Chapter 12: Monopolistic Competition and Oligopoly

Most markets lie somewhere in between perfect competition and monopoly. In some markets, there are many firms, but firms do not sell exactly the same product. In other markets there are only a handful of firms, so that each has market power but must consider the behavior of its rivals when setting prices and quantities. In this chapter, we will examine some of the possibilities that arise in these situations, and how they differ from our "extreme" cases of perfect competition and monopoly.

Monopolistic Competition

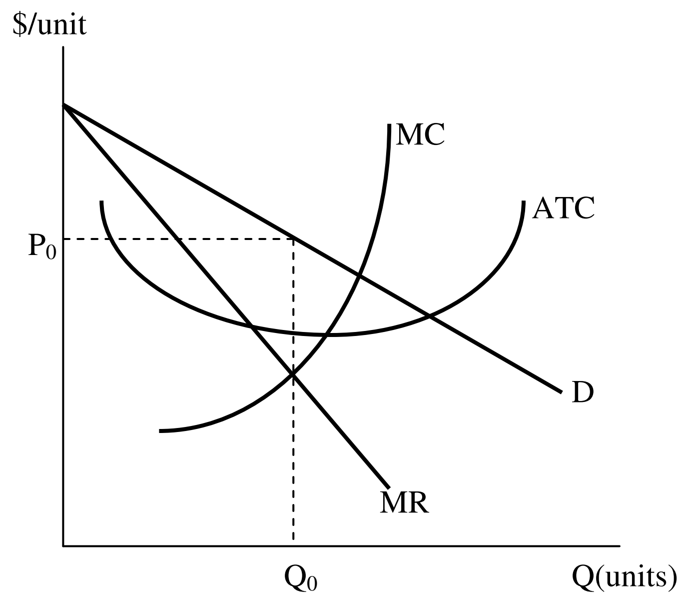
Until now, we have assumed that products produced by firms are homogenous. We will now relax that assumption and allow a products in a market to vary on different dimensions. Products are said to be differentiated when consumers perceive a particular firm's product as being only an imperfect substitute for those sold by its rivals. Differentiated products are quite common.

The basic idea behind product differentiation is that it gives firms market power. If Coke and Pepsi were 100% identical, nobody would pay more for Coke than Pepsi. By differentiating their products, Coke knows that some consumers will pay more for Coke than for Pepsi, so Coke can raise its price without having all its sales drop to zero. Likewise, Pepsi knows that some consumers will be willing to pay more for its product than Coke's, so Pepsi can raise its price above MC as well.

The most common way to discuss product differentiation is through the model of monopolistic competition, which involves two main assumptions:

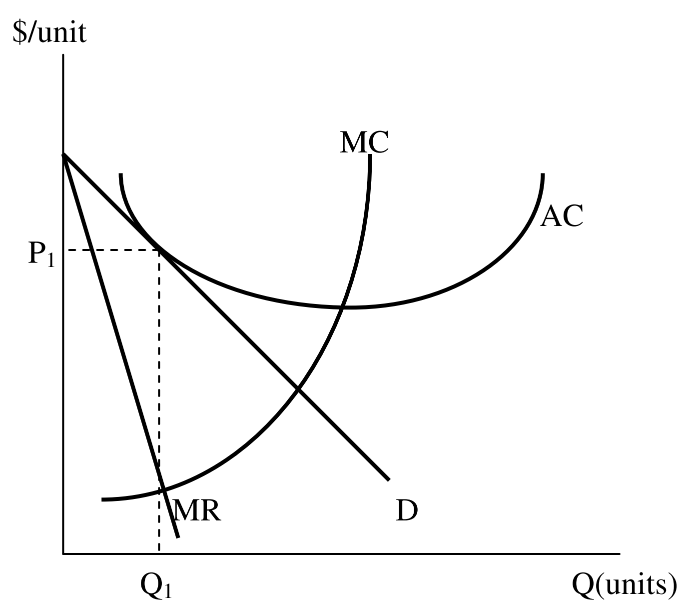
1) Products are heterogeneous. That is, each firm faces a downward-sloping demand curve.
2) Freedom of entry and exit

Being a profit-maximizer, each firm sets MC = MR as usual, resulting in the following familiar diagram:



Note that this firm is earning positive economic profits because P > ATC at the profit-maximizing quantity.

