- Companies earn excess returns when they anticipate market changes and rapidly exploit new opportunities.
- Companies earn excess returns when they possess a sustainable competitive advantage over potential and actual rivals.

Most economists working in the strategy area in the 1970s focused their attention on the first source of advantage. All of the work by Porter and his followers in the industrial organization area looked to industry forces and particularly entry conditions as the main source of excess returns. Work in the general management literature, on the other hand, has traditionally looked to the latter two sources of advantage. Increasingly, firms, too, have focused on competitive advantage and particularly speed as a way to outperform the market. Jack Welch, current CEO of General Electric, writing in 1991, emphasized the importance of speed in corporations as a way "to propel ideas and drive processes right through functional barriers, sweeping bureaucrats and their impediment aside in the rush to get to the marketplace."[12] In markets such as the ones General Electric competes in, where the overall level of competition is high, time to market is critical. In these markets, one of the goals of strategy is to help managers "imagine the future"[13] as a way to exploit it. So, too, in the battle for markets, is it critical for firms to develop and sustain an edge over real and potential competitors and, when possible, to work to change the environments in which they operate to favor their own core competencies. In 1991, Robert Allen, CEO of AT&T, focused on AT&T's research and development as the source of their "powerful competitive advantage," the advantage which would deliver returns even in the face of substantial competition. In firms throughout the world, we see managers struggling to identify the right strategic path to superior performance.

But in this struggle for profitability and growth, the power of new entry and the lure of the $20 bill serves as a constant backdrop. Managers must constantly ask themselves in testing out alternative strategies the fundamental market question: What protects my strategy against encroaching entry and imitation by existing rivals? And, if I am not protected, and imitation does occur, what can I do to maintain good performance in the new era? We turn in the next several chapters to these questions.

3

Industry Analysis

As we saw in Chapter 2, the natural trajectory of a market is towards a point at which the risk-adjusted returns to that market are the same as those earned in the market as a whole. Nevertheless, in some markets, over some periods of time, most of the firms in the industry seem to do better than the market as a whole. In these markets, the overall environment is less hostile than average. This is the first source of superior performance outlined in Chapter 2. In this chapter we explore some of the reasons for industry-wide high profitability and the factors that contribute to a more forgiving environment. An analysis of these forces forms an essential part of the environment analysis task of strategic planning.

PORTER'S FIVE FORCES MODEL OF INDUSTRY

One way to organize information about an industry that shows us the potential attractiveness of that industry is the **Five Forces Model** developed by Michael Porter and reproduced as Figure 3.1.[1] Taken together, the five forces shown in the figure help to explain the overall level of profitability one might expect in a given industry. We look to these forces to help explain why one industry is profitable while another is not; why, for example, the disposable diaper industry is highly profitable, while the pulp and paper industry is not. Porter's model suggests that, to a large extent, these industry differences can be explained by five factors: the current intensity of competition, the presence of substitute products, the power of buyers, the power of suppliers, and new entry. In this chapter, we explore the first four of these factors. Since entry is so vital to understanding industry dynamics, as we have already seen in Chapter 2, we will leave that topic for separate coverage in Chapter 4.

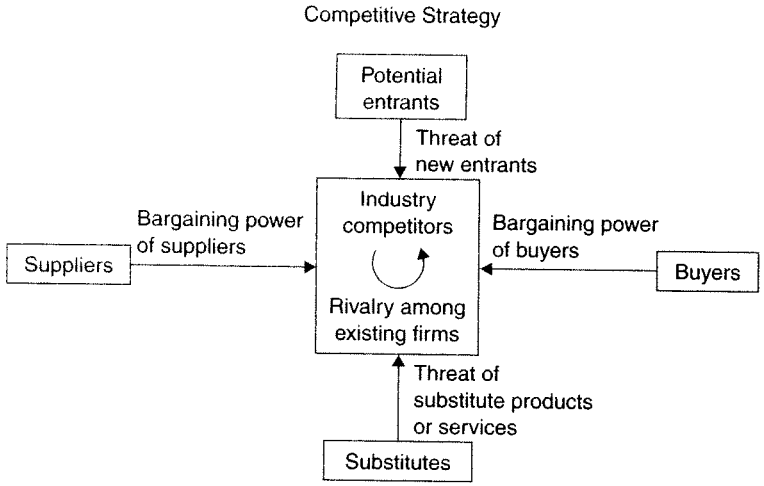

Figure 3.1 Porter's Five Forces Model.

Porter Force I: Intensity of Competition

We begin in the industry analysis by looking at existing relations within the industry or the level of rivalry among firms. This is represented by the center box in the diagram. In 1997, the branded disposable diaper market was dominated by two firms, Procter and Gamble (P&G) and Kimberly Clark (KC). These two firms controlled approximately 80 percent of the market. But the market shares of these two firms were extremely unstable, as was the relationship between them as they warred via research and marketing. In 1982, for example, P&G controlled 65 percent of this $2-billion-a-year market with Luvs and Pampers, while KC, with its Huggie brand, had a modest share. By 1985, P&G's share had dropped to 52 percent and most of the difference had been taken up by KC. In 1985, in response to its market share decline, P&G introduced a new "thin" version of diapers and its share again rose. By 1986, KC had developed its own thin diaper and was recovering market share. In the early 1990s, KC continued to innovate with training pants, overnight diapers, and the like, becoming the dominant player in the market. In 1997, P&G struck back with a lotion-added diaper, hoping to recover market share. Product innovations proliferate in this market and market shares respond, creating considerable instability in the competitive positions of the two industry giants. Moreover, despite the large size of the market and the relatively lucrative margins, a number of large firms—Scott Paper, Union Carbide, and Johnson and Johnson—had been forced to exit with large losses from the industry since the mid-1970s. In 1992, a consortium of Weyerhauser, Johnson and Johnson, and Unicharm, a large Japanese firm, also failed in their bid to sell a premium disposable diaper. By almost any measure, the rivalry in the disposable-diaper business was and is intense.

