

Chapter 5: Firms and Industries

“By establishing its Web presence, although belatedly, Barnes and Noble is in effect destroying its traditional value chain.”

Larry Downes and Chunka Mui, A conversation with the authors of *Unleashing the Killer App*, www.hbsp.harvard.edu/products/press/books/killerapp/ka_qa.html

Prologue

Downes and Mui’s best-selling business book, *Unleashing the Killer App*, illustrates the importance of economic ideas. They reached back to the 1937 work of Ronald Coase, a Nobel Prize winning economist, for some of their arguments. The concept of the value chain was introduced in the 1980s by influential management guru Michael Porter, and has formed the basis for much of the subsequent thinking on business strategy formulation and implementation. A value chain, in the words of Downes and Mui, is the series of functions that create and distribute a company’s goods or services. They argue that e-commerce alters value chains dramatically or destroys them altogether. Part of their argument is based on Coase’s emphasis on transaction costs of market exchange as leading to the existence of firms. They assert that technology is increasing market efficiency faster than the internal efficiency of firms, leading to what they call the Law of Diminishing Firms – “as transaction costs in the open market approach zero, so does the size of the firm.”

Will firms disappear? Downes and Mui are more cautious here. In their ‘conversation’ with their publisher, they say:

No, but the nature of firms will change and already is changing. Increasing market efficiency puts tremendous pressure on companies to shed less efficient parts of their operations. The concept of a firm as a physical entity, now defined by its employees and fixed assets, will give way to what some would call a “virtual organization.”

What exactly did Coase say? Is that all that there is to the modern economic theory of the firm? And how do transaction costs relate to the value chain? What are the dimensions of firm strategy? Is the value chain concept irrelevant, to be replaced by “virtual organizations”? We will answer some of these questions in later chapters, but clearly, the claim is that something dramatic is going on. Where are we actually starting from? How different, really, is Amazon from Barnes and Noble, or from Wal-Mart? How is Microsoft different from IBM?

Read on!

5.1 Introduction

We reviewed some of the key aspects of the standard economic analysis of the firm in Chapter 4. In that analysis, the firm was presented in a stripped-down fashion, with much of what actually goes on within a firm being swept under the rug. This allowed us to focus on some important insights about the capture of value by firms as sellers in the marketplace. E-commerce, because it potentially changes the way that firms are organized, as well as how they operate in the marketplace, forces us to look inside the 'black box' that constitutes the firm of Chapter 4. In doing so, we link up with the management approach to firms and their behavior.

The management approach analyzes the details of the workings of firms, to understand how they create value as well as capture it. The starting point for us will therefore be the 'value chain', which describes the firm as a sequence of value-creating activities (Section 5.2). We can think of this also as an extension of the concept of the production function (Chapter 4). We can extend the value chain idea to include activities that take place beyond the boundaries of the firm. Thus the firm's 'supply chain' becomes a part of this extended value chain.