Therefore, this diagram cannot represent a long run equilibrium, because these profits will attract new entrants into the market. The entry of new, rival firms will cause the demand for the incumbent firm's product to fall. This entry will occur as long as the firm continues to earn a profit, so D (and MR) will continue to fall until P = ATC at the quantity where MR = MC. This will eventually result in the following situation:



At $Q_1 < Q_0$, $P_1 = ATC$ so the firm is now making zero economic profits.
The long-run result of monopolistic competition is somewhat like that of perfect competition in the sense that profits are driven to zero. However, notice that monopolistic competition still results in inefficiently low output, because each firm produces at a quantity where P > MC. This means that at least some consumers would like the firm to produce more, and they would be willing to pay enough for this extra output to cover the (marginal) production cost of this extra output. From the standpoint of social efficiency, the extra output should be produced. However, the firm is unwilling to produce any more because it would be forced to lower its price on *all* its output, leading to a net reduction in profits. There is still a deadweight loss despite the fact that profits have been driven to zero.

Also, notice that AC is not minimized in the long run equilibrium for a monopolistically competitive industry. We know that PC minimizes AC in the long run. Thus, monopolistically competitive firms produce at a higher AC than PC firms, another source of inefficiency. Sometimes this phenomenon is called <u>excess capacity</u> (as in P&R) because this firm is not producing enough, given its AC curve. Since it should be producing more if it were to minimize AC, we can think of the firm as running its factory at only 75% capacity, for instance.

However, we cannot conclude that this market structure is itself inefficient. Yes, P > MC and quantity is "too low," but in this case consumers gain something that might offset this loss in standard consumer surplus. Namely, they benefit from product differentiation. Consumers who prefer Coke are better off that all cola drinks are not Pepsi, and consumers who prefer Pepsi are better off that all cola drinks are not Coke. In other words, having a choice of substitutes makes consumers better off. Whether the gains from product differentiation outweigh the loss associated with reduced output cannot be determined from this analysis.

Oligopoly – Quantity vs. Price Competition

Suppose now that firms sell homogenous products, but that there are only a few firms in the market (an <u>oligopoly</u>). The important thing to recognize here is that firms will have to take their rivals' reactions into account when they set P and/or Q.

<u>Quantity Competition -- Cournot</u>

There are a couple of ways to think about oligopoly. On one hand, we can imagine that firms compete by setting quantity. That is, each firm chooses the amount of output it wants to produce, and then price is dictated by the market demand curve. This is the basis of the <u>Cournot model</u> (named after Augustin Cournot, who developed this model in the 1800s.

In this model, each firm chooses its quantity, taking as given the quantity chosen by other firms. We can see how this works in the following example:

EXAMPLE: Suppose that two firms collectively face a linear demand curve given by

$$Q = 1500 - 50P$$

Note that Q is total industry output, so that if q_1 and q_2 denote the output of firm 1 and firm 2 respectively, $Q = q_1 + q_2$. Let us further suppose that each firm faces a constant marginal cost of 1. If there are no fixed costs, then MC = AC = 1.

Each firm will choose its output to maximize its profits given its residual demand. In other words, each firm takes its rival's actions as given and chooses the profit-maximizing output level based on the demand "left over" after its rival has produced its output.

In this context, solving the firm's maximization problem is simple. We know that each firm will choose its output to equate MC and MR. Rewriting the demand curve as

$$q_1 + q_2 = 1500 - 50P$$

allows us to write firm 1's demand curve as

$$q_1 = 1500 - 50P - q_2$$

This is the amount of output that firm 1 can sell, given the output of firm 2 and the market price. So its inverse demand is

$$P = (1500 - q_2)/50 - (1/50)q_1$$

A rule of thumb will make this a bit easier. If the firm faces a linear demand curve, as this firm does, we can simply double the slope of this demand curve to get its MR. Doubling the slope of this (inverse) demand curve gives us the MR curve for firm 1:

$$MR = (1500 - q_2)/50 - (2/50)q_1$$

Now firm 1 can set MR = MC and solve for it's optimal quantity.

$$\begin{aligned} MR = MC &\Rightarrow (1500 - q_2)/50 - (2/50)q_1 = 1 \\ &\Rightarrow 1500 - q_2 - 2q_1 = 50 \\ &\Rightarrow q_1 = 725 - (1/2)q_2 \end{aligned}$$

Notice that the optimal q_1 is a function of firm 2's output. This makes sense because of the interaction of the two firms. If firm 2 produces a lot of output (flooding the market), there isn't much demand left for firm 1 to fill, so it produces relatively little. Firm 1's output, expressed in terms of its rival's output, is usually called the firm's reaction function or reaction curve because it gives firm 1's best response to any output choice by firm 2 (sometimes called a "best-response" function). We will write it as $q_1(q_2)$, indicating that q_1 is a function of q_2.