In contrast, the educational book publishing market is a highly *fragmented* one. Approximately 60 to 70 major publishers compete in the elementary and secondary school market, with the top five in the industry accounting for only 40 percent of the sales in the industry. All of the major players in the market—McGraw Hill, Simon and Schuster (Prentice Hall), Harcourt Brace Jovonovich, Macmillan—have been important influences in the market for decades. Market shares are reasonably stable, and exit of major players has been rare, though many of the once-independent firms are now part of larger corporate families. Editors from the various publishing houses, as well as many of the people on the "business" side of the operations, frequently socialize with one another and larger publishers support and encourage smaller operations. Consider the following observation from Winthrop Knowlton, former president of Harper and Row: "Small presses often come to us for advice and help and sometimes marketing services. We are happy to oblige because we cannot cover the editorial waterfront ourselves."[2]

By most measures, rivalry in the book publishing industry is modest and cooperation among the staff of various publishing houses fairly common.

What accounts for differences in the levels of rivalry among industries? The question is clearly of interest both to potential new entrants in an industry, who are seeking to understand what lies ahead, and to current market participants involved in changing environments, for differences in the levels and variance of profits are explained in part by differences in levels of rivalry.

Intense rivalry among firms in an industry reduces average profitability.

It is easy to demonstrate that in any given industry, some measures of coordination is to the *collective* good of industry participants. Total profits in an industry are greater with coordination than without. In the pricing area, for example, perfect coordination would allow industry participants to avoid costly price wars, and in the extreme even to practice monopoly pricing as a group. In the area of branding, coordination would allow firms to avoid duplicating each other's products. In the area of research, coordination would permit firms to avoid costly duplicate research efforts.[3] Nevertheless, while the collective interest is served by coordination, the individual self-interests of the firms in the industry may not be perfectly consonant.[4] A few examples will help to clarify the seeming paradox.

Suppose we start with an industry characterized by pricing coordination among firms and above-normal profits. That is, in this industry, prices exceed long-run marginal costs. Collectively, we would say that the indus-

try is doing very well. Now let us take the perspective of one of the firms within that industry, perhaps one of the smaller firms. This firm, too, is doing well. On the other hand, the fact that prices exceed marginal costs is quite seductive to the firm. By lowering prices just a little, the firm can increase its market share and increase its own profits, since the incremental production can be sold at prices in excess of their costs. The firm may well believe that it can cut prices in this way without upsetting the overall industry pricing consensus. But, of course, if enough firms yield to this temptation, the market price will drift down, and industry coordination will be threatened. Similar stories can be told about the advantages to the individual firm of shading just a bit on other areas of industry coordination.

In an industry in which firm coordination yields excess profits, there are incentives for individual firms to shade prices.

We see then that there is a tension within an industry. Coordination increases the collective good, but individual cheating on that coordination may increase the individual good, as long as not too many other firms cheat at the same time. One way of analyzing the level of rivalry in an industry is to focus on the costs and benefits of coordination versus cheating from the perspective of the industry participants. This is the perspective we will adopt in our discussion.

From this perspective, we find that there are a number of characteristics of an industry that help determine the level of rivalry in the industry and that should be considered when we look at box one of the Porter model. Principal among them are the number of competitors, the size distribution of those competitors, the homogeneity of competitors, the fixed nature of investment in the industry, and the stability of demand in the industry.

The Number of Competitors

When the number of competitors in an industry is large, all else equal, we expect more competition in the industry. Under these circumstances, each firm believes itself to be only a minor player in the industry, and thus will act as if its price cuts or marketing investments will have only a small effect on the industry as a whole. Thus, firms tend to act more individualistically when there are many firms in the market. Moreover, under these conditions, coordination among the players—even if sought—would be quite difficult to consummate owing to sheer numbers. The larger the number of firms in an industry, the greater the uncertainty there is likely to be about relative costs and other operating factors of those firms, and this, too, complicates coordination.[5]

Large numbers of firms in a market reduce coordination opportunities.

Interestingly, in industries with many players, there is often considerable competition but very little direct rivalry among firms. In dairy farming, for example, competition is intense in the sense that any individual farmer would find it difficult to charge a price higher than that of his neighbors, since the individual output of any single farm is small relative to the total market. And yet the rivalry in the dairy industry is quite impersonal. Neighboring farmers, for example, will often cooperate in the lending of capital equipment, R&D expenditures, and so on. In other industries—like computers—rivalry is both intense and personal. Netscape thinks of Microsoft as a rival and devises its strategy accordingly.

A commonly used measure of the number and relative power of firms in an industry is given by the **concentration ratio.** In the United States, the concentration ratio is generally calculated as the percent of the total industry sales or employment accounted for by the largest four firms in that industry. Higher figures represent more concentration. Data on the C4, as it is known, is calculated by the United States Census and printed in *Census of Manufacturing.* Industries vary widely in their concentration. Logging, for example, is a relatively unconcentrated industry with a C4 of 18 percent, while cigarette manufacturing is relatively concentrated with a C4 of 85 percent. We should note that the concentration ratio is keyed to the position of the largest firms in the industry, given that it measures the combined share of the top four firms in an industry. For some situations, it is also useful to know whether any appreciable fringe firms exist, since fringe firms can sometimes have a big effect on the overall stability of an industry.

We have suggested that as the number of firms in an industry grows, opportunities for coordination diminish. Some flavor for this relationship can be seen if we look at the history of antitrust litigation. Greer and Fraas gathered data on the 606 illegal price-fixing agreements prosecuted in the period between 1910 and 1973. They then determined the number of firms typically involved in the industry conspiracies and computed a frequency distribution.[6] Their results are given in Figure 3.2. As we can easily see, the

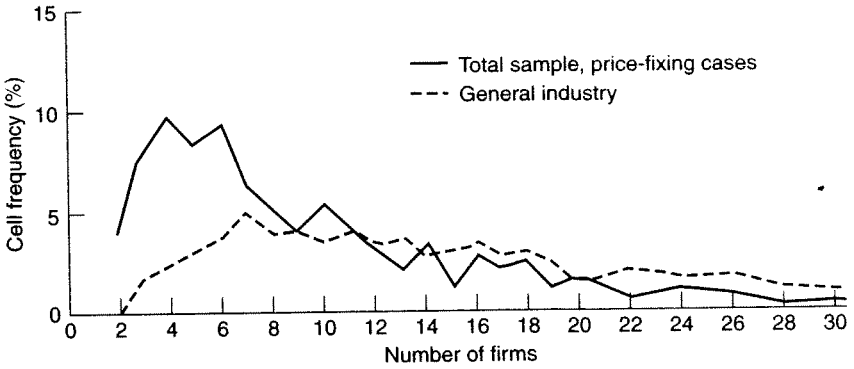

Figure 3.2

typical price conspiracy prosecuted by the government involves few firms relative to the average industry. About 10 percent of the price-fixing cases, for example, involved industries with only four firms, while in toto only about 2 percent of United States industries contain only four firms. This numerical pattern may in part reflect government interest in concentrated industries rather than actual price-fixing behavior, but the results are nevertheless interesting and broadly consistent with the general theme of small numbers providing more opportunities for cooperation.

