9

Static Games and Cournot Competition

One of the most successful companies in the history of business is Coca-Cola. Indeed, "Coca-Cola" is said to be the second most well-known phrase in the world, the first being "okay."[1] Yet despite its iconic status in American popular culture, Coca-Cola is not a monopoly. Coca-Cola shares the carbonated soft drink market share with its archrival PepsiCo. An ongoing battle for market share has engaged these two companies for almost 100 years. The cola wars have been fought with a number of strategies, one of which is the frequent introduction of new soft drink products. Pepsi launched *Pepsi Vanilla* in the summer of 2003 in response to the year-earlier introduction of *Vanilla Coke*. In 2006, Coke initiated its biggest new brand campaign in twenty-two years for its new diet drink, *Coke Zero*. This followed Pepsi's revitalization of its *Pepsi One* brand made with Splenda sweetener instead of Aspartame.

In fighting these cola wars, each company must identify and implement the strategy that it believes is best suited to gaining a competitive advantage in the soft drink industry. If Coca-Cola were a monopoly, it would not have to worry about any rivals. Nebbither does a perfectly competitive firm. The pure monopolist has no rivals and the company in a perfectly competitive market has no effect on its rivals. Each such firm is so small that its decisions have no effect at all on the industry.

However, Coke, Pepsi, and many if not the majority of other real world firms live in the middle ground of *oligopoly*, in which firms have visible rivals with whom strategic interaction is a fact of life. Each firm is aware that its actions affect others, and therefore, prompt *re*actions. Each firm must, therefore, take these interactions into account when making a decision about prices, output, or other business actions. *Game theory* is the analytic framework used to formally analyze strategic interaction. As a result, game theory and the study of oligopoly are closely intertwined. In this chapter we introduce some basic game theoretic analysis to show how it may be used to understand oligopoly markets.

Game theory itself is divided into two branches: *noncooperative* and *cooperative* game theory.[2] The essential difference between these two branches is that in noncooperative

[1] "Coca-Cola is okay" has been claimed to be understood in more places by more people than any other sentence (Tedlow 1996).

[2] A good textbook that offers a more formal treatment of game theory and its applications to economics is Rasmusen (2007).

games, the unit of analysis is the individual decision-maker or player, for example, the firm. By contrast, cooperative game theory takes the unit of analysis to be a group or a coalition of players, for example, a group of firms. We will focus almost exclusively on noncooperative game theory. The individual player will be the firm. The *rules of the game* will define how competition between the different players or firms takes place. The noncooperative setting means that each player is concerned only with doing as well as possible for himself or herself, and not in advancing a more general group interest. As we shall see though, such noncooperative behavior can sometimes look very much like cooperative behavior because cooperation sometimes turns out to maximize the well-being of each individual player, as well.

Two basic assumptions underlie the application of noncooperative game theory to oligopoly. The first is that *firms are rational*. They pursue well-defined goals, principally profit maximization. The second basic assumption is that firms apply their rationality to the process of *reasoning strategically*. That is, in making its decisions, each firm uses all the knowledge it has to form expectations regarding how other firms will behave. The motivation behind these assumptions is that our ultimate goal is to understand and predict how real firms will act. We assume that firms are rational and reason strategically because we suspect that real firms do precisely this or will be forced to do so by market pressures. Hence, understanding what rational and strategic behavior implies ought to be useful for understanding and predicting real world outcomes.

There is one caution that any introduction to the study of oligopoly must include. It is that, unlike the textbook competition and monopoly cases, there is no single, standard oligopoly model. Differences in the rules of the game such as the number of players, the information available to the various players, and the timing of each player's actions all conspire to yield a number of possible scenarios. Yet while there is not a single theory or model of oligopoly, common themes and insights from the various models of oligopoly do emerge. Understanding these broad concepts is our goal for the next three chapters. Moreover, we should add that the lack of one single oligopoly model is not entirely a disadvantage. Rather, it means that one has a rich assortment of models from which to choose for any particular investigation. One model will be appropriate for some settings, a different model for other settings. Because the real business world environment is quite diverse, it is useful to have a variety of analyses from which to draw. We will present three different oligopoly models. In this chapter we introduce the Cournot (1836) model of oligopoly, in the next chapter the Bertrand model, and then in Chapter 11 the Stackelberg model.

9.1 STRATEGIC INTERACTION: INTRODUCTION TO GAME THEORY

In game theory, each player's decision or plan of action is called a *strategy*. A list of strategies showing one particular strategy choice for each player is called a *strategy combination*. Any given strategy combination determines the *outcome* of the game, which describes the payoffs or final net gains earned by each player. In the context of oligopoly theory, these payoffs are naturally interpreted as each firm's profit.

For a game to be interesting, at least one player must be able to choose from more than one strategy so that there will be more than one possible *strategy combination*, and thus more than one possible outcome to the game. Yet while there may be many possible outcomes, not all of these will be *equilibrium* outcomes. By equilibrium we mean a *strategy combination* such that no firm has an incentive to *change* the strategy it is currently using given that no other firm changes its current strategy. If this is the case, then the combination of strategies across firms will remain unaltered because no one is changing his or her behavior. The market or game will come to rest. Nobel Laureate John Nash developed this notion of an equilibrium strategy combination for a noncooperative game. In his honor, it is commonly referred to as the Nash equilibrium concept.[3]

In the oligopoly models studied in the next three chapters, a firm's strategy focuses on either its price choice or its output choice. Each firm chooses either the price it will set for its product or how much of that product to produce. A corresponding Nash equilibrium will, therefore, be either a set of prices, one for each firm, or a set of production levels, again one for each firm, for which no firm wishes to change its price (quantity) decision given those of all the other firms.

We note parenthetically here that, unlike the monopoly case, the price strategy outcome differs from the quantity strategy outcome in oligopoly models. For a monopolist, the choice of price implies—via the market demand curve—a unique output. In other words, the monopolist will achieve the same market outcome whether the profit-maximizing price or the profit-maximizing output is chosen.[4] Matters are different in an oligopoly setting. When firms interact strategically, the market outcome obtained when each firm chooses the price will usually differ from the outcome obtained when each firm chooses its output level. The fact that the outcome depends on whether the rules of the game specify a price strategy or a quantity strategy is just one of the reasons that the study of oligopoly does not yield a unique set of theoretical predictions.

Because interaction is the central fact of life for an oligopolist, rational strategic action requires that such interaction be recognized. For example, when one firm in an oligopoly market lowers its price, its rivals will notice the effect as they lose customers to the price-cutter. If these firms then lower their price too, they may win back their original customers. Because prices have fallen throughout the industry, the quantity demanded at each firm will increase. However, each firm will now be meeting that demand at a lower price that earns a lower mark-up. Our assumption that the oligopoly firm is a rational strategic actor means that the firm will understand and anticipate this chain of events *and* that the firm will include this information in making the decision of whether or not to lower prices in the first place.

Our opening story about carbonated beverages is an example of such an interaction, except that instead of a price decision, Coca-Cola and Pepsi were making product design choices. In doing so, each forms some idea as to how its rival will react. It would be *irrational* for Coke to anticipate no reaction from Pepsi, when, in fact, Coke understands that not reacting is not in Pepsi's interest. Similarly, if Coke lowers the price of its soft drinks, it doesn't make sense for Coke to hope that Pepsi will continue to charge a high price if Coke knows that Pepsi would do better to match its price reduction.

How can an oligopolist anticipate what the response of its rivals will be to any specific action? The best way to make such a prediction is to have information regarding the structure of the market and the strategy choices available to other firms. In a symmetric

3 Nash shared the 1994 prize with two other game theorists, R. Selten and J. Harsanyi. The award to the three game theorists served as widely publicized recognition of the importance game theory has achieved as a way of thinking in economic analysis. See Schelling's (1960) seminal work for much of the intuitive foundation of strategic analysis.

4 Competitive firms have no option as to which choice variable—price or quantity—to select. Competitive firms by definition cannot make a price choice. They are price-takers and can only choose the quantity of output they sell.

situation in which all firms are identical, such information is readily available. Any one firm can proceed by asking itself, "What would I do if I were the other player?" Sometimes, even when firms are not symmetric, they will still have enough experience, business "savvy," or other information to be fairly confident regarding their rivals' behavior. As we shall see later, precisely what information firms have about each other is a crucial element determining the final outcome of the game.

Another crucial element in determining the outcome of the game is the time-dimension of the strategic interaction. In a two-firm oligopoly or *duopoly*, such as Coca-Cola and Pepsi, we can imagine that one firm, say Coca-Cola, makes its choice to introduce *Vanilla Coke* first. Then in the next period, the other firm, Pepsi, follows with its choice. In that case, the strategic interaction is *sequential*. Each firm moves in order and each, when its turn comes, must think strategically about how the course of action it is about to choose will affect the future action of the other firm and how those *re*actions will then feed back on its own future choices. Chess and Checkers are each a classic example of a two-person, sequential game. Sequential games are often called dynamic games.