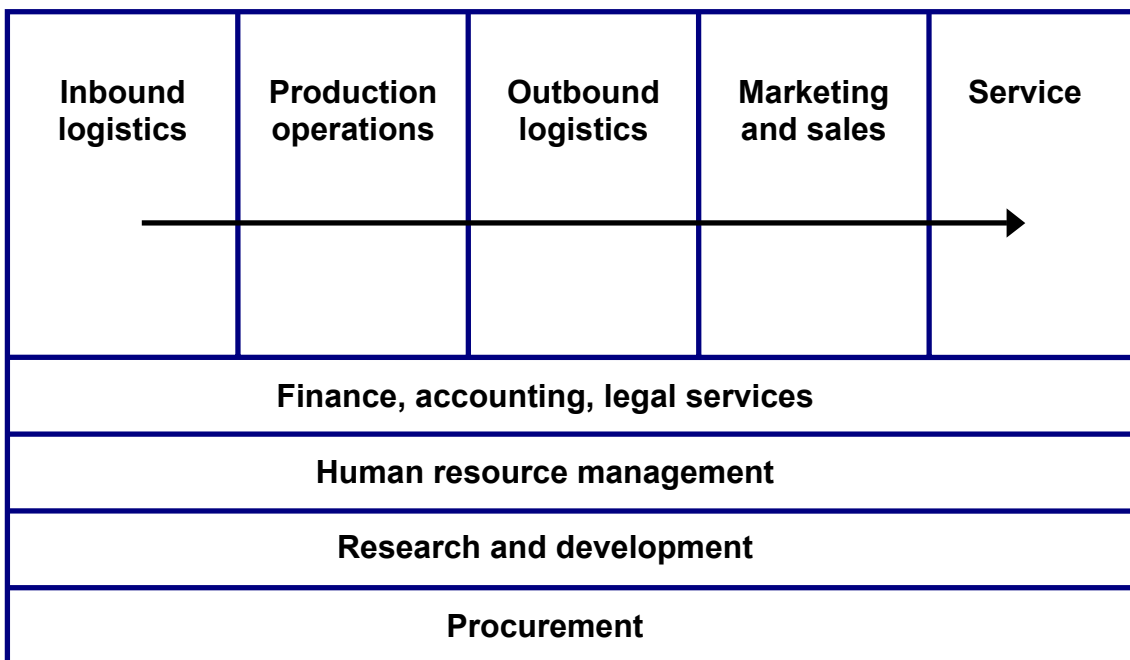
What exactly is a firm, and why do firms exist? Very abstractly, firms are ongoing collective action mechanisms. This concept is more general than firms, of course: it includes non-profit organizations, governments, and even households. Firms are distinguished from these other organizations by their legal structures, and by their objectives. However, this perspective is useful in identifying some of the fundamental reasons for the existence of firms, or any of these other collective action mechanisms. We discuss this in the context of the ideas of economies of scale and scope (Section 5.3). This forms the basis of a discussion of the boundaries between one firm and another, and between the firm and the market (Section 5.4). Why are some activities performed within firms, while others take place through market transactions? For example, when will a firm use an employee, and when will it use an independent contractor, or when will it outsource production of an input, instead of making it in-house? The answers to these questions turn out to be rich in possibilities. They also point the way toward understanding how the Internet may change the way firms are constituted: we take this up in Chapters 10 and 11.

We round out the chapter with several topics. We provide a discussion of the location of economic activity, and how products and services are moved around geographically (Section 5.5). This is obviously central to a revolution that promises the 'death of distance'. We add some more detail to the discussion in Chapter 4 of industry structure and its consequences for strategic behavior (Section 5.6). In particular, we provide an overview of the dimensions of strategic decision-making that confront firms (Section 5.7). This lays some more groundwork for the treatment of strategy in Chapters 13 through 17.

5.2 Value Chains and Supply Chains

The value chain is essentially a stylized listing of the sequence of activities performed by a firm. Each activity in this sequence adds further value to the product. You can think of the value chain as an expansion of the transformation expressed in the production function, $Q = F(L, K)$. All the activities shown in the value chain illustrated in Figure 5.1, including sequential operational and selling activities along the top, and the support activities below, are thus subsumed in the mathematical function, $F(L, K)$. The arrow in the figure merely indicates the direction of activities. Clearly one could further disaggregate processes, to capture the technological details of production operations, for example, but the boundaries in the typical value chain are meant to capture points at which value creation can be reasonably measured.

Figure 5.1: The Value Chain



An example will illustrate the stages of value creation behind the value chain concept.¹ In this example, we neglect the support activities, which are common to all the stages of the value chain, and would therefore need to have their costs allocated somehow across the steps of the chain. The example concerns Ekin, a not-so-well-known, rather backward manufacturer of athletic shoes. The firm sells its product through three channels:

1. Under its brand name, to independent wholesalers, who distribute them to retailers.
2. Under its brand name, directly to retailers.
3. Unlabeled, to firms that attach their own house brands and sell them.

¹ This example is adapted from *The Economic of Strategy*, by Besanko, Dranove and Shanley, Chapter 12.