Of course, we can perform exactly the same analysis for firm 2. Its residual demand is

$$q_2 = 1500 - 50\,P - q_1$$

So its inverse demand is

$$P = (1500 - q_1)/50 - (1/50)q_2$$

Doubling the slope of this (inverse) demand curve gives us the MR curve for firm 2:

$$MR = (1500 - q_1)/50 - (2/50)q_2$$

and setting MR = MC yields firm 2's reaction function.

$$MR = MC \Rightarrow (1500 - q_1)/50 - (2/50)q_2 = 1$$
$$\Rightarrow 1500 - q1 - 2q2 = 50$$
$$\Rightarrow q_2(q_1) = 725 - (1/2)q_1$$

Notice that $q_1(q_2)$ and $q_2(q_1)$ are symmetric. This will always be the case when the duopolists have the same MC.

So we know firm 1's optimal reaction to any output produced by firm 2, and we know firm 2's optimal reaction to any output produced by firm 1. For an equilibrium, we need each firm to be responding optimally to the actions of its rival (this concept is called Nash equilibrium, and specifies that each firm be doing the best it can, given the actions of its rival). In other words, the Nash equilibrium in this industry occurs at the point that the firm's reaction functions intersect.

Another way to look at this is to note that firm 1 knows that firm 2 will compute $q_2(q_1)$, and firm 1 knows that firm 2 will produce this amount. Firm 1 will set its own quantity based on the assumption that firm 2 will respond optimally. Therefore, firm 1 will assume that firm 2 will produce $q_2(q_1)$ and choose q_1 accordingly. Thus, we find q_1 as

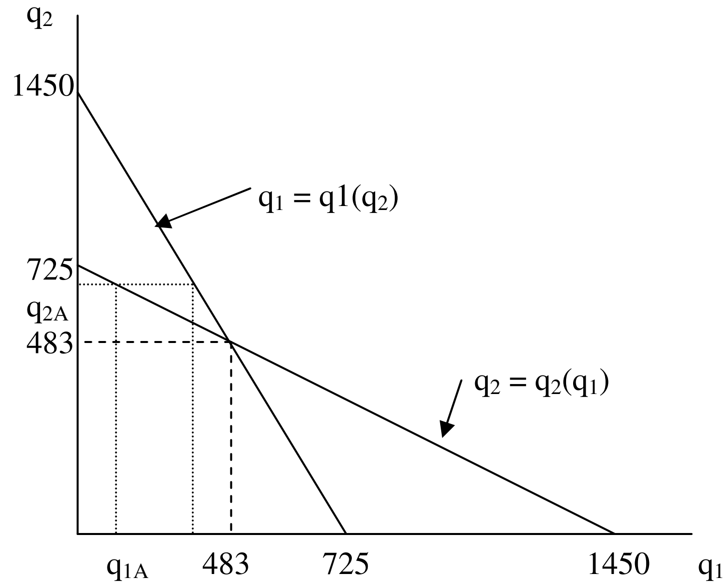
$$q_1 = 725 - (1/2)q_2(q_1) = 725 - (1/2)[725 - (1/2)q_1]$$
$$\Rightarrow q_1 = (1/2)(725) + (1/4)q_1$$
$$\Rightarrow (3/4)q_1 = (1/2)(725)$$
$$\Rightarrow q_1 = (2/3)(725) = 483.33$$

If $q_1 = 483.33$ then $q_2 = 725 - (1/2)q_1 = 483.33$. Since the reaction functions are symmetric, it should not be surprising that each firm ends up producing the same quantity.

Note that the output levels $(q_1, q_2) = (483.33, 483.33)$ is, in fact, a Nash equilibrium because neither firm regrets its output decision given the action of the other firm. Given that firm 2 produces 483.33, firm 1 wants to produce 483.33, which makes firm 2 want to produce 483.33, thus confirming the beliefs of firm 1.

In our Nash equilibrium, Q = 966.67 so P = $10.67. Note that each firm earns a profit of about $4673.

We can graph the firms' reaction functions to see how this equilibrium is achieved.