The Size Distribution of Market Participants

The size distribution of firms within an industry may be very important in determining the intensity of rivalry. Are the top three or four firms in the industry similar in terms of market share or quite different?

In general, in industries in which the major firms are all similarly sized, rivalry is more intense.

In industries in which one large firm dominates the market, that firm can often impose a certain discipline on the rest of its market, for its relative size allows it to speak with some authority. In industries in which several firms are equally sized, more rivalry develops as each vies for the leadership position.

An interesting example of the intense rivalry that can occur among equal-sized rivals is provided by the European automobile market. There are seven auto makers in Europe, with market shares in the range of 10 to 15 percent, among them Volkswagen, Fiat, and Ford. As trade barriers across Europe crumble further, most observers believe the level of competition among these balanced rivals will be substantial.

A convenient measure of the balance in an industry is provided by the **Herfindahl index (HI),** which is used by government to help decide when mergers among industry participants should be allowed. The Herfindahl is equal to:

$$HI = 10{,}000\Sigma_i Si^2$$

where S_i is the market share of the *i*th firm. A simple example will be helpful to show us how to calculate the HI. Suppose we have an industry with three firms with market shares as follows: .5, .25, .25. The HI would equal $(10{,}000)((.5)^2 + (.25)^2 + (.25)^2)$ or 3750. If the industry was configured .33, .33, .33, it would have an HI of 3267. The unbalanced size distribution is associated with a higher index. With an unbalanced distribution and thus a high index, the market leader can help promote coordination.

The Herfindahl index can vary from 0 (perfectly competitive industry) to 10,000 (a monopoly). Generally speaking, HIs in excess of 1,800 are

thought to characterize industries with reduced rivalry. Indeed, at HIs of this level, the Department of Justice investigates merger activity fairly carefully, since it believes that indexes in this range are associated with markets in which collusion is more easily achieved. Thus, mergers which increase HIs above this level are often challenged.

In January 1986, Pepsico proposed to acquire 7-Up from Phillip Morris for $380 million in cash. Pepsi argued that, unless 7-Up was bought by a company with large marketing resources, it could not hope to survive. A month later, Coca-Cola responded by announcing its intentions to acquire Dr. Pepper for $470 million. Even before the two acquisitions were proposed, the concentration in the soft drink industry was fairly high. In 1985, Coca-Cola and Pepsico had 66 percent of the market, with an HI of 2362. If both mergers had occured, the HI would have increased to 3258, a level far in excess of current federal guidelines. And, in fact, the mergers were blocked by the Federal Trade Commission. Indeed, the federal judge who ruled on an injunction in the case called the proposed acquisitions "a stark, unvarnished attempt to eliminate competition that totally lacks any apparent redeeming feature."[7] Dr. Pepper, we might note, was eventually sold in a leveraged buyout. This is an excellent example of a case in which the antitrust laws and the use of the Herfindahl index has had a big impact on the industry structure.

Homogeneity of Firms

Industries with relatively few firms tend to develop at least some cooperative mechanisms in their dealings with one another. The more alike are those firms, the easier it will be for cooperative mechanisms to develop. If firms are very similar, the **symmetric solution**—everyone behaves in the same way and reaps the same reward—becomes a natural solution. Then, too, if demand and/or technological forces perturb the industry, firms are more likely to be affected similarly than if they began in quite different positions. Thus, similarity among firms may dampen industry uncertainty.

The more similar are firms in a market, all else equal, the easier will be coordination of those firms.

A classic example of the importance of similarity in starting points is provided by Schelling. In lecturing to a class of Harvard undergraduates, Schelling posed the following problem:[8] You are to meet a friend somewhere in New York at 12 noon. But your friend has neglected to indicate the meeting place. Choose the most likely place to meet.

There is one overwhelmingly favorite choice among Schelling's Harvard students: Under the clock at Grand Central Station. Students arrive at this choice without any direct discussion. Apparently, Harvard students typically travel to and from New York by train. For this population, Grand Cen-

tral Station is a **focal point,** a solution to which similarly minded individuals converge without any discussion.

In our example, the focal point was a place. Other common focal points include prices—$3.95, not $4.01, is a common choice among price setters, for example. We also see focal points among bargaining rules: "split the difference" is a bargaining focal point. We tend to move to this solution unless there are good arguments to the contrary.

The ability to find a focal point depends in part on similarities among participants. Schelling's group consisted of undergraduates used to going to New York on the train. This common experience led to a common solution. In the same way, common management background and firm characteristics can ease the ability of organizations to find common ground. Even membership in common clubs can play a role in facilitating the search for commonality. Ouchi's work on "clans" suggests that homogeneity within organizations allows managers to act more or less autonomously because the organization's goals are congruent with their own.[9] So, too, homogeneity across firms allows for congruent behavior without explicit coordination. To the extent that firms in an industry differ significantly in the products they sell or the production processes they use, it may be harder for them to coordinate. In the same way, the entry of new firms into an industry (and particularly foreign firms) often proves to be quite destabilizing. New players and new ideas often have an influence on the industry well beyond their direct one.

Asset Specificity

Virtually all organizations require some fixed assets to conduct their enterprises. But there is wide variance in the extent to which fixed assets are used, or, stated alternatively, in the ratio of variable costs to total costs within the industry. In the consulting business, for example, the fixed assets employed are typically quite small relative to total costs. Most of the costs of the enterprise are labor costs. In heavy manufacturing, on the other hand, fixed costs form a substantial part of the total costs of the operation, and, moreover, are quite long-lived.