Alternatively, both players might make their choices *simultaneously*, thereby acting without knowledge as to what the other player has actually done.[5] Yet even though the other player's choice is unknown, knowledge of the strategy choices available to the other player permits a player to think rationally and strategically about what other players will choose. The childhood game, "Rock-Paper-Scissors" is an example of a simultaneous two-person game. Such simultaneous games are often called static games.

Whether the game is sequential or simultaneous, the requirement that the strategic firm rationally predicts the choices of its rivals is the same. Once it has done this, the firm may then choose what action is in its own best interest. In other words, being rational means that the firm's choice of strategy is the optimal (profit-maximizing) choice against the anticipated optimal actions of its rivals. When each firm does this, and when each has, as a result of rational strategizing, correctly predicted the choice of the others, we will obtain a Nash equilibrium. In this chapter we will focus on solving for Nash equilibria in simultaneous or static games.

9.2 DOMINANT AND DOMINATED STRATEGIES

Sometimes Nash equilibria are rather easy to determine. This is because some of a firm's possible strategies may be *dominated*. For example, suppose that we have two firms in a market, A and B, and that one of A's strategies is such that it is *never* a profit-maximizing strategy regardless of the choice made by B. That is, there is always an alternative strategy for firm A that yields higher profits than does the strategy in question. Then we say that the strategy in question is dominated. It will never rationally be chosen. Player A would never choose a dominated strategy because to do so would be to guarantee that A's profit was not maximized. No matter what B does, the dominated strategy does worse for A than one of A's other strategies. In turn, this means that in determining the game's equilibrium, we do not have to worry about any strategy combinations that include the dominated strategy. Because these will never occur, they cannot possibly be part of the equilibrium outcome.

[5] The important aspect of simultaneous games is not that the firms involved actually make their decisions at the same time. Rather, it is that no firm can *observe* any other firm's choice before making its own. This lack of information makes the actions of each firm effectively simultaneous.

Dominated strategies can be eliminated one by one. Once the dominated strategies for one firm have been eliminated, we can turn to the other firms to see if any of their strategies are dominated given the strategies still remaining for the first firm that we examined. We can proceed firm-by-firm, eliminating all dominated strategies until only non-dominated ones remain available to each player. Often but not always, this iterative procedure of eliminating dominated strategies leaves one or more players with only one strategy choice remaining.[6] It is then a simple matter to determine the game's outcome because, for such firms, their course of action is clear.

As an example, consider the case of two airlines, Delta and American, each offering a daily flight from Boston to Budapest. We assume that each firm has already set a price for the flight but that the departure time is still undecided. Departure time is the strategy choice in this game. We also assume that the two firms choose departure times simultaneously. Neither can observe the departure time selected by the other before it makes its own departure time selection. Managers for each airline do realize, however, that at the very time American's managers are meeting to make their choice, Delta's managers are as well. The two firms are engaged in a strategic game of simultaneous moves.

In part, the choice of departure time will depend upon consumer preferences. Suppose that market research has shown that 70 percent of the potential clientele for the flight would prefer to leave Boston in the evening and arrive in Budapest the next morning. The remaining 30 percent prefer a morning Boston departure and arrival in Budapest late in the evening of the same day. Both firms know this distribution of consumer preferences. Both also know that, if the two airlines choose the same flight time, they split the market. Profits at each carrier are directly proportional to the number of passengers carried so that each wishes to maximize its share of the market.

If they are rational and strategic, Delta's managers will reason as follows: If American flies in the morning, then we at Delta can either fly at night and serve 70 percent of the market or, like American, depart in the morning in which case we (Delta) will serve 15 percent of the market (half of the 30 percent served by the two carriers in total). On the other hand, if American chooses an evening flight time, then we at Delta may choose either a night departure as well, and serve 35 percent (half of 70 percent) of the market or, instead, offer a morning flight and fly 30 percent of the market.

A little reflection will make clear that Delta does better by scheduling an evening flight *no matter which departure time American chooses*. In other words, choosing a *morning* departure time is a dominated strategy. If Delta is interested in maximizing profits, it will never select the morning flight option. But of course, American's managers will reason similarly. They will recognize that flying at night is their best choice regardless of Delta's selection. The only equilibrium outcome for this game is to have both airlines choose an evening departure time.

Table 9.1 illustrates the logic just described. The table shows four entries, each consisting of a pair of values. These entries describe the payoffs, or market shares, associated with the four feasible strategy combinations of the game. American's strategy choices are shown as the columns, while Delta's choices are shown as the rows. The pair of values at each row-column intersection gives the payoffs to each carrier if that particular strategy combination occurs. The first (left-hand) value of each pair is the payoff—the percent of

[6] If the process continues until only one strategy remains for each player, then we have found an iterated dominance equilibrium.

Table 9.1 Strategy combinations and firm payoffs in the flight departure game

		American	
		Morning	Evening
Delta	Morning	(15, 15)	(30, 70)
	Evening	(70, 30)	(35, 35)

the total potential passenger market—that goes to Delta. The second (right-hand) value is the payoff to American.

Now we put ourselves in the shoes of Delta's managers and ask first what Delta should do if American chooses a morning flight. The answer is obvious. If Delta also chooses a morning flight then Delta's market share will be 15 percent, whereas if Delta chooses an evening flight its market share will be 70 percent. The evening flight is clearly the better choice. Now consider Delta's response should American choose an evening flight. If Delta opts for a morning departure, its market share is 30 percent, whereas if it goes for an evening departure its market share is 35 percent. Once again, the evening departure is the better choice. In other words, no matter what American does, Delta will never choose to depart in the morning. Whatever the equilibrium outcome is, it must involve Delta choosing an evening flight.

If we now place ourselves in American's shoes, we get the same result. A morning flight is never American's rational choice. Just as it was for Delta, flying in the morning is a dominated strategy for American. Hence, just like Delta, American will always choose the evening departure time.

The outcome of the game is now fully determined. Both carriers will choose an evening departure and share equally the 70 percent of the potential Boston-to-Budapest flyers who prefer that time. That this is a Nash equilibrium is easy to see by virtue of the dominated strategy argument. Clearly, neither carrier has an incentive to change its choice from evening to morning because neither carrier would ever choose a morning flight time in any case.

Solving the flight departure game was easy because each carrier had only two strategies and for each player one of the strategies—the morning flight—was dominated. To put it another way, we might refer to the evening departure strategy as *dominant*. A dominant strategy is one that outperforms all of a firm's other strategies *no matter what its rivals do*. That is, it leads to higher profits (or sales, or growth, or whatever the objective is) than any other strategy the firm might pursue regardless of the strategies selected by the firm's rivals. This does not imply that a dominant strategy will lead a firm to earn higher profits than its competitors. It only means that the firm will do the best it possibly can if it chooses such a strategy. Whether its payoff is as good as, or better than the payoffs obtained by its rivals depends on the structure of the game.

Except when the number of strategy choices is two, a firm may have some *dominated* strategies—choices that are never good ones because better ones are available—but not have any *dominant* strategies, or a choice that always yields better results than all others. Sometimes, a firm will have neither a dominant nor a dominated strategy. But for a firm that has a dominant strategy, the choice is clear. Use it! Such a firm really does not have to think very much about what other firms do.

Table 9.2 Strategy combinations and firm payoffs in the modified flight departure game

		American	
		Morning	Evening
Delta	Morning	(18, 12)	(30, 70)
	Evening	(70, 30)	(42, 28)

Let's rework the departure time game so that at least one firm has no dominated or dominant strategy. To do this, we will now suppose that because of a frequent flyer program, some of the potential Boston-to-Budapest flyers prefer Delta even if the two carriers fly at the same time. Specifically, assume now that departing at the same time does not yield an even split of customers between the two carriers. Instead, whenever the two carriers schedule identical departure times, Delta gets 60 percent of the passengers and American gets only 40 percent. Table 9.2 depicts the new payoffs for each strategy combination.

As can be seen from the table, an evening flight is still the dominant strategy for Delta. It always carries more passengers by choosing an evening flight than it would by choosing a morning flight, regardless of what American does. However, American's strategy choices are no longer so clear. If Delta chooses a morning flight, American should fly at night. But if Delta chooses an evening departure time, American does better by flying in the morning.

Assuming that each carrier knows the payoffs described by Table 9.2, however, the game's outcome remains clear. Looking at the table, American can readily determine that Delta will always select an evening flight. Knowing that Delta will choose an evening departure, it is then an easy matter for American to select a morning departure as its best response. The equilibrium outcome for this modified departure time game is therefore just as transparent as that for the earlier version. In this case, the equilibrium involves Delta choosing an evening flight and American opting to fly in the morning. Again, it is easily verified that this equilibrium satisfies the Nash criteria.