The first two channels require the manufacturer to spend on product promotion and advertising. Using the second channel means that the firm also has to bear the distribution costs itself. The simplified value chain for Ekin is shown in Figure 5.2, including three stages in the value chain: production operations, brand management and distribution. In one way, the value chain is more complicated than that of Figure 5.1 because it shows the three different channels the firm uses. Activities performed outside the firm's boundaries are shown by separate rectangles. This illustrates an idea we shall return to shortly, namely that the value chain can encompass the activities of more than one firm.

Figure 5.2: Ekin's Value Chain

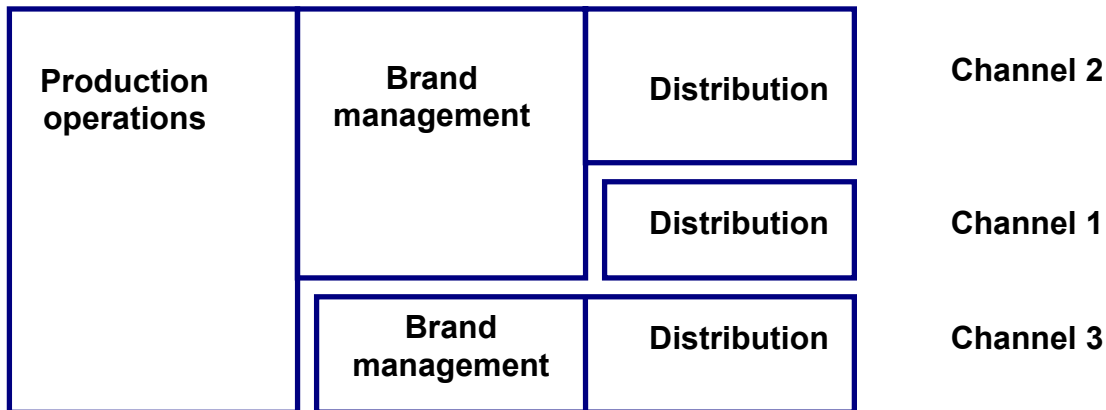


Table 5.1 provides hypothetical data for Ekin, for production quantities and costs, and sales through its three channels. This data allows us to calculate the value created, or the **value added** by Ekin at each stage of the value chain. Note that while the value created is obvious for the production stage, since the shoes are actually manufactured then, brand management creates value in the minds of customers through advertising, marketing and promotions. As long as customers are willing to pay more for the brand, value is created, irrespective of whether any objective improvement in the shoes occurs as a result of branding. Distribution creates value by relocating shoes closer to where customers can conveniently buy them, as we discuss further in Section 5.5.

For the purposes of the following calculation, we will assume that the costs reported include the full economic costs (including the implicit cost of capital), so that the value added we calculate is true economic value added. The economic profit for Ekin is given by the sum of the net revenues from each sales channel, minus the brand management costs, as follows:

$$70,000(\$20 - \$5.25) + 10,000(\$25 - \$5.25 - \$3) + 20,000(\$8 - \$5) - \$900,000 = \$360,000.$$

This profit can be disaggregated into value added per stage as follows:

- Value added in production = profit if all shoes were sold unbranded
 = $100,000(\$8 - \$5) = \$300,000$ or \$3 per pair

- Value added in brand management = extra profit if all branded shoes were sold to wholesalers instead of unbranded
 $= 80,000[(\$20 - \$5.25) - (\$8 - \$5)] - \$900,000 = \$40,000$ or \$0.50 per pair
- Value added in distribution = extra profit from self-distributing shoes to retailers instead of selling to wholesalers
 $= 10,000(\$25 - \$20 - \$3) = \$20,000$ or \$2 per pair