A second distinction among assets has also proved to be extremely important from a strategic perspective: the *specificity* of the assets used in a business.[10] One of the fixed assets of an airline is its airplanes. Moreover, airplanes have no alternative uses: They can only be used to fly. In this sense, an airplane might be thought to be a **specific asset** from the perspective of the airline. On the other hand, if we think about the air market as having a geographic component, the specificity of the asset diminishes considerably. That is, an airplane currently used in the New York to Denver market could easily be transferred to the New York to Los Angeles market. This transfer could be accomplished either by the initial airline, or through the competitive resale market. Contrast this example with that of the railroad and its

track. The track between New York and Denver is a fixed asset quite specific to both use and market. It cannot be used to serve the New York to Maine market, for example. Track is clearly a more specific asset for a railroad than are airplanes for the airline.[11]

Organizations with large amounts of specific assets find exit from particular markets quite difficult. A brand name can be a valuable specific asset. Information about the way a particular market works can also be an important specific asset. High asset specificity can be said to raise **barriers to exit** for an organization. These organizations find themselves with assets valuable only in the particular ventures to which they are currently committed. Lacking any alternative use for the asset, the organization will find it rational to continue deploying those assets in markets in which the accounting returns are exceedingly low. These assets can be said to have low opportunity cost. In the case of specific assets there are few if any alternative opportunities for the asset.

Industries which have substantial specific assets exhibit high barriers to exit and intensified rivalry.

Heavy reliance on specific assets encourages firms to stay in an industry even when times are bad, simply because there is nothing else they can do with these assets. Industries characterized by high specific assets are typically also characterized by intense rivalry during downturns. Firms will fight hard for market share to help them to cover at least part of their large fixed costs. Indeed, railroads, which are clearly an extreme in the high–asset specificity group, were regulated at least in part as a response to the fierce competition which arose from the operation of this market.

If we return once again to the comparison of diapers versus book publishing, it is clear that asset specificity plays a major role in distinguishing the two industries. The production of disposable diapers requires an extremely expensive, very specialized assembly line. Indeed, the equipment is so specialized that recent product innovations by Procter & Gamble required a major investment of hundreds of millions of dollars to put in a new assembly line. Book publishing, in contrast, uses a modest level of relatively unspecialized fixed assets. And, we recall, the diaper market is considerably more rivalrous than is publishing.

The extent to which a particular organization uses specific assets is in part under the control of the organization. In the case of airplanes, for example, I can lease a plane or buy it. In the case of a machine shop, I can buy a very effective but highly specialized machine or a general-purpose machine which is perhaps somewhat less effective. A railroad company could rent trackage rights rather than own them. The decision to use specific assets owned by the organization itself is a complex one, and will be discussed in Chapter 11. The central point to note here is that the decision to use spe-

cific assets affects the firm interaction in an industry, and thus looking at the extent of these assets is another part of our examination of the first Porter box.

Changing Conditions of Demand and Supply

One of the features that distinguishes the diaper and educational book publishing industries is the *stability* in the environments in which each industry is embedded. In text book publishing, demand has been relatively stable over time, paralleling school and college enrollments. Of course, there have been good years and bad, but variations have been relatively small and the growth rate on average only just above GNP growth. The market share of particular firms in the industry has also been relatively constant. The disposable-diaper industry, on the other hand, has exploded in the last twenty years as families switched from cloth diapers, going from a $20 million business in 1967 to a $4.4 billion business today. And, as we suggested, market shares have been volatile as well. In publishing, technological change has been modest—most has come from industries that supply publishing, for example, from the printing equipment industry. In the disposable diaper industry, on the other hand, technological changes and product changes have been large and rapid, often involving complete changeovers in product design and assembly line configuration. On many levels, then, publishing appears to be embedded in a considerably more stable environment than the disposable diaper industry.

As a related matter, we also note that industries differ considerably in the extent to which they face cyclical demand. In the diesel engine market, for example, sales and production vary dramatically over time, largely in response to business cycle movements. In 1975, the United States economy was in a recession and sales of diesel engine suffered as a result. The recession of 1982 produced a similar sales decline, as did the recession of the 1990–91 period. The diesel market is sensitive to GNP levels as diesel equipment is used in such raw materials industries as housing construction. The cylicity of the underlying user industry is thus pushed back into the engine market. In this industry, variable demand is a fact of life, outside the control of individual firms. But it is fact of life with substantial strategic effects.

Industries with unstable demand are often characterized by intense rivalry, in particular if this cyclical demand is coupled with high fixed costs, a factor we will discuss below. In the five-year period between 1979 and 1984, for example, one producer of diesel engines, Cummins, saw its market share of engines fluctuate over ten percentage points. Detroit Diesel, another major player in the industry, ranged in share from under 20 percent of the market to one third of the market. Such instability carries with it substantial managerial and strategic costs. Indeed, several industries that early economists described as having cutthroat competition are precisely the industries that suffer from cyclical demand. In these industries, when demand

is high, firms are led to increase capacity to serve that demand; in the subsequent slump, excess capacity encourages fierce battles for share. The identity of the market share leader changes over time, and the overall organization of the industry is fragile.

Variability in demand creates more rivalry within an industry.

But why does variability in demand and supply increase the level of rivalry in an industry? Increased rivalry occurs for several reasons. In the first place, coordination of pricing, branding, or research effects requires considerable agreement among the industry participants. Often, of course, agreement is tacit. Through the business press, for instance, firms make it known that modest price increases seem likely, and sure enough, modest price increases occur. But such coordination, tacit or otherwise, requires agreement among firms as to what the desirable outcome might be. And such agreement is much harder to arrive at when conditions of demand and supply are constantly changing.

Variability of supply and demand often creates additional uncertainty in an industry. This uncertainty will in turn influence the ease with which coordination can be maintained across disparate firms. We have already indicated, for instance, that firms face some temptation to move away from a coordinated price. In part, firms resist this temptation because they realize that other firms will observe them and retaliate. If demand fluctuates in unpredictable ways, however, then it will be very difficult to detect departures from the common ground.

Demand fluctuations, even when predicted, may have a second impact on industry structure. When demand is variable, *flexibility* begins to play a more important role in the industry. Under these conditions, there is often an opportunity for technically diverse firms to compete. In particular, small firms may be able to use their flexibility in these industries to compete against scale advantages of larger firms. Thus, demand fluctuations may increase organizational diversity in an industry.[12] And diversity reduces the ease of coordination among firms.