In solving both the previous games, we made extensive use of the ability to rule out dominated strategies and, when possible, to focus on dominant ones.[7] We showed that the outcomes obtained by this process were Nash equilibrium outcomes. However in many games no dominated or dominant strategies can be found. In such cases, the Nash equilibrium concept becomes more than just a criterion to check our analysis. It becomes part of the solution procedure itself. This is because rational, strategic firms will use the Nash concept to determine the reactions of their rivals to their own strategic choice. In the modified departure time game just described, for instance, Delta can work out that if it selects an evening departure, then its rival American will choose a morning flight. Delta can therefore deduce that the strategy combination of both carriers flying at night can never be an equilibrium—in the Nash sense—because if that outcome occurred, American would have a clear incentive to change its choice.

[7] Some care needs to be taken in ruling out dominated strategies. While one can eliminate *strictly* dominated strategies as a rational choice, *weakly* dominated strategies cannot be so ruled out. A strategy is *weakly* dominated if there exists some other strategy, which is possibly better but never worse, yielding a higher payoff in some strategy combinations and never yielding a lower payoff. The Nash equilibrium may be affected by the order of exclusion of *weakly* dominated strategies. See Mas-Colell, et al. (1995), 238–41.

9.3 NASH EQUILIBRIUM AS A SOLUTION CONCEPT

In order to understand more deeply how to use the Nash equilibrium concept to solve a game, let's change the Boston-to-Budapest game to a game in prices.

Although many price strategies are available to each firm in reality, let's limit ourselves here to just three—a low, medium, and high price. This yields nine possible strategy combinations. We will assume that the payoffs (in this case profits) for each such combination are described by the new airfare game matrix shown in Table 9.3. As is now our convention, the payoff entries in each row-column intersection show the profit of the row player (Delta) as the first entry in each case.

We could in principle check each cell of the matrix in Table 9.3 to determine if it satisfies the Nash equilibrium requirement that neither airline wants to change its strategy. For example, examination of the upper left-hand corner cell quickly reveals that the strategy combination $P_D = P_A = Low$ is not a Nash equilibrium, because each airline in that case would want to switch to a medium price. Similarly, the middle cell of the left-hand column with $P_D = Medium$ and $P_A = Low$ is not a Nash equilibrium. While Delta cannot improve its profit given that American is charging a low fare, American would do better by raising its own fare to the medium level given that Delta is already doing just that. If American followed this inclination, however, it would move the game into the $P_A = P_D = Medium$ cell in the middle of the matrix, and it is easy to see that that strategy combination *is* a Nash equilibrium. Indeed, following this same procedure for all other cells would reveal that $P_A = P_D = Medium$ is the unique Nash equilibrium for this game.

A more appealing approach is to start by identifying that the strategy of setting a high price is never the best strategy for either firm. That is, $P_D = High$ and $P_A = High$ are each dominated strategies for Delta and American, respectively. We can therefore eliminate the third column and third row cells from consideration. Having done that, we are left with the simple 2×2 game in which each firm sets either a low or medium price. Within those choices however, we now find that $P_D = Low$ and $P_A = Low$ are each dominated by, respectively, $P_D = Medium$ and $P_A = Medium$. Thus, peeling off the dominated strategies as suggested earlier allows us to reach the result that $P_D = P_A = Medium$ is the unique Nash equilibrium for this game.

The fact that there are a number of ways of identifying the Nash equilibrium for any game suggests that behind each is a common but deeper meaning of the Nash concept. This more fundamental insight may be made clear by noting that if players are rational, we would not expect the game above to evolve along the path suggested by our first solution technique in which both firms start by setting $P_D = P_A = Low$, followed by Delta choosing $P_D = Medium$ and then American doing the same thing. Instead, given the fact that both players know the above payoff matrix and the rules of the game, each should choose a medium price right from the start. Why? Consider again Delta's management. Looking at

Table 9.3 Payoff matrix for the airfare game

		American		
		$P_A = Low$	$P_A = Medium$	$P_A = High$
Delta	$P_D = Low$	\$15,000 \$15,000	\$25,000 \$22,000	\$40,000 \$20,000
	$P_D = Medium$	\$22,000 \$25,000	\$35,000 \$35,000	\$38,000 \$33,000
	$P_D = High$	\$20,000 \$40,000	\$33,000 \$38,000	\$30,000 \$30,000

the game, they can see that American will never set $P_A = High$. Thus, recognizing that $P_A = High$ is a dominated strategy allows Delta to restrict American's possible choices to either $P_A = Low$ or $P_A = Medium$. However, further consideration reveals that American would only ever choose $P_A = Low$ if it thought that Delta was going to set $P_D = High$, for if Delta chooses a low or medium price, American does better by choosing a medium price as well. Yet it is not rational for American ever to think that Delta would choose $P_D = High$ because looking at the same table, American can see that whether it sets a low or medium price, Delta always does best by choosing $P_D = Medium$. Therefore, it would be foolish for American ever to set $P_A = Low$. Thus, Delta can set $P_D = Medium$ with a complete confidence that American will do the same thing and, more importantly, that when Delta's choice of a medium price is revealed, the predicted American strategy of $P_A = Medium$ is in fact precisely the choice American would want to make given Delta's choice of $P_D = Medium$. Of course, American can follow the exact same reasoning. As a result, each firm will select a medium pricing strategy from the start. There will be no action and reaction that eventually converge to the $P_D = P_A = Medium$ combination. It is this deeper logical consistency—the fact that each firm's strategy choice reflects its best response to the choice predicted for its rival which, in turn, is the rival's best response to the firm's own selected strategy—that gives the Nash equilibrium concept its true power.

9.1

Practice Problem

Firm 1 and Firm 2 are movie producers. Each has the option of producing a blockbuster romance or a blockbuster suspense film. The payoff matrix displaying the payoffs for each of the four possible strategy combinations (in thousands) is shown below, with Firm 1's payoff listed first. The game is played simultaneously. Determine the Nash equilibrium outcome.

		Firm 2	
		Romance	Suspense
Firm 1	Romance	(\$900, \$900)	(\$400, \$1,000)
	Suspense	(\$1,000 \$400)	(\$750, \$750)

9.4 STATIC MODELS OF OLIGOPOLY: THE COURNOT MODEL

All the games of the previous section are single period or static. Delta and American, for example, are assumed to choose either their departure times or their airfares simultaneously and without regard to the possibility that, at some later date, they might play the game again. This is a feature of earlier work on modeling oligopoly markets. Firms in these models "meet only once" and the market clears once-and-for-all. There is no sequential movement over time and no repetition of the interaction. These may be limitations. Yet the analysis is still capable of generating important insights. Moreover, studying such static models is a good preparation for later examining dynamic models.

The most well known static oligopoly models are the Cournot and Bertrand models, each named after its respective author. Both works were completed in the 19th century, hence these models were not expressed in the formal language of modern game theory. Nevertheless, the equilibrium proposed by the author of each model fully anticipates the

Nash concept. Indeed, for this reason one often sees that outcome referred to as the Cournot-Nash or the Bertrand-Nash equilibrium in tribute to these two early scholars and their anticipation of Nash's result. Presenting both models also helps us to underscore a critical point that should be recognized by game theory students everywhere, namely, that the rules of the game matter . . . a lot. The rules for the Cournot and Nash game are identical in all respects except for the variable of strategic choice. In the Cournot model, that variable is the firm's output level, whereas in the Bertrand model it is price. Yet this seemingly small change has enormous implications for the outcome of the game.

Augustin Cournot, a French mathematician, published his model in 1836. Although its insights remained largely unrecognized for the next 100 years, it is now at the foundation of models of oligopolistic markets. The story that Cournot told to motivate his analysis went as follows. Assume a single firm wishes to enter a market currently supplied by a monopoly. The entrant is able to offer a product that is identical in all respects to that of the incumbent monopolist and to produce it at the same unit cost. Entry is attractive because under the assumption of constant and identical costs, we know that the monopolist is producing where price is greater than marginal cost, which means that the price also exceeds the marginal cost of the would-be entrant. Hence, the entrant firm will see that it can profitably sell some amount in this market. However the new entrant will, Cournot reasoned, choose an output level that maximizes its profit, *after taking account of the output being sold by the monopolist*.

Of course, if entry occurred and the new firm produced its chosen output, the monopolist would react. Before entry, the monopolist chose a profit-maximizing output assuming no other rivals. Now, the former monopolist will have to re-optimize and choose a new level. In so doing, the monopolist will (as did the new entrant previously) choose an output level that maximizes profits *given the output sold by the new rival firm*.