Table 5.1

Volumes, Prices and Costs	
Total quantity of shoes manufactured, per year	100,000 pairs
Branded shoes, sold to wholesale distributors, per year	70,000 pairs
Branded shoes, self-distributed, per year	10,000 pairs
Unbranded shoes sold to private label firms, per year	20,000 pairs
Selling price, branded shoes, sold to wholesale distributors	\$20 per pair
Selling price, branded shoes, self-distributed	\$25 per pair
Selling price, unbranded shoes	\$8 per pair
Production cost per pair, unbranded shoes	\$5 per pair
Production cost per pair, branded shoes	\$5.25 per pair
Distribution cost per pair	\$3 per pair
Total brand promotion and advertising expenses	\$900,000

Interestingly, the value added calculations show that brand management in this example lags both in terms of total value added (behind production) and in terms of value added per pair (behind production and distribution). Does this mean that Ekin is better off not spending so much on brand management? That is one possible answer, but we shall see, in Section 5.6, that other answers are also possible, depending on the strategic situation that the firm is in.

We now return to the idea that the firm may outsource part of its value chain. In Figure 5.2, Ekin sold part of its output to another firm that performed the branding, and presumably the brand management. This is not quite what we normally mean by outsourcing. Suppose instead that Ekin used an outside agency for all its marketing, promotion, advertising and so on, but on behalf of its own Ekin brand. This is what outsourcing typically refers to. Sections 5.3 and 5.4 tackle the varied possible reasons for outsourcing and other issues in redrawing the boundaries of the firm. Figure 5.3 shows the same value chain as in figure 5.1, but with Human Resource Management and Customer Service both outsourced to specialist providers of these functions.

Figure 5.3: Outsourcing and Splitting the Value Chain



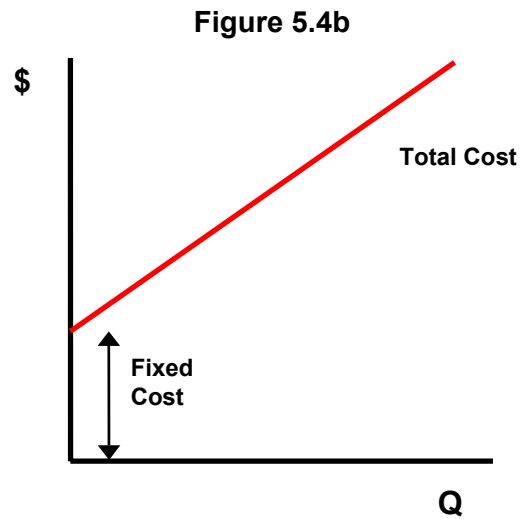
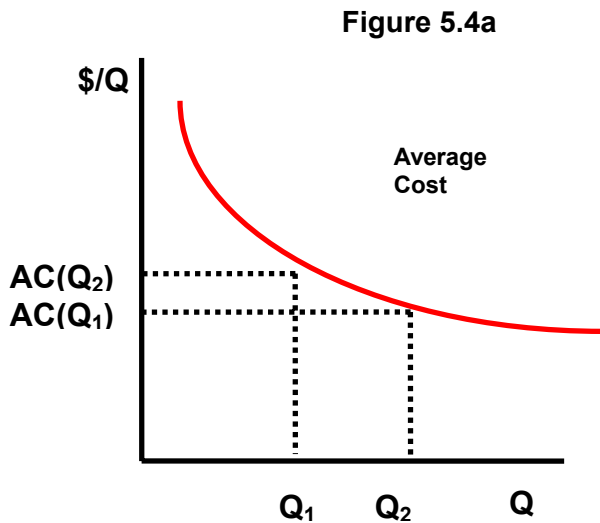
Much attention in e-commerce has recently been placed on moving B2B transactions to online exchanges. For the buyer in such a transaction, the seller is a supplier, and is therefore part of its **supply chain**. The supply chain is therefore a part of the overall value chain. Some portions of the supply chain may be within the value chain of the firm, while other portions are handled by other firms. We will use the example of Wal-Mart to illustrate this point. Wal-Mart's supply chain includes the manufacturing of the products it sells, transportation of these products to its distribution centers, and then transportation from the distribution centers to its retail stores. The first two parts of this supply chain are handled by other firms, while the final portion is under Wal-Mart's control. Activities such as marketing and customer service are part of Wal-Mart's value chain, but they are not supply chain activities. Logistical elements of storage and transportation have traditionally been important in business analyses of supply chains. We discuss some of these in Section 5.5. In the 'new economy', however, information exchange and timing become more significant aspects of the supply chain, as we discuss in Chapter 11.