Competitive Perspective

The catalogue of industry characteristics just described as leading to industry rivalry are all **structural elements.** By and large, the traditional Porter industrial-organization framework focuses on these structural features of an industry as a way to do industry diagnosis. More recent work in economics and management, however, suggests some ways in which more subtle industry attitudes and perspectives can affect industry dynamics.[13]

In the classic economic framework used by Porter and others, firms within the same industry typically think of each other as rivals for the same

fixed pie. Hence, the competitive process we have just been describing. Brandenburger and Nalebuff, however, argue that in many cases firms would do better to think of themselves as **complementors** as well as competitors. In this way, firms can increase the size of the pie rather than compete over the slices. For example, while Compaq and Dell compete with each other for customers, they also complement one another in increasing Intel's sales and thus supporting its research and development efforts. As Intel's chips improve, the marketplace in which Dell and Compaq sell their wares expands, potentially helping them both. Brandenburger and Nalebuff argue that focusing on complementarities with one's rivals in a process they refer to as **coopetition** can create profitable opportunities in industries which otherwise look competitive.

Porter Force 2: Presence of Substitute Products

Thus far, we have been discussing the effects on industry profits of the intensity of rivalry within the industry. But firms are also affected by competition from related markets. Air conditioners compete with fans, disposable diapers with cloth, opera with musicals and dramatic theater, and beer with wine. In each of these cases, the availability of substitutes influences the ability of a firm to raise its price or change the attributes of its products. These substitutes are especially important in markets in which there are few rivals in the narrow market or ones in which it is difficult to increase industry supply quickly. In these cases, we would normally expect some excess profits to accrue to firms in the industry. The presence of good substitute products limits those profits.

How do we identify those products and services which are substitutes for a given product? Here, we wish to identify products and services which serve more-or-less the same function for more-or-less the same people as the product we are evaluating. Conceptually, we wish to ask: What set of products *constrains* the ability of firms in this industry to substantially raise their prices? Typically, when we ask this question we find a chain of substitute goods. Consider the market analysis done from the perspective of Kimberly-Clark, looking at its Huggies brand disposable diaper. The closest substitute to this product are Luvs and Pampers, the other premium brand products made by Procter & Gamble. Next in line come the regional and generic brand diapers made by Weyerhauser, Veragon, and a number of other small producers. Beyond these disposable diapers come the cloth diapers, provided by diaper services. Finally, we have cloth diapers laundered at home. For each of these products, we wish to ask: How much does the presence of this product influence Kimberly-Clark's ability to raise its price on Huggies? The **cross elasticity of demand** is one measure of this effect. It is the ratio of the percentage change in the demand for one good in response to a 1 percent increase in the price of a second good. Substitute goods have positive cross-

elasticities: As the price of branded disposable diapers increases, we would expect an increase in the demand for generic brands. The higher the cross-elasticity, the closer the substitutes. Very close substitutes should be analyzed as part of the within-market competition which we covered in the last section of this chapter. Further-away substitutes, while out of the market proper, may nevertheless have an effect on pricing behavior. In the case of Huggies, I think it is reasonable to include all disposable diapers as part of the market proper, while cloth diapers should be analyzed as one of the substitute products outside the market.

Substitute products play an uneven role in industry dynamics. In highly competitive industries, or during periods of excess capacity, substitute products play a very modest role. All of the action in determining industry profitability is within the industry. But, in times of rapidly increasing demand—or in industries in which there are few competitors—substitute products may become quite important. An excellent example of the growth in the importance of substitutes in a dynamic environment is provided by the Encyclopedia Britannica.[14] In the 1990s, Britannica's sales of encyclopedias plummeted, not as a consequence of new competition from other encyclopedias, but rather from the information on the Web. Internet information has become a more potent substitute in the information industry, and as a consequence, Britannica's sales declined by more than 50 percent in just a few years.

Porter Force 3: Buyer Power

All firms need to pay attention to what their customers want. Nevertheless, there are considerable differences across markets in how powerful buyers are and in how able they are to force down prices or influence product quality levels. The steel industry sells a substantial amount of its output to the auto firms. Clearly, this set of buyers has a lot more power than the individual consumers who buy from Coca-Cola. For Coca-Cola, on the other hand, the large fast-food chains like Burger King and McDonalds wield considerable power.

The first factor to look at in determining buyer power is the number of buyers and the distribution of their purchases. The larger the number of buyers and the smaller their individual purchases, the less power each one will have. Secondly, there are some characteristics of the product itself. Standardization of products increases buyer power since it typically reduces switching costs of those buyers and allows them to more easily play one supplier against a second. Third, when buyers can integrate backwards, producing the good for themselves, this also increases their bargaining power. Electric utilities, ostensibly regulated monopolies, have faced increased pressure in the last decade from customers who are able to **bypass** the power network and generate their own electricity. Finally, there are institutional

factors associated with the way transactions are done in the industry that are relevant to buyer power. The more open are transactions, the more power buyers have, in part because this reduces their costs of search among producers.

Porter Force 4: Power of Suppliers

In the same way that powerful buyers can squeeze profits by putting downward pressure on prices, suppliers can squeeze profits by increasing input costs. And, we will see, the same factors that determine the power of buyers also determine the power of suppliers.

First, we consider the number of suppliers available. The more suppliers, the better. Product standardization is also important: Standardized products reduce a firm's vulnerability to supplier pressures. In the semiconductor industry, the government—an early large buyer of semiconductors—required firms to license other firms to serve as second sources for government contracts. This practice insured the government purchaser against hold-up by suppliers. In the long run, it also reduced the profitability of the semiconductor industry. If firms can pose a credible threat to backward integration, this will further reduce supplier power. Most of the major computer companies produce at least some chips themselves. This partial integration insulates them from pressure by sellers, increasing the profitability of the computer industry at the expense of the semiconductor industry. Finally, the more open is information in the industry, the less power will be held by suppliers.

One tool that can be useful in understanding the balance of power in the bargaining between a firm and its suppliers is the **supplier matrix.**

The matrix, given in Figure 3.3, characterizes buyer-supplier relations in two ways. On the vertical axis is measured the supplier's monopsony power over the buyer. Here we ask the question: What fraction of a particular input is supplied by a given vendor? On the horizontal axis, we measure the firm's power over that same buyer. Here, we ask: What fraction of the supplier's output does the buyer use? In the discussion that follows, I focus on the actual numbers of buyers and vendors engaging in a transaction. Of course, the presence of a fringe of equally able potential vendors has the same effect on industry suppliers' power as would actual vendors.