This process of each firm choosing an output conditional on the other's output choice is to be repeated—at least as a mental exercise. For every output choice by the incumbent, firm 1, the entrant, firm 2, is shown to have a unique, profit-maximizing response and vise-versa. Cournot called the graph representations of these responses Reaction Curves. Each firm has its own Reaction Curve that can be graphed in the q_1q_2 quadrant. That Cournot anticipated Nash is evidenced by the fact that he described the equilibrium outcome of this process as that pair of output levels at which each firm's output choice is the profit-maximizing response to the other's quantity. Otherwise, Cournot reasoned, at least one firm would wish to change its production level. A further appealing aspect of Cournot's duopoly model is that the equilibrium price resulting from the output choices of the two firms is below that of the pure monopoly outcome. Yet it is also greater than that which would occur if there were not two firms but many firms and pure competition prevailed. In other words, the Cournot model carries the intuitive implication that more competition is better than less.

To present Cournot's analysis more formally we assume that the industry *inverse* demand curve[8] is linear, and can be described by:

$$P = A - BQ = A - B(q_1 + q_2) \tag{9.1}$$

where Q is the sum of each firm's production, i.e., the total amount sold on the market; q_1 is the amount of output chosen by firm 1, the incumbent firm; and q_2 is the amount of

[8] See Green and Newberry (1992) for a model in which firms compete in supply schedules relating the volume of output offered to a specific price. See also Wolfram (1999).

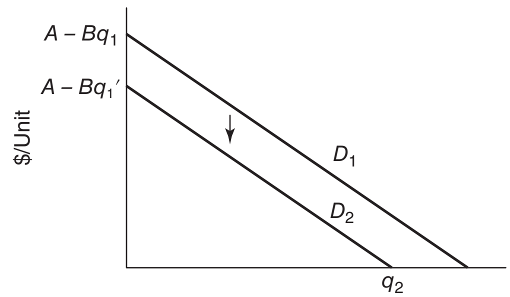

Figure 9.1 Firm 2's demand curve in the Cournot duopoly game depends on firm 1's output
An increase in q_1 to q_1' shifts D_2, the demand curve facing Firm 2, downwards.

output chosen by firm 2, the new competitor. As noted earlier, we shall also assume that each firm faces the same, constant marginal cost of production, c.

If we now consider firm 2 alone, and take firm 1's output, q_1, as given, the inverse demand curve, facing firm 2 is:

$$P = A - Bq_1 - Bq_2 \tag{9.2}$$

which is formally identical to (9.1). However, from firm 2's perspective, the first two terms on the right-hand side are not part of its decision-making, and can be taken as given. In other words, those two terms together form the intercept of firm 2's perceived demand curve so that firm 2 understands that the only impact *its* output choice has on price is given by the last term of the equation, namely, $-Bq_2$. Note, however, that any change in the anticipated output choice of the firm would be communicated to firm 2 by means of a shift in firm 2's perceived demand curve. Figure 9.1 illustrates this point.

As we can see from Figure 9.1, a different choice of output by firm 1 will imply a different demand curve for firm 2 and, correspondingly, a different profit-maximizing output for firm 2. Thus, for each choice of q_1 there will be a different optimal level of q_2. We can solve for this relationship algebraically, as follows. Associated with each demand curve illustrated in Figure 9.1 there is a marginal revenue curve that is twice as steeply sloped as discussed in Chapter 2. That is, firm 2's marginal revenue curve is also a function of q_1 given by:

$$MR_2 = (A - Bq_1) - 2Bq_2 \tag{9.3}$$

Marginal cost for each firm is constant at c. Setting marginal revenue MR_2 equal to marginal cost c, as required for profit-maximization, and solving for q_2^* yields firm 2's Reaction Curve. Thus we have $MR_2 = c$, which implies that $A - Bq_1 - 2Bq_2^* = c$ or $2Bq_2^* = A - c - Bq_1$. Further simplification then gives the Reaction Curve for firm 2:

$$q_2^* = \frac{(A - c)}{2B} - \frac{q_1}{2} \tag{9.4}$$

Equation (9.4) describes firm 2's best output choice, q_2^*, for every choice of q_1. Note that the relationship is a negative one. Every increase in firm 1's output lowers firm 2's demand and marginal revenue curves and, with a constant marginal cost, also lowers firm 2's profit-maximizing output.

Of course, matters work both ways. We may symmetrically re-work the industry demand curve to show that firm 1's individual demand depends similarly on firm 2's choice of output, so that as q_2 changes, so does the profit-maximizing choice of q_1. Then, we may analogously derive firm 1's Reaction Curve giving its best choice of q_1 for each alternative possible value of q_2. By symmetry with firm 2, this is given by:

$$q_1^* = \frac{(A-c)}{2B} - \frac{q_2}{2} \tag{9.5}$$

As was the case for firm 2, firm 1's profit maximizing output level q_1^* falls as q_2 increases.[9] The Reaction Curve for each firm is shown in Figure 9.2 in which the strategic variables for each firm and firm outputs are on the axes.

Consider first, the Reaction Curve of firm 1, the initial monopolist. This curve says that if firm 2 produces nothing, then firm 1 should optimally produce quantity $\frac{(A-c)}{2B}$, which is, in fact, the pure monopoly level, at which we assumed firm 1 to be producing in the first place. Now consider the Reaction Curve for firm 2. That curve shows that if firm 1 were producing at the assumed level of $\frac{(A-c)}{2B}$, then firm 2's best bet is to produce at level $\frac{(A-c)}{4B}$, that is, firm 2 should enter the market. However, if firm 2 does choose that level, then firm 1 will no longer do best by producing the monopoly level. Instead, firm 1 will maximize profits by selecting quantity $q_1 = \frac{3(A-c)}{8B}$.

As Cournot understood, none of the output or strategy combinations just described corresponds to an equilibrium outcome. In each case, the reaction of one firm is based upon a choice of output for the other firm that is not, itself, that other firm's best reaction to the initial chooser's selection. For the outcome to be an equilibrium, it must be the case that each firm is responding optimally to the (optimal) choice of its rival. For that to be the case

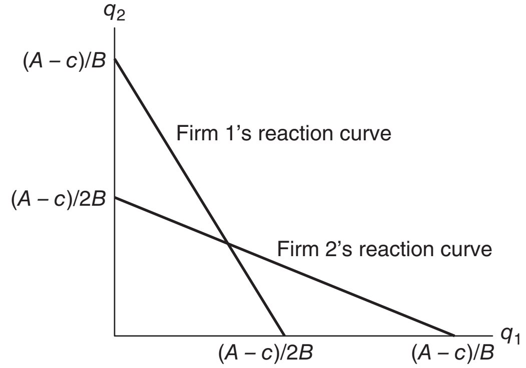

Figure 9.2 Best response (Reaction) curves for the Cournot duopoly model

9 We could alternatively solve for q_2^* by writing firm 2's profit function, Π^2, as revenue less cost, or: $\Pi_2(q_1, q_2) = (A - Bq_1 - Bq_2)q_2 - cq_2 = (A - Bq_1 - c)q_2 - Bq_2{}^2$. When we differentiate this expression with respect to q_2 and set the result equal to 0, (first order condition for maximization), and then solve for q_2^*, we get the same result as equation (9.4). A similar procedure may be used to obtain q_1^*.

though, requires that *both* firms be on their respective Reaction Curves. This happens at only one point in Figure 9.2, namely, the intersection of the two Reaction Curves.

To see how this works, recall the Reaction Function for firm 2: $q_2^* = \frac{(A-c)}{2B} - \frac{q_1}{2}$. We know and firm 2 knows that in an equilibrium, firm 1 must also be on its Reaction Curve, or that $q_1^* = \frac{(A-c)}{2B} - \frac{q_2}{2}$. Substituting this into firm 2's Reaction Curve allows firm 2 (and also us) to solve for:

$$q_2^* = \frac{A-c}{2B} - \frac{1}{2}\left(\frac{A-c}{2B} - \frac{q_2^*}{2}\right) = \frac{A-c}{4B} + \frac{q_2^*}{4} \tag{9.6}$$

so that $\frac{3q_2^*}{4} = \frac{A-c}{4B}$. In turn, this implies: $q_2^* = \frac{(A-c)}{3B}$. Because the two firms are formally identical, it follows that $q_1^* = \frac{(A-c)}{3B}$ as well. Total output for this market is $Q^* = \frac{2(A-c)}{3B}$. Substituting this into the demand function gives the equilibrium price: $P = A - BQ = \frac{A+2c}{3}$. Profit for each firm is total revenue less total cost, which can be solved as $\pi_i = \frac{(A-c)^2}{9B}$.