Suppose that firm 1 produces less than 483, say q_{1A}. Then firm 2's best response is to produce q_{2A}. But if firm 2 is producing q_{2A}, firm 1 no longer wants to produce q_{2A}! Thus, q_{2A} cannot be firm 1's equilibrium output. This logic holds for all output levels other that that given by the intersection of the reaction functions.

<u>Price Competition -- Bertrand</u>
45 years after the development of the Cournot model, Joseph Bertrand argued that firms actually set price, not quantity. The only difference between Bertrand's model and Cournot's model is that in the Cournot model, firms choose their output and then the market price is determined by the demand curve. In the Bertrand model, firms choose their prices and then each firm's output (sales) is determined by its demand. As it turns out, this seemingly minor change makes a tremendous difference in the resulting Nash equilibrium.

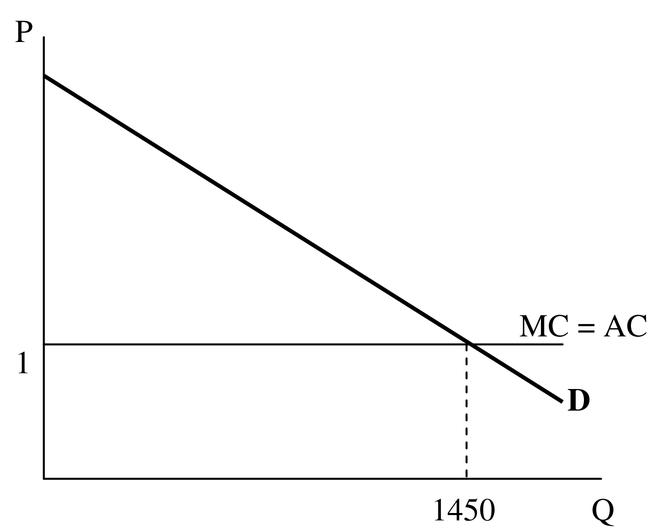
To reiterate, we have two firms selling homogenous products. Each firm has still has constant returns and the same MC, which we will assume is still MC = AC = 1 and the market demand function facing the duopoly is still Q = 1500 – 50 P.

Since the firm's products are identical, consumers will buy from whichever firm sets the lower price. Therefore, whichever firm sets the lower price captures the entire market and its rival's sales fall to zero.

The immediate implication of this observation is that the only Nash equilibrium occurs when both firms set P = MC = \$1. To see why, suppose that either firm, say firm 1, set $P_1 > \$1$. Then firm 2 could set its price, P_2, just a little lower than P_1 and capture the entire market and earn economic profits. In this case, firm 1 will regret its choice of P_1, since it could now set its price just a little lower than P_2, capture the entire market, and earn positive economic profits. But if firm 1 regrets its choice of P_1, it is impossible for any set of prices including P_1 to have been a Nash equilibrium, and the only thing we know about P_1 is that it was greater than MC. Therefore, in a Nash equilibrium, no firm can charge more than MC.

Note that if each firm chooses to set P = MC, this price combination (P_1, P_2) = (MC, MC) *is* a Nash equilibrium. If my rival sets P = MC, I have no incentive to raise my price, since I will now be charging a higher price than my rival, my sales will fall to zero, and I will make no profits (the same result that I face now). Likewise, I clearly have no incentive to undercut my rival, since in that case I would be choosing P < MC which guarantees a loss when I could have broken even by keeping P = MC. Therefore, the only Nash equilibrium in this model is for both firms to choose P = MC, resulting in zero economic profits.

We can represent this situation graphically.



If either firm sets a price greater than $1, it will be undercut. Thus, P is driven down to $1 and neither firm earns any economic profit. We don't know how much each firm will sell because consumers will be indifferent when choosing between the two firms, but we know that industry output will be 1450.

Notice that this is, of course, the perfectly competitive result. The price competition of Bertrand oligopoly forces the price down and output up, maximizing consumer (and social) surplus.

So which is it, Cournot or Bertrand? Both models attempt to explain the same phenomenon yet wind up with wildly different conclusions. Generally, the Cournot model is more appropriate for industrial and manufacturing industries (such as the auto industry) where, because it takes time to produce output, firms tend to choose their output levels and leave prices up to the market. A firm in such an industry couldn't accommodate the rush of orders it would get if it undercut its rivals.

On the other hand, the Bertrand model is more appropriate for modeling industries in which firms *can* accommodate a sudden rush of customers and therefore are price-setters. The airline industry is a classic example. American Airlines, for example, sets its fares very close to MC on routes where it faces competition. If it steals all of its rival's customers, so be it.

