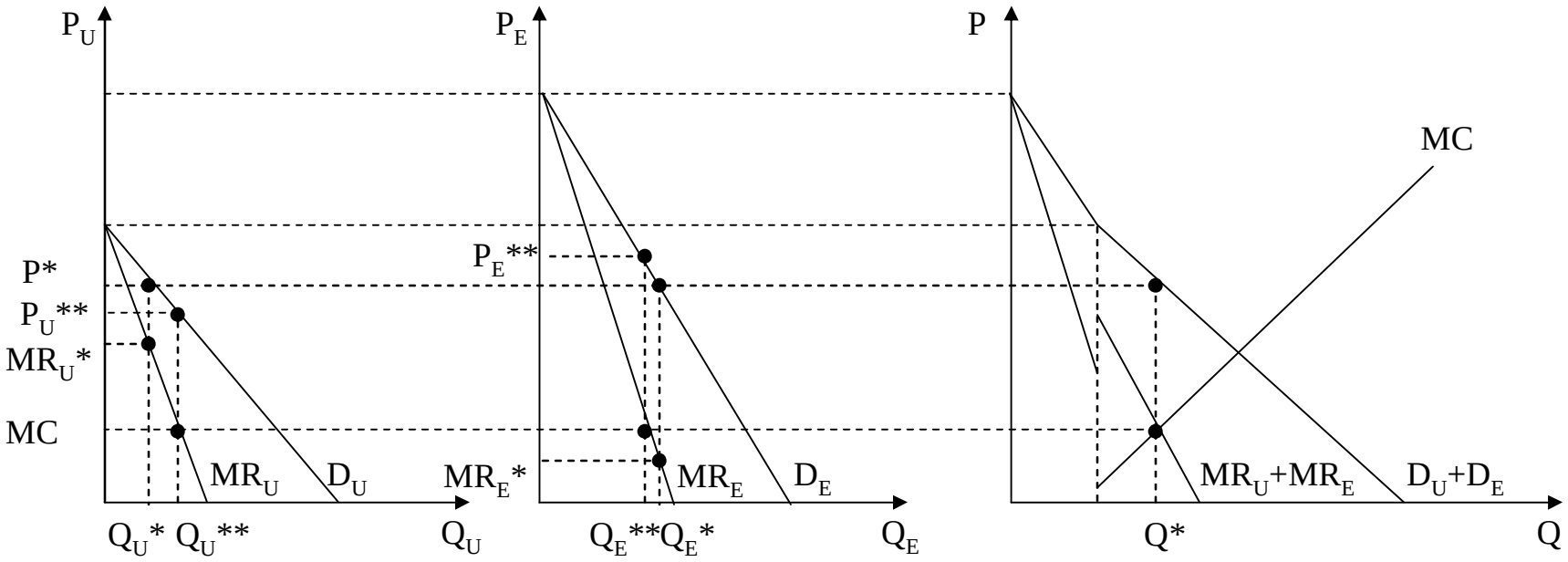


Figure 10: Price Discrimination Between U.S. (U) and European (E) Markets



Note: $Q_U^* + Q_E^* = Q^*$

$Q_U^{**} + Q_E^{**} = Q^*$

Figure 11: Perfect (First-Degree) Price Discrimination, Efficiency, and Equity

Consumer (CS)
Producer (PS), and
Total (TC) Surplus

Monopoly
CS = 0
PS = area **abc**
TS = area **abc**

Perfect Competitor
CS = area **abd**
PS = area **bcd**
TS = area **abc**

Surplus Lost Due To
Monopoly = 0