As Figure 9.2 makes clear, the Cournot duopoly model just presented has a unique Nash equilibrium. Unfortunately, Cournot's original discussion somewhat obscures the underlying Nash insight that we emphasized earlier. In particular, Cournot's story suggests a kind of trial-by-error learning process by which the two firms act and react and thereby move along their Reaction Curves until the equilibrium is achieved. But the power of the Nash equilibrium is that it makes it unnecessary to play out such an iterative procedure in real time. If the entrant, firm 2, is rational and strategic then in choosing its own production level, it *must anticipate* that the incumbent, firm 1, will do whatever maximizes its profits. An expectation, for instance, by firm 2 that the incumbent will continue to produce the monopoly output $\frac{(A-c)}{2B}$ after firm 2 enters and produces $\frac{(A-c)}{4B}$ is not a rational expectation as that is not firm 1's best response to the production choice by firm 2. That is, firm 2 ought never to predict $q_1 = \frac{(A-c)}{2B}$. The only rationally prediction that firm 2 can possibly make is in fact that $q_1 = \frac{(A-c)}{3B}$ the value of q_1 in the Nash equilibrium. For if firm 2 predicts q_1 to be equal to $\frac{(A-c)}{3B}$, then firm 2 will *optimally* choose that output level, too. In turn, this output choice by firm 2 is such that firm 1 should indeed produce at the level of $\frac{(A-c)}{3B}$ if it wishes to maximize its profits. Of course, the same holds true for firm 1's expectations of firm 2's output. In other words, for each firm the only logically consistent expectation is that its rival will produce $\frac{(A-c)}{3B}$ in which case each firm will also choose to produce exactly that amount and thereby fulfill the expectation. Rationally strategic firms can work through the Cournot model as a pure thought experiment, and select the unique Nash equilibrium output $q_i^* = \frac{(A-c)}{3B}$ without any time-consuming real world trials and errors. For this reason, many economists, including the authors of this book, prefer to use the term "*best response function*" instead of "Reaction Curve." The point is to emphasize that the correct Nash interpretation of the Cournot model is not one of action and reaction but one in which equilibrium is achieved directly by each seller's recognition that its choices be logically consistent.[10]

As a numeric example, consider two firms, Untel and Cyrox, who supply the market for computer chips for toaster ovens. Untel's chips are perfect substitutes for Cyrox's chips and vice versa. Market demand for chips is estimated to be $P = 120 - 20Q$, where Q is the total quantity (in millions) of chips bought. Both firms have a constant marginal cost equal to 20 per unit of output. Untel and Cyrox independently choose what quantity of output to produce. The price then adjusts to clear the market of the total quantity of chips produced. What quantity of output will Untel produce? What quantity of output will Cyrox produce? What will be the price of computer chips and how much profit will each firm make?

Let's put ourselves on the management team at Untel to see the problem from its perspective. The demand curve that Untel faces can be written as $P = 120 - 20q_c - 20q_u$, where q_c is the output of Cyrox and q_u is the output of Untel. Untel's marginal revenue curve is $MRu = 120 - 20q_c - 40q_u$. To maximize profit, Untel chooses a quantity of output q_u^* such that its marginal revenue is equal to marginal cost. That is, $120 - 20q_c - 40q_u^* = 20$. This condition for profit maximization implies that:

$$q_u^* = \frac{120 - 20}{40} - \frac{20}{40}q_c \quad or \quad q_u^* = \frac{5}{2} - \frac{1}{2}q_c \tag{9.7}$$

This is Untel's best response function describing its optimal choice for any given level of output by Cyrox. In addition however, Untel knows that Cyrox is also a profit maximizer, and so Untel anticipates that Cyrox will want to produce q_c^* to satisfy the same best response condition for its profit maximization. That is, by precisely the same argument that we have just gone through, Untel knows that Cyrox's best response function is similarly: $q_c^* = \frac{120 - 20}{40} - \frac{20}{40}q_u$ *or* $q_c^* = \frac{5}{2} - \frac{1}{2}q_u$. Untel can recognize that Cyrox's choice of output depends on Untel's. Untel also knows that Cyrox knows that Untel is a profit-maximizer, and that Cyrox will anticipate that Untel will choose a profit-maximizing level of output q_u^*. Therefore, Untel predicts that Cyrox will choose $q_c^* = \frac{5}{2} - \frac{1}{2}q_u^*$. Substituting this prediction into Untel's best response function, equation (9.6), leads Untel to produce

$$q_u^* = \frac{5}{2} - \frac{1}{2}q_c^* = \frac{5}{2} - \frac{1}{2}\left(\frac{5}{2} - \frac{1}{2}q_u^*\right) \Rightarrow q_u^* = \frac{5}{3}$$

Now let's put ourselves on the management team at Cyrox and repeat the exercise. Because the two firms are *identical* there is no reason why Cyrox would do anything different from Untel, and so we can quickly jump to the conclusion that Cyrox will also produce $q_c^* = \frac{5}{3}$. Note that when Untel produces $\frac{5}{3}$, Cyrox's best response is to produce $q_c^* = \frac{5}{3}$, and similarly when Cyrox produces $\frac{5}{3}$, Untel's best response is to produce $q_u^* = \frac{5}{3}$. Aggregate market output therefore is $Q^* = \frac{10}{3}$, and so the price that clears the market

[10] Friedman (1977) includes a brief discussion of these issues, particularly valuable to those interested in the history of economic thought. He notes that Cournot's fate was not quite one of total obscurity owing to his friendship with the father of the French economist Walras. The English economist Marshall apparently was also well aware of and influenced by Cournot's work.

is $P^* = 120 - 20\left(\frac{10}{3}\right) = \$53.33$. For each firm the margin of price over unit cost is \$33.33 so that each firm makes a profit of \$55.55.

9.2 **Practice Problem**

Assume that there are two identical firms serving a market in which the inverse demand function is given by $P = 100 - 2Q$. The marginal costs of each firm are \$10 per unit. Calculate the Cournot equilibrium outputs for each firm, the product price, and the profits of each firm.

9.5 VARIATIONS ON THE COURNOT THEME: MANY FIRMS AND DIFFERENT COSTS

Cournot's model is insightful in its treatment of the interaction among firms and remarkably modern in its approach. Yet these are not its only strengths. Cournot's analysis has the further advantage that, as noted earlier, the results also blend well with economic intuition. In the simple Cournot duopoly model described above, each firm produces its Nash equilibrium output of $\frac{(A-c)}{3B}$, implying that total industry output is $\frac{2(A-c)}{3B}$. This is clearly greater than the monopoly output for the industry, which would be $Q^M = \frac{(A-c)}{2B}$. Yet it is also less than the perfectly competitive output, $Q^C = \frac{(A-c)}{B}$, where price equals marginal cost. Accordingly, the market-clearing price in Cournot's model $P = \frac{(A+2c)}{3}$ is less than the monopoly price $P^M = \frac{(A+c)}{2}$ but it is higher than the competitive price, c, which is equal to marginal cost. That is, Cournot's duopoly model has the intuitively plausible result that the interaction of two firms yields more industry output at a lower price than would occur under a monopoly, but less output at a higher price than results under perfect competition.

It is natural to ask if the foregoing result generalizes to the Cournot model when the number of firms grows to three, four, or N. That is, would introducing a third firm bring the industry still closer to the competitive ideal and, if so, would a fourth bring us closer still? Is the Cournot model consistent with the notion that when there are many firms the price converges to marginal cost?

The answer to this question is straightforward in the linear demand framework we have been using. To see this, assume that instead of two there are N identical firms, each producing the same homogenous good and each with the same, constant marginal cost c. Industry demand is again given by $P = A - BQ$ where Q is aggregate output. However, now we have that $Q = q_1 + q_2 + ... + q_N = \sum_{i=1}^{N} q_i$ so that $P = A - B\sum_{i=1}^{N} q_i$, where q_i is the output of the ith firm. In turn, this means that we can write the demand curve facing just a single firm, say firm 1, as: $P = (A - Bq_2 - Bq_3 - ... - Bq_N) - Bq_1$. The parenthetical expression reflects the fact that for firm 1, this sum is beyond its control and merely appears as the intercept in firm 1's demand curve. It is convenient to use the notation Q_{-1} as a shorthand method of denoting the sum of all industry output *except* that of firm 1's. Using this notation, we can write firm 1's demand curve even more simply as: $P = A - BQ_{-1} - Bq_1$. Clearly, firm 1's profits depend on both Q_{-1}, over which it has no control, and its own production level, q_1, which it is free to choose. Given its constant unit cost of c, firm 1's profits Π^1 can be written as: $\Pi^1(Q_{-1}, q_1) = (A - BQ_{-1} - Bq_1)q_1 - cq_1$. In turn, the twice-as-steep rule implies that firm 1's marginal revenue is given by $A - BQ_{-1} - 2Bq_1$. Hence, the condition $MR = MC$ necessary for profit maximization implies:

$$(A - BQ_{-1}) - 2Bq_1^* = c \tag{9.8}$$

Solving this equation for q_1^* gives us the *best response* function for firm 1:

$$q_1^* = \frac{(A-c)}{2B} - \frac{Q_{-1}}{2} \tag{9.9}$$

Because all firms are identical, we can extend this same logic to develop the best response function for any firm. Using the same shorthand notation, we can use Q_{-i} to mean the total industry production excluding that of firm i. The best response function for any firm i is:

$$q_i^* = \frac{(A-c)}{2B} - \frac{Q_{-i}}{2} \tag{9.10}$$

In a Nash equilibrium, each firm i chooses a best response, q_i^*, that reflects a correct prediction of the outputs that the other $N - 1$ firms will choose. Denote by Q_{-i}^* the sum of all the outputs excluding q_i^* when each element in that sum is *each firm's best output response decision*. Then an algebraic representation of the Nash equilibrium is:

$$q_i^* = \frac{(A-c)}{2B} - \frac{Q_{-i}^*}{2}; \quad \text{for } i = 1,2, \ldots N \tag{9.11}$$