Collusion

Of course, if there are only a few firms in the industry, they could collude. In other words, the firms could all agree to charge the monopoly price, and divide the monopoly quantity and profits among them. This would be better than either of the two outcomes discussed above.

The problem with collusion is that is appears to be unsustainable, especially if firms do not have a long history of interaction with one another. The reason is that when firms form a cartel, they all benefit from higher profits. However, if any firm in the cartel cheats (by selling more than its quota or undercutting its rivals), it can earn enormous profits at the expense of the other firms. Thus, while all firms benefit from forming a cartel, each individual firm has an incentive to cheat.

To see why this poses such a problem for colluding firms, let's digress for a moment to consider a famous game: the Prisoners' Dilemma.

Prisoners' Dilemma: Consider the following scenario: two criminals (accomplices) are arrested for murder. The criminals are placed in separate rooms and offered the following deal by the district attorney:

> We don’t have enough evidence to convict both of you of murder, so if you both deny that you’ve committed the crime, the best I can do is to convict you of a lesser offense (burglary). You’ll each go to prison for 3 years. If you both confess to the crime, I’ll send you to both to prison for 25 years. However, if your accomplice denies that he committed the crime, and you confess and testify against him, I’ll reward your help by letting you go free and sending your accomplice to prison for 50 years.

Notice that each player (each prisoner) has two possible actions: confess or deny. We can set up a playoff matrix to make this easier to analyze.

		Player B	
		Confess	Deny
Player A	Confess:	(25, 25)	(0, 50)
	Deny:	(50, 0)	(3, 3)

Each cell in the matrix shows the payoffs to each player given their choices of actions. For instance, if A denies and B confesses, A gets 50 years and B gets 5 years. Note that the payoffs go (A’s payoff, B’s payoff), and that, in this case, a lower number is better.

Notice that no matter what B does, A does best to confess. If B confesses, A should confess since he would rather go to prison for 25 years than 50 years. If B denies, A should again confess since he would rather go free than go to prison for 3 years. No rational person in prisoner A’s shoes would do anything other than confess.

Note that “confess” is also the correct strategy for prisoner B, although this doesn’t really matter. B knows that A will confess because that is A’s correct strategy. Therefore, B does best to confess since he would rather go to prison for 25 years rather than 50.

Remember that a Nash equilibrium is a set of strategies in which each player is doing the best he can given the strategies chosen by the other players. Notice that (confess, confess) is a Nash equilibrium because neither player can improve his position given that the other is confessing. Furthermore, it is easy to verify that (confess, confess) is the only Nash equilibrium.

e.g.) Is (deny, deny) a Nash equilibrium? No. If B denies, A does better to confess. Therefore (deny, deny) cannot be a NE.

The perverse result of this game (the “dilemma” faced by the prisoners) is that both end up going to prison for 25 years, while they hypothetically could have cooperated and faced only 3 years in prison. The strategic nature of the game, and the interdependence of their actions, precluded this cooperation.

Back to Collusion: You may have noticed that our analysis of a cartel's collapse sounds a lot like the Prisoners' Dilemma: everyone is better off if they all cooperate, but everyone has an incentive to cheat, so the collusive outcome cannot be reached. Actually, it is quite easy to represent the maintenance of a cartel as a Prisoners' Dilemma.

Suppose for simplicity that there are only two firms in the market and they have the possibility of forming a cartel. Let's suppose that each firm has two options: setting a high price and setting a low price. (In the analysis above, this might correspond to setting $P = P^*$ and $P = P^C$. Alternately we could think of this as setting a high or low Q. We've already seen that the two approaches amount to the same thing and this should already be intuitively clear).

If the firms both set the low (competitive) price, they earn low profits. If they each set the high (monopoly) price, they each earn high profits. If one firm sets a high price while the other sets a low price, the low-price firm captures the entire market and makes an especially large profit while the high-priced firm sells nothing and earns no profit. We can represent the strategic form game as follows:

		Firm B	
		High Price	Low Price
Firm A	High Price:	(30, 30)	(-10, 50)
	Low Price:	(50, -10)	(0, 0)

You should be able to verify that the only E is (low, low). Both firms would like to cooperate and set a high price. But given that your rival is setting a high price, you do better to "cheat" and set a low price. Hence collusion is unsustainable since it is not a Nash equilibrium.