The four cells in Figure 3.3 depict four different dependency relationships between suppliers and buyers. In the upper left corner, the supplier wields most power; here, the supplier dominates in its supply of an input to the firm, but the firm is a small player to the supplier. I call this position *vulnerable;* buyers here are often subject either to price pressure or supply interruption. The bottom left corner represents anonymous markets; individual buyers and sellers are of little import to one another. Transactions

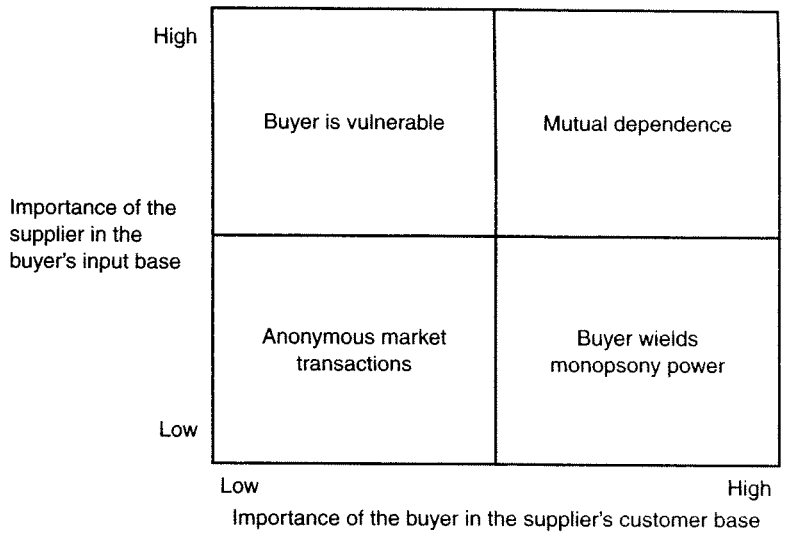

Figure 3.3 A Supplier Matrix.

here occur without much angst. In the bottom right, the buyer wields power; here, the firm has multiple suppliers, each dependent on it for substantial business. Finally, in the top right box, we have the most complex bargaining situation: Buyer and supplier are mutually dependent. In this situation, long-term relationships—sometimes even mergers—are likely.

The supplier matrix can be used to dissect buyer-supplier bargaining positions in a broad industry setting to forecast which party is likely to get most of the profit. Consider the personal computer market. In its entry into the PC market, IBM decided early to contract out control of the operating system to Microsoft, then a small company. At this point, IBM had considerable strength relative to Microsoft. Indeed, I would place IBM in this transaction in the bottom right corner of the matrix, able to extract most of the value. As time progressed, however, two things occurred. First, more firms entered the PC market, thus reducing Microsoft's dependence on IBM as each of the new entrants became a potential customer for its operating system. This shifted the transaction leftward on the matrix. At the same time, Mircosoft's control over its technology and the development of volumes of software written for this proprietary operating system shifted the transaction upward. As a consequence, today in the IBM-Microsoft transaction, IBM is firmly in the upper left corner, with Mircosoft in a position to extract most of the value in the deal.

We will explore the use of supplier matrices further in later chapters as we look at alternative organizational relations between firms and their suppliers.

ANOTHER APPROACH TO DETERMINING PROFITS

Throughout this chapter, we have focused on the industry. Our approach implicitly assumes that it is possible to group firms into industries and to make some meaningful generalizations about the firms within those industries. Another approach to strategy work is to forego industry analysis altogether and focus instead on the transaction level. Here, the central question is: Who captures the value associated with a particular transaction and how can I help insure it is me?[15] Before we proceed to look at entry issues, it is useful to summarize this alternative approach to analyzing profitability.

The **value-based or value-added approach** begins, as does the Five Forces model, with economics. Any firm can be thought of as lying between a collection of its suppliers on the one hand and its buyers on the other. The profits that can be taken from any transaction are bounded by the demands of the buyer on one side and the costs of the suppliers on the other. In particular, no customer will pay more for a good or service than his or her maximum **willingness to pay**, and no supplier will accept for its input any less than the **marginal costs** of producing that input. From these basic principles of microeconomics comes the following.

The value created in a transaction is the difference between consumers' willingness to pay and the opportunity cost of production.

The question value-added analysis is intended to address is which party captures this added value in any particular transaction. The answer depends in large measure on how much competition each party faces. To the extent that a supplier has an advantage over its rivals, either in cost or product quality terms, that supplier will have bargaining power and be able to capture value in the transaction. Indeed, in the formal analysis, the value any party can capture is limited by the incremental value he or she brings to the transaction. We explore some of the mechanics of the added-value approach in a later chapter.

THE ROLE OF GOVERNMENT IN DETERMINING INDUSTRY PROFITABILITY

Virtually all of American business is, in one way or another, influenced by government. The government, while not explicitly one of Porter's Five Forces, has an overarching effect on all of the other players in the market. Government action affects levels of rivalry, buyer and supplier power, the importance of substitutes, and the power of entry. In a later chapter, we will explore some of the strategic implications of pervasive government regulation. But at this point it is worth noting that heavy-handed regulation and antitrust laws may dramatically affect industry profitability.

Regulation may have a dramatic effect on rivalry within an industry.

The banking industry provides an excellent example of the effect of regulatory changes on relations within an industry. Banking regulations affect virtually all facets of the bank's operations: where it can operate, what prices it can pay suppliers (i.e., interest rates), and which businesses it can enter. Originally, restrictions were imposed by Congress during the Great Depression in an attempt to limit the risks faced by banks and depositors and prevent certain conflicts of interests or abuses of the type associated with the crash of 1929. The effect of the regulations was thus to dampen variability within the industry. And, for a long period, the industry was considered a relatively stable, cooperative one.

Since the early 1970s, regulations within banking have eased considerably. Interest rate ceilings were first raised, and then removed; some limits on branch banking have been removed; some of the regulatory limits on diversification have been relaxed. The effect of deregulation has been to considerably increase the uncertainty in the industry, since matters that were controlled by regulatory fiat are now under managerial control. And, as a consequence, the industry is becoming truly rivalrous for the first time.