However, since the N firms are identical, each will produce in equilibrium the *same* output, that is, $q_1^* = q_2^* = \ldots = q_N^*$, or just $q*$ for short. Hence, for any firm i, we must have $Q_{-i}^* = (N-1)q^*$. Therefore, we can write equation (9.11) as:

$$q^* = \frac{(A-c)}{2B} - \frac{(N-1)q^*}{2} \tag{9.12}$$

from which it follows that the equilibrium output for each firm in the Cournot-Nash equilibrium is:

$$q^* = \frac{(A-c)}{(N+1)B} \tag{9.13}$$

If each of the N firms produces q^* as given by equation (9.13), then we may easily derive the Cournot-Nash equilibrium industry output, $Q^* = Nq^*$, and the Cournot-Nash equilibrium industry price, $P^* = A - BQ^*$, as:

$$Q^* = \frac{N(A-c)}{(N+1)B}; \quad P^* = \frac{A}{(N+1)} + \frac{N}{(N+1)}c \tag{9.14}$$

Examine the two equations in (9.14) carefully. Recall that $\frac{(A-c)}{B}$ is the competitive industry output. Thus, for the linear demand and constant cost case, the N -firm Cournot model industry output is $\left(\frac{N}{N+1}\right)$ times the competitive output. This rule, sometimes called the $\left(\frac{N}{N+1}\right)$ rule predicts (correctly) that the monopoly output is ½ of the competitive output. As we move to a duopoly, the rule says that industry output rises to $2/3$ of the competitive output. At three firms, the proportion rises to $3/4$ and so on. In other words, our earlier result does generalize. The Cournot market output rises progressively closer to the competitive output as the number of firms N rises. Likewise close inspection of the price equation (9.14) reveals that the equilibrium Cournot price correspondingly moves progressively closer to the competitive equilibrium price $P = c$ as N rises.

In short, the Cournot model implies that as the number of identical firms in the market grows, the industry equilibrium gets closer and closer to that prevailing under perfect competition. That is, for this symmetric case, the Cournot model has the appealing feature that it predicts market outcomes will improve as market concentration falls.

What if the firms competing in the market are not identical? Specifically, what if each firm has a different marginal cost? We first handle this question for the case of two firms. Assume that the marginal costs of firm 1 are c_1 and of firm 2 are c_2. We use the same approach as before with the duopoly model, starting with the demand function for firm 1, which we can write as:

$$P = (A - Bq_2) - Bq_1$$

The associated marginal revenue function is $MR_1 = (A - Bq_2) - 2Bq_1$.

As before, firm 1 maximizes profit by equating marginal revenue with marginal cost. Thus setting $MR_1 = c_1$ and solving for q_1 gives the best response function for firm 1 as:

$$q_1^* = \frac{(A - c_1)}{2B} - \frac{q_2}{2} \tag{9.15a}$$

By an exactly symmetric argument, the best response function for firm 2 is:

$$q_2^* = \frac{(A - c_2)}{2B} - \frac{q_1}{2} \tag{9.15b}$$

Notice that the only difference from our initial analysis of the Cournot model is that now each firm's best response function reflects its own specific marginal cost.

An important feature of these best response functions that is obscured when the firms are identical is that the *position* of each firm's best response function is affected by its marginal cost. For example, if the marginal cost of firm 2 increases from say, c_2 to c_2', its best response curve will shift inwards.

Figure 9.3 illustrates this point. It shows the best response function for each firm assuming initially that each firm has identical costs as in Figure 9.2. It then shows what happens when firm 2's unit cost rises. As equation (9.15b) makes clear, this cost increase *lowers* firm 2's best output response for any given level of q_1. That is, it shifts firm 2's best response curve inward. This change in firm 2's best response function affects the equilibrium outputs that the two firms will choose. As you can see from the diagram, an increase in firm 2's

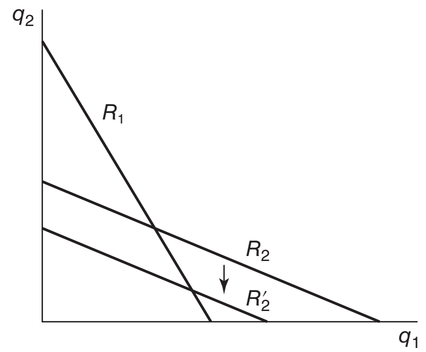

Figure 9.3 The Cournot duopoly model with different costs across firms
A rise in Firm 2's unit cost shifts the Firm 2 Best Response Function downward from R_2 to R_2'. In the new equilibrium, Firm 1 produces more and Firm 2 produces less than previously.

marginal cost leads to a new equilibrium in which firm 1 produces more than it did in the initial equilibrium and firm 2 produces less. This makes intuitive sense. We should expect that low-cost firms will generally produce more than high-cost firms. The changes are not offsetting, however. Firm 2's output falls by more than firm 1's production rises so that the new equilibrium is characterized by less output in total than was the original equilibrium. (Can you say why?)

The Cournot-Nash equilibrium can be obtained as before by substituting the expression for q_2^* into firm 1's best response to solve for q_1^*. Then we may use this value to solve for q_2^*. In other words, we have:

$$q_1^* = \frac{(A - c_1)}{2B} - \frac{1}{2}\left(\frac{(A - c_2)}{2B} - \frac{q_1^*}{2}\right)$$

which can be solved for q_1 to give the equilibrium:

$$q_1^* = \frac{(A + c_2 - 2c_1)}{3B} \tag{9.16a}$$

By an exactly symmetric argument, the equilibrium output for firm 2 is:

$$q_2^* = \frac{(A + c_1 - 2c_2)}{3B} \tag{9.16b}$$

It is easy to check that the relative outputs of these two firms are determined by the relative magnitudes of their marginal costs. The firm with the lower marginal costs will have the higher output.

Let's return to our Untel and Cyrox example of the two firms who produce computer chips for toaster ovens but now change this story a bit. While we still assume that Untel's chips are perfect substitutes for Cyrox's chips and vice versa, we no longer assume that they have identical costs. Instead, we now assume that Untel is the low-cost firm with a constant unit cost of 20, and Cyrox is the high-cost producer with a constant unit cost of 40. Market demand for chips is still estimated to be $P = 120 - 20Q$, where Q is the total quantity (in millions) of chips bought. What now happens when Untel and Cyrox independently choose the quantity of output to produce? What quantity of output will Untel produce? What quantity of output will Cyrox produce?

Again we put ourselves on the management team at Untel to see the problem from Untel's perspective. The demand curve that Untel faces is still $P = 120 - 20q_c - 20q_u$, where q_c is the output of Cyrox and q_u is the output of Untel. Untel's marginal revenue is again $MR_u = 120 - 20q_c - 40q_u$. To maximize profit, Untel should sell a quantity of output q_u^* such that at that quantity marginal revenue is equal to marginal cost. That is, $120 - 20q_c - 40q_u^* = 20$, and so the condition for profit maximization implies that:

$$q_u^* = \frac{5}{2} - \frac{1}{2}q_c \tag{9.17}$$

Untel's profit-maximizing choice of output still depends on the output that the higher cost rival, Cyrox, chooses to produce. Equally importantly, a comparison of equation (9.16) with (9.6) indicates that Untel's best response function is unaffected by the assumed increase in Cryox's marginal cost. What about Cryox? By the same argument, the demand curve that Cryox faces is $P = 120 - 20q_u - 20q_c$ and its marginal revenue curve is $MR_c = 120 - 20q_u - 40q_c$. Equating this with marginal cost of 40 and solving for q_c gives the best response function for Cryox of $q_c^* = \dfrac{120-40}{40} - \dfrac{20}{40}q_u$ *or* $q_c^* = 2 - \dfrac{1}{2}q_u$. As we expected, the best response function for Cryox is shifted downward by the assumed increase in its marginal cost.