5.3 Scale and Scope

Fundamentally, a firm is a method of ongoing collective action, with the objective of producing something for sale and profit. Production may not involve a material product, but may be the provision of a service. For example, buying finished goods and transporting them to another location for resale fits this description. It is possible to think of one person acting as a firm, but typically even a sole proprietorship requires coordination of the activities of more than one person. Other people involved may not even be employees, but simply providers of inputs or services along the value chain,

through market transactions. We shall partly explore the determinants of these choices in this section, and return to it in broader context in the next section.

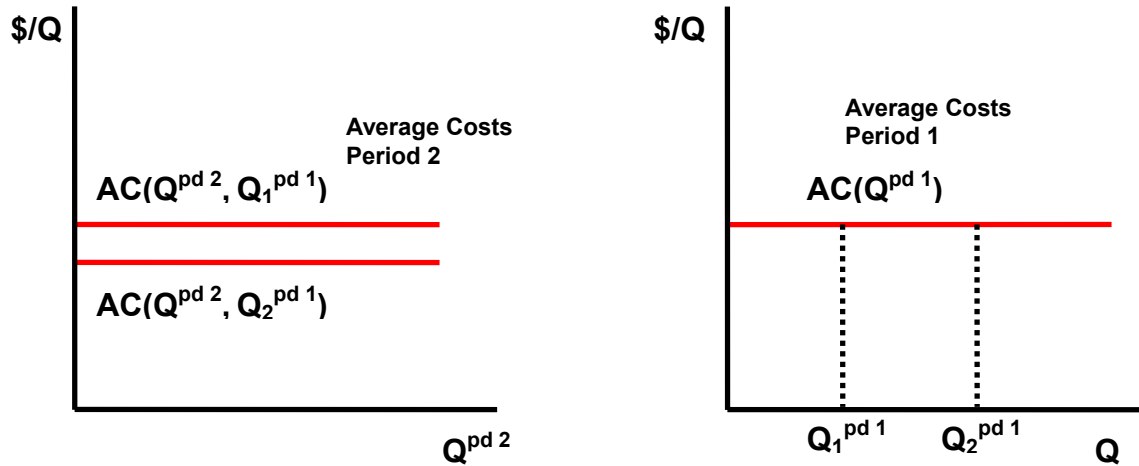
Economies of Scale A simple place to begin analyzing firms is through the properties of the cost function. Recall that the cost function gives the lowest cost of producing any given level of output. If the output is Q units per period, the lowest cost of production is denoted by $C(Q)$, measured in dollars. A cost function displays increasing returns to scale, or economies of scale, if the average cost decreases as output per period increases. In mathematical terms, we have that $AC(Q_2) < AC(Q_1)$ whenever $Q_2 > Q_1$, as shown in Figure 5.4a. This concept was reviewed and illustrated in Chapter 4 (Figures 4.2 and 4.3). Here we offer a slightly different view, the increasing returns to scale coming from the existence of fixed costs, that is, costs that are borne at a fixed level irrespective of the level of output. A total cost function that has this property, and can underlie the average cost function in Figure 5.4a, is shown in Figure 5.4b. An equivalent term for increasing returns to scale is **economies of scale**.



Economies of specialization In the standard concept of economies of scale, producing more in a period reduces the cost per unit of production. We can extend this idea to costs over more than one period. Suppose that producing more in one period reduces the average cost of production in the next period. This can be thought of as **dynamic economies of scale**. Mathematically, using superscripts to denote outputs in periods 1 and 2. $AC(Q_2^{pd 2}, Q_2^{pd 1}) < AC(Q_2^{pd 2}, Q_1^