For many banks, these new pressures have created management concerns that were simply not present fifteen years ago. During the period of heavy regulation, marketing was the primary arena for competition among banks; now pricing is moving more to center stage. In the regulated period, mergers and outside growth possibilities were limited; since deregulation, all banks have had to adopt a much more external focus to their planning. State legislation, which has eased the rules on interstate banking mergers, has encouraged the development of super-regional banks, institutions often based in the hinterlands that now can operate across state boundaries. Banks like Fleet Financial in Rhode Island and Wells Fargo, based in San Francisco, are beginning to offer tough competition for the New York money center banks. The effect of deregulation on the relations among firms in the banking sector has been profound and has included both substantial mergers and many bank failures.[16]

Two useful additional illustrations of the importance of regulation are provided by the airline industry and the telecommunications industry. In the period before 1982, the United States airline industry was heavily regulated by the Civil Aeronautics Board (CAB). The CAB restricted entry and exit into markets, and regulated fares as well. Aircraft type, frequency of flight, and service quality, on the other hand, were not controlled. As a result, airlines competed almost entirely on the basis of quality: flights were frequent, meals and movies elaborate, aisles wide.[17] The effect of deregula-

tion in this case has been to change the *form* of competition in the industry as well as its intensity. In many instances, it is easier to collude on service than on price; prices can be changed rather quickly in response to competitor action, whereas a change in service characteristics often requires a longer-run strategic decision. Though an airline cannot widen its aisles overnight, it can decide, almost overnight, to change the quality of meals. As deregulation altered the possible arenas for competition, it had the effect of changing the level of intensity as well. We have all seen the results in terms of falling prices and changes in the product mix of the airline industry. We have also seen the failure of long-term industry participants such as Eastern, Braniff, and TWA.

Regulation has had a rather different effect in the telecommunications industry. Throughout most of the history of the United States telecommunications industry, the long-distance market was serviced entirely by AT&T. Entry into the market was controlled by the Federal Communications Commission (FCC). This began to change with a series of court cases in the late 1980s, until the market was completely opened up by the Modified Final Judgement settlement of an antitrust case against AT&T separating long-distance telecommunications from local telephone operations. With one fell swoop, much of the protection afforded AT&T was removed, and Sprint, MCI, and numerous small competitors rapidly gained share. In just a few years, due in large measure to regulatory changes, the market has gone from one without rivals to an extremely rivalrous market in which the competitive weapons range from pricing and marketing to lawsuits.

In the mid 1990s, the European telecommunications industry began opening up in similar ways, and companies like Deutsche Telekom are now facing true competition for the first time. Alliances with other telecommunications firms, both European and American, are developing as one response to the new marketplace.

Antitrust Laws

The main statute governing the behavior of firms in the United States on antitrust issues is the Sherman Antitrust Act. The Sherman Act was passed in 1890 at least partially in response to complaints by farmers and small consumers about the pricing practices of newly created trusts and big businesses. Virtually all of U.S. federal antitrust law rests on the Sherman Act and its two later companions, the Clayton Act of 1914 and the FTC Act of 1914. The Sherman Act limits the ability of firms in an industry to coordinate their behavior.

Section 1 of the Sherman Act proscribes "every contract, combination . . . or conspiracy in restraint of trade or commerce among the several states." A reading of the case law in this area indicates that a variety of agreements among firms in an industry are *per se* illegal; in other words, these actions

are intrinsically illegal, regardless of whether or not there is evidence as to negative social effect. Acts which are *per se* illegal include agreements among competing firms to fix prices, to share markets, or to restrict or pool output. Any express agreements among firms in an industry in these areas is a violation of the law, regardless of ultimate effect.

In many situations, it may be possible for firms to reach pricing or product agreement without express agreements or even direct discussion. Firms can at times send signals to one another through the media, for example. Information relevant to pricing decisions may be disseminated by a trade association. At present in the United States, collusion without formal agreement—sometimes called tacit collusion—is in a grey area of the law. The *spirit* of the Sherman Act is clearly against such collusion, but in practice the likelihood of litigation in areas of tacit collusion is relatively small.

Antitrust laws attempt to increase firm rivalry

In cases of tacit collusion, the Justice Department has two remedies it can use under the Sherman Act. First, it can institute a civil case and seek, for example, an injunction against the prohibited act. Or, it can institute a criminal proceeding, leading to fines and imprisonment for corporate personnel. In the well-known 1960 conspiracy case against electrical equipment manufacturers, seven senior managers were sent to prison as a consequence of a price-fixing agreement.[18]

The Sherman Act clearly imposes boundaries on the kinds of coordination that can occur in industries. It is interesting to contrast the United States' situation with the situation abroad. While most other industrialized countries have some antitrust laws, "agreements" are not *per se* illegal. It is common for Japanese firms to work together on plans for industry cutbacks in depressed industries, with some firms, for example, compensating other, higher-cost firms for lowering production. Such coordination would be illegal in the United States. One explanation for the greater accommodation of the Japanese government is that, since many of the Japanese industries are primarily export markets, the world is perhaps thought to provide sufficient discipline for firms.

HISTORY AND INSTITUTIONS

The Porter model outlines a way to structurally dissect an industry. But in many ways, it ignores industry history and institutions. Yet these factors may be very important in understanding the level of profitability in an industry. To the extent that firms have existed in an industry for long periods, confronting one another in many and varied submarkets, cooperative mechanisms have had time to grow and develop. A firm that anticipates be-

ing in an industry with a given set of other firms for some period will have quite a different posture than will a firm only temporarily active in that same industry, even if the industry structure is otherwise identical. The knowledge that I will face the same rival tomorrow surely influences how aggressive I am likely to be today. Thus, history has an overarching effect on industry profitability, in much the same way that government does. In some industries, as time passes, institutional mechanisms develop that improve the ability of firms to coordinate. These institutional mechanisms are often referred to as **facilitating devices.** Facilitating devices may evolve as a result of government intervention in the market or through private initiative. They may at times be devised carefully to serve as facilitators or they may arise innocently to serve another purpose and have facilitation as a byproduct. In either case, knowledge of the existence of these facilitators is an essential part of industry analysis.

Facilitating devices can help coordinate an industry through one of two ways: *information exchange* and *incentive management.*[19] An excellent example of an information-exchange device is the practice of preannouncing price changes, a practice common in the steel industry as well as in a number of other raw materials industries. By announcing a price change to take place six weeks hence, raw materials suppliers ease the transition for their buyers. They also eliminate some of the uncertainty associated with price competition in the industry and reduce any transitional gains from price discounts. In the absence of preannouncements, some firms can build share by underselling their competitors. By announcing price changes, firms make following those changes easier. Preannouncements also make it easier for firms to see if price increases will stick.