Untel knows that higher cost Cyrox is also a profit maximizer and therefore anticipates that Cyrox will want to produce q_c^* that maximizes its profit. It is also the case, as it was before, that Untel knows that Cyrox knows that Untel is a profit-maximizer, and so knows that Cyrox will anticipate that Untel will choose a profit-maximizing level of output q_u^*. All of this implies that Untel predicts that Cyrox will choose $q_c^* = 2 - \dfrac{1}{2}q_u^*$. Substituting this new prediction into Untel's best response function leads Untel to produce $q_u^* = \dfrac{5}{2} - \dfrac{1}{2}q_c^* = \dfrac{5}{2} - \dfrac{1}{2}\left(2 - \dfrac{1}{2}q_u^*\right) \Rightarrow q_u^* = 2$.

Now we put ourselves on the management team at Cyrox and repeat the exercise. To cut to the chase, we know that Cyrox's best response is $q_c^* = 2 - \dfrac{1}{2}q_u$. Moreover we know that Cyrox will predict that Untel will produce a best response that is based on a prediction that Cyrox will also produce a best response. That is, Cryrox predicts that Untel will produce $q_u^* = \dfrac{5}{2} - \dfrac{1}{2}q_c^*$. Substituting this prediction into Cyrox's best response function leads to: $q_c^* = 2 - \dfrac{1}{2}q_u^* = 2 - \dfrac{1}{2}\left(\dfrac{5}{2} - \dfrac{1}{2}q_c^*\right) \Rightarrow q_c^* = 1$. Again note that when Untel produces 2, Cyrox's best response is to produce $q_c^* = 1$, and similarly when Cyrox produces 1, Untel's best response is to produce $q_u^* = 2$.

Although the foregoing analysis is limited to just two firms, it still yields important insights. First, there is a disadvantage to being a high-cost firm. It means a smaller market share and less profit. Second, while the disadvantage is real it is not fatal. A high-cost firm is not driven out of business, It is simply less successful than its ost rival. Finally, when costs vary across firms, the equilibrium Cournot output Q^* is not only too low (i.e., less than the competitive level), it is also produced inefficiently. As we know from

Chapter 4, efficient production among two or more firms would allocate output such that, in the final configuration, each firm's marginal cost is the same. This would be the outcome, for example, if the industry were comprised of a single, profit maximizing multiplant monopolist. It will also obtain under perfect competition. However, the non-cooperative feature of the Cournot-Nash equilibrium means that firms' marginal costs are not equalized.[11] Hence, the output allocation at equilibrium when costs are different is not an efficient one.

9.3 Practice Problem

What is aggregate output, market price, Untel's profit, and Cyrox's profit for the above case in which Untel is the low-cost producer and Cyrox the high-cost one? Compare your answers to the ones you work out when the two firms are identical and have a constant unit cost of 20.

9.6 CONCENTRATION AND PROFITABILITY IN THE COURNOT MODEL

Let's now try to combine the case of many firms together with the assumption of *non*-identical costs. That is, let's analyze the Cournot model with N firms, each with its own (constant) marginal cost such that the marginal cost of firm i is c_i. We can use the first order condition for profit-maximization for each firm i, equation (9.8), and substitute c_i for c in this equation. This gives us the following:

$$A - BQ_{-i} - 2Bq_i^* - c_i = 0 \tag{9.18}$$

where Q_{-i} again is shorthand for the industry production accounted for by all firms other than the ith one.

In a Nash equilibrium, the equilibrium output q_i^* *for each firm* i must satisfy the first-order profit-maximizing condition. Hence, in the Nash equilibrium, the term Q_{-i} must be the sum of the *optimal* outputs q_j^* for each of the "not i" firms. Denote this equilibrium sum as Q_{-i}^*. Then we can rewrite (9.18) as:

$$A - BQ_{-i}^* - 2Bq_i^* - c_i = 0 \tag{9.19}$$

By definition, the total equilibrium output, Q^*, equals the sum of Q_{-i}^* and q_i^*. Hence, (9.19) implies that

$$A - B(Q^* - q_i^*) - 2Bq_i^* - c_i = 0$$

which can be reorganized to give:

$$A - BQ^* - c_i = Bq_i^* \tag{9.20}$$

[11] Our example assumed constant but different marginal costs across firms. The same insight could be easily obtained for the more general presentation in which the marginal cost of firm i, c_i, is a general function of its output, q_i, as in $c_i = c_i(q_i)$.

We also know that the Nash equilibrium price, P^*, is obtained by substituting the Nash equilibrium output into the industry demand curve yielding, $P^* = A - BQ^*$. Substitution into equation (9.20) then yields:

$$P^* - c_i = Bq_i^* \tag{9.21}$$

Dividing both sides of equation (9.21) by P^* and multiplying the right-hand side by $\frac{Q^*}{Q^*}$, we obtain:

$$\frac{P^* - c_i}{P^*} = \frac{BQ^*}{P^*} s_i^* \tag{9.22}$$

where $s_i^* = \frac{q_i^*}{Q^*}$ is the ith firm's market share in equilibrium.

Let's consider equation (9.22) step-by-step. The left-hand-side term is the difference between the price and firm i's marginal cost as a proportion of market price. This is just the *Lerner Index of Monopoly Power* that we met in Chapter 3. The notion is that the greater firm i's market power, the greater its ability to keep price above marginal cost.

The right-hand-side of (9.22) has two terms. The first is the slope of industry demand curve times the ratio of industry output to price. But the slope is just $B = -\Delta P/\Delta Q$ so that we have $\frac{BQ^*}{P^*} = -\frac{\Delta P}{\Delta Q}.\frac{Q^*}{P^*}$. Recall the definition of the price elasticity of demand: $\eta = -\frac{\Delta Q}{\Delta P}.\frac{P}{Q}$. So the first term on the right-hand side of equation (9.22) is just the inverse of the price elasticity of demand. The second term is just the market share of the ith firm, that is, its output relative to total industry output. Hence, equation (9.22) may be rewritten as:

$$\frac{P^* - c_i}{P^*} = \frac{s_i^*}{\eta} \tag{9.23}$$

where η is the price elasticity of industry demand.

Equation (9.23) is a further implication of the Cournot model, now extended to allow for many firms with differing costs. What it says is this: A firm that produces in an industry where demand is relatively inelastic and where it has a relatively large market share will also be a firm with a substantial degree of market power as measured by the Lerner Index or the firm's price-marginal cost distortion.

The relationship described in equation (9.23) tells us about market power at the level of the firm. In Chapter 3, we discussed the structure-conduct-performance (SCP) paradigm in industrial organization that linked market power, as measured by the Lerner Index, to the structure of the *industry*. The question that remains is whether we can extend the relationship in equation (9.23) at the firm level to the level of the entire industry.

To see how, let's first multiply each side of equation (9.23) by the firm's market share, s_i^*. Then add together the modified equation for firm 1 with that for firm 2 and that for firm 3 and so on until we add together all N modified (9.23) equations. The left-hand side of this sum of N equations is:

$$\sum_{i=1}^{N} s_i^* \left(\frac{P^* - c_i}{P^*}\right) = \frac{\left(\sum_{i=1}^{N} s_i^* P^* - \sum_{i=1}^{N} s_i^* c_i\right)}{P^*} = \frac{P^* - \overline{c}}{P^*}$$

Reality Checkpoint

Cournot Theory and Public Policy: The 1982 Merger Guidelines

In our review of antitrust policy in Chapter 1, we noted the dramatic change in policy regarding the treatment of mergers that occurred in 1982. In that year, the Department of Justice issued a new version of its *Horizontal Merger Guidelines*. This version replaced the original guidelines issued in 1968. Like that first set of guidelines, the 1982 document specified the conditions under which the government would challenge horizontal mergers. Unlike their predecessor, however, the new guidelines were based explicitly on the Herfindahl Index. Specifically, they stated that a merger would not be challenged if the industry Herfindahl Index was less than 1,000. A merger would also not be challenged if the index was over 1,000 but less than 1,800 *and* if the merger did not raise the Herfindahl Index by over 100 points. If the Herfindahl Index exceeded 1,800 points, then any merger that raised the index by over 50 points would cause concern and likely be challenged.

We will discuss these guidelines and their more recent modifications again in Chapter 15. For now, the point to note is that the explicit use of the Herfindahl Index may be viewed as a bow to the Cournot model which, as shown in the text, directly connects that index to the price-cost margin measure of monopoly power.

Source: Department of Justice, *Horizontal Merger Guidelines* (1982, 1984).

where $\overline{c}$ is the weighted average unit cost of production, the weights being the market shares of the firms in the industry. The right-hand side of the summed N equations is:

$$\frac{\sum_{i=1}^{N}(s_i^*)^2}{\eta} = \frac{HI}{\eta}$$

where HI is the Herfindahl Index that we defined as a measure of concentration in Chapter 4 (here expressed using fractional shares, e.g., a 10 percent share is recorded as $s_i = 0.10$). Therefore equation (9.23) aggregated at the level of the industry implies that:

$$\frac{(P^* - \overline{c})}{P^*} = \frac{HI}{\eta} \tag{9.24}$$

Our generalized Cournot model thus gives theoretical support for the view that as concentration (here measured by the industry's Herfindahl Index) increases, prices also rise further and further above marginal cost.

A variant of the relationship in equation (9.24) was tested in Marion et al. (1979) for food products. They collected price data for a basket of ninety-four grocery products, and market share data for thirty-six firms operating in thirty-two US Standard Metropolitan Statistical Areas, and found that price is significantly higher in markets with a higher Herfindahl Index. Likewise, Marvel (1989) found that for twenty-two US cities, concentration in the retail market for gasoline, as measured by the Herfindahl Index, had a significant impact on the average price of gasoline.

9.7 EMPIRICAL APPLICATION: COURNOT COMPETITION IN AN AIRLINE DUOPOLY

The examples above suggest that at least some of the implications of the Cournot model are consistent with real world data. However, such results might also be compatible with other models as well that differ in an important respect from the precise assumptions of the Cournot model presented in this chapter. One such difference concerns the Cournot firm's understanding of how rivals will react to its output choice.

Consider again the basic duopoly model with a linear demand curve: $P = A - BQ$. As before, assume as well that each firm has a constant marginal cost c_i. As we described above, the industry demand perceived by each firm is described by $P = A - Bq_1 - Bq_2$. Consider firm 1. The Cournot model hypothesis is that in calculating the impact of changes in its output, i.e., in calculating its marginal revenue, firm 1 assumes that firm 2's output is unchanged. Hence, as in equation (9.3), the profit-maximizing output for firm 1 satisfies $\text{MR}_1 = A - Bq_2 - 2Bq_1 = c$. However, we want to write this equation slightly differently to make a key point of the Cournot model. Recognizing that $A - Bq_1 - Bq_2 = P$, we can rewrite firm 1's profit-maximizing condition as follows:

$$P - q_1 B = c \tag{9.25}$$

The term $B = -\left(\dfrac{\Delta P}{\Delta Q}\right)$ is of course the slope term that simply characterizes the (negative) effect on the industry price of changes in industry output. Its presence reflects the fact that when firm 1 makes a small marginal increase in its own output, $\Delta q_{1,}$ the Cournot assumption that q_2 does not change implies that $\Delta Q = \Delta q_1$ so that the resultant fall in price is B. Multiplying this over firm 1's current output then captures the downside of further increases in q_1. Similarly, the Cournot model assumes firm 2 also expects no reaction to a small change in its output q_2 so that it perceives $\Delta Q = \Delta q_2$.

It is possible, however, that instead of no response each firm anticipates that its rival will adjust its output as the initial firm changes its own production. That is, firm 1 may believe that $\Delta q_2 = v\Delta q_1$ so that when firm 1 changes its output, the total change in industry output is now $\Delta Q = (1 + v)\ \Delta q_1$. Likewise, firm 2 may believe that $\Delta Q = (1 + v)\ \Delta q_2$. As a result, the movement in price from a small increase at the margin in firm 1's output will be $B(1 + v)$. Hence, the profit-maximizing first-order condition in equation (9.25) now becomes:

$$P - q_1 B(1 + v) = c \tag{9.26a}$$

Symmetrically, firm 2's profit maximizing condition is now:

$$P - q_2 B(1 + v) = c \tag{9.26b}$$

Because the parameter v reflects the firm's conjecture regarding the output response their own choices will induce, it is often referred to as the conjectural variations parameter. With a little intuition, we can use equations (9.26a) and (9.26b) to put some bounds on the values that v can take. In this regard, observe that if $v = -1$, the two equations imply $P = c$. That is, $v = -1$ corresponds to perfect competitive behavior in which price is driven to marginal cost. To see the implication if $v = 1$, recall from Chapter 2 that the monopoly output is $Q^M = \dfrac{A - c}{2B}$, so that the monopoly price is $P^M = \dfrac{A + c}{2}$. Assume that each firm produces half the monopoly output as would be the case if the firms acted like a cartel. Hence, $q_1 = q_2 = \dfrac{A - c}{4B}$. Then our adjusted equations each yield $P = \dfrac{A + c}{2}$, i.e., the monopoly price. Thus, our conjectural variations parameter must lie between a lower bound of $v = -1$ reflecting perfectly competitive behavior and an upper bound of $v = +1$ reflecting cartel or monopoly behavior. The intermediate value of $v = 0$ reflects Cournot behavior.

Table 9.4 Estimated conjectural variation parameter *v* for American airlines and United airlines in 33 city markets paired with Chicago

Pair City	*American v Estimate*		*United v Estimate*
Grand Rapids	0.82		1.02
Indianapolis	−0.11		1.95
Columbus	1.46		0.41
Des Moines	1.50		0.34
Omaha	1.74		0.02
Buffalo	0.39		0.35
Rochester	0.39		0.81
Tulsa	0.00		1.08
Wichita	0.25		0.90
Syracuse	0.19		0.59
Baltimore	0.62		0.16
Oklahoma	0.17		0.82
Albany	−0.10		0.82
New York	0.20		0.48
Charleston	−0.50		−0.62
Hartford	0.79		0.01
Dallas	−0.39		1.33
Providence	−0.14		0.28
Austin	−0.72		−1.43
San Antonio	−0.13		−0.20
Albuquerque	−0.78		−0.31
Phoenix	−0.08		−0.70
Tucson	−0.79		−0.17
Las Vegas	−0.75		−1.15
Reno	−0.80		−0.77
Ontario, CA	−0.28		−0.49
San Diego	−0.28		−0.58
Seattle	−0.11		−0.32
Los Angeles	−0.16		0.02
Portland	−0.84		−0.35
Sacramento	−0.08		−0.31
San Jose	0.33		0.04
San Francisco	0.26		−0.20
Mean	**0.06**		**0.23**
Overall Mean		**0.09**	

Let us now allow for the two firms to have different marginal costs c_i and therefore different market shares s_i. Then recognizing that firm i's output $q_i = s_i Q$ we may generalize equations (9.25a) and (9.25b) as follows:

$$P - s_i QB(1 + v_i) = c_i;\ i = 1,2 \tag{9.27}$$

Solving for v_i and recognizing that $B = -\left(\frac{\Delta P}{\Delta Q}\right)$ so that $\frac{1}{BQ} = \frac{\eta}{P}$, where η is the elasticity of demand (expressed as a positive number), equation (9.27) may be rewritten as:

$$v_i = \left(\frac{P - c_i}{P}\right)\frac{\eta}{s_i} - 1 \tag{9.28}$$

Brander and Zhang (1990) use data from thirty-three airline routes emanating from Chicago in the 1980s that were served by just the two carriers, American Airlines and United Airlines. Their baseline case uses an assumed elasticity of $\eta = 1.6$ (based on other research) and marginal cost estimates derived from the operating expenses for each firm (including an allowance for marginal cost to vary with the length of the flight with an elasticity of 0.5). Their calculated v_i values for each firm in each market are shown in Table 9.4 above.

The rough calculation behind these measures implies that they will include a fair bit of noise. Even so, the overwhelming majority or 57 of the 66 estimated v_i parameters lie within the theoretical upper and lower bounds of 1 and -1. More importantly, the estimated mean is very close to 0 for both firms and across the entire sample. Classical statistical tests would reject the null hypothesis that v equals either 1 or -1 They would not reject the Cournot hypothesis that $v = 0$. Thus, this rough exercise gives fairly strong support to the Cournot model as a useful description of real world behavior.

Summary

For industries populated by a relatively small number of firms, strategic interaction is a fact of life. Each firm is aware of the fact that its decisions have a significant impact on its rivals. Each firm will want to take account of the anticipated response of its rivals when determining its course of action. It is reasonable to believe that firms' anticipations or expectations are rational.

Game theory is the modern formal technique for studying rational strategic interaction. Each player in a game has a set of strategies to choose from. A strategy combination is a set of strategies—one for each player. Each such strategy combination implies a particular payoff or final outcome for each player. A Nash equilibrium is a strategy combination such that each player is maximizing his or her own payoff *given* the strategies chosen by all other players. In a Nash equilibrium, no player has an incentive to change behaviors unilaterally.

In this chapter, we presented the well-known Cournot model of competition. It is a static or single market period model of oligopoly. Although this model was developed prior to the development of formal game theory, the outcome proposed by Cournot captures basic game theoretic principles, specifically the Nash equilibrium solution.

The Cournot model makes clear the importance of firms recognizing and understanding their interdependence. The model also has the nice intuitive implication that the degree of departure from competitive pricing may be directly linked to the structure of the industry as measured by the Herfindahl Index. While the Cournot model is not the only game-theoretic model of