Trade associations are another institution that may improve industry coordination through information exchange. A **trade association** is an organization of firms in a given industry, designed to promote the overall industry and collect and disseminate information about that industry. Often the information collected and published by these associations is extremely helpful; indeed, this information is often used by industry analysts as they report on trends. But information disseminated by the trade association can also serve to improve industry coordination by reducing collective uncertainty about the future. Informal contacts made at trade association meetings can also help prepare the way for the kind of focal point agreements described by Schelling.

Facilitating devices also act to alter the incentives of firms. Consider the **most-favored-nation clause** (MFN) found in some sales contracts. Under this clause, the seller agrees that if the firm offers *any* buyer a lower price during a specified period of time for the particular item, that same price will be given to the firm with which the MFN has been signed. In other words, selective price discrimination is reduced. I cannot offer to sell jeans to Macy's at one price and Bloomingdale's at another. The MFN clause thus creates a penalty for price cuts, since cuts must be extended to more customers. If an entire industry uses MFN, the effect will be to stabilize industry prices.[20]

Recent work suggests that the most-favored-nation requirements imposed by the federal government in its application of the Medicaid reimbursement program have substantially increased pharmaceutical prices.[21]

Another facilitating device that operates through the incentive principle is the **Meeting competition** clause. Under this provision, a seller offers to meet the price of any responsible rival. Such clauses, which are common in long-term contracts and even in retail-outlet advertising, perform different functions in different contexts. They protect sellers against losing a sale to a price cutter; alternatively, they reduce incentives to cut prices, since such price cuts will be matched automatically by the firms operating with meeting-the-competition clauses.

Best price provisions combine the features of meeting competition and most-favored-customer protection. Sellers must meet the best price offered to a buyer by rival *and* make this price available to any buyers with whom an agreement has been signed. Provisions of this form, which were found in the gasoline-additives market, also appear to reduce incentives of firms to compete on prices and stabilize prices at levels above competitive ones.[22]

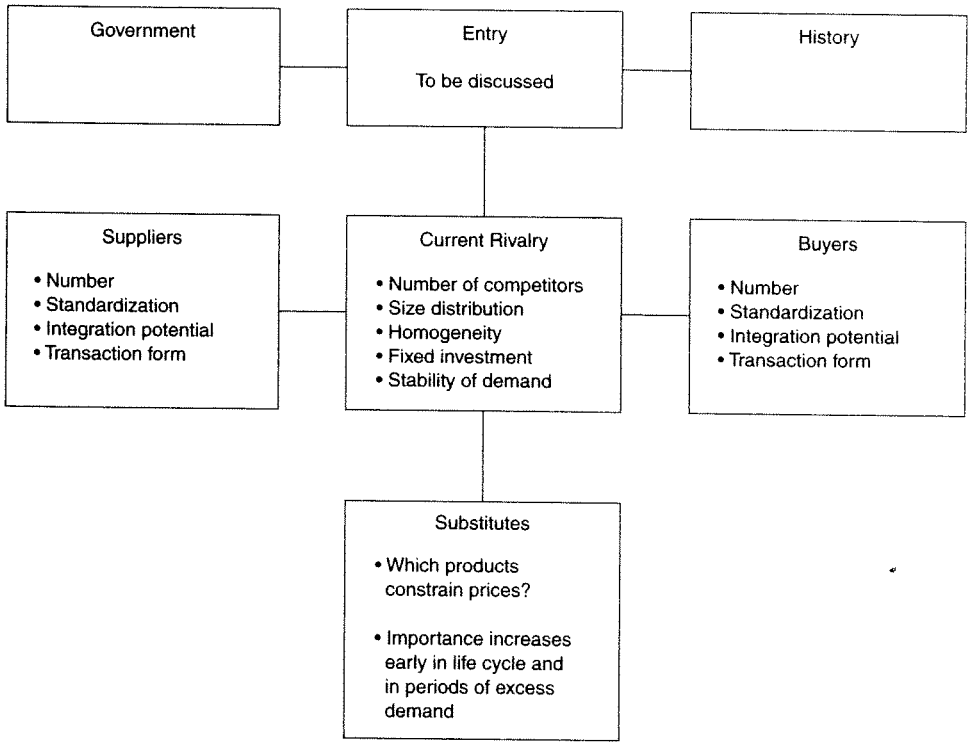

Figure 3.4 Schematic for determining intensity of competition.

Facilitating devices can ease coordination difficulties.

SUMMARY

We have now reviewed four of the five forces explaining profitability within an industry, as well as the overarching influence of government and history. These factors are summarized in Figure 3.4 and provide a guide to industry analysis.

4

Understanding the Impediments to Entry

In the last chapter, we explored four of the five factors which affect industry profitability, as well as overarching roles of government and history. Here, we complete the analysis by looking at the final and, arguably, the most important force: the ease of entry into an industry.

The difficulties of entry into new markets is well exemplified by AT&T's experiences in the computer market. AT&T first attempted to enter the computer industry in the early 1980s. Many would have predicted that with its strong propriety operating system Unix, substantial R&D resources, and strong complementary telecommunications business, AT&T would have an easy passage into the computer market. Nevertheless, the company lost $2 billion in an entry attempt in the 1980s. Between 1991 and 1994, AT&T tried again with the purchase of NCR. In 1994, AT&T dropped NCR's brand name and tried to establish a new brand, Global Information Systems (GIS). This, too, failed to take off. Overall, in a fifteen-year period AT&T lost more than $10 billion in its entry attempt. In this chapter, we explore some of the features of a market that can, at times, make entry so difficult.

THE ENTRY DECISION

Suppose we have a relatively new industry in which excess profits are being earned. What factors would influence a new firm's decision whether or not to enter this industry? What, in particular, would discourage entry given the presence of excess profits? Since, as we have seen, entry is a prime vehicle for reducing rates of return, factors that discourage entry will be viewed as favorable from the perspective of industry incumbents. Thus, the analysis of entry is a fundamental part of the environmental analysis of a strategic plan, the fifth force in the traditional Porter analysis. It is clear that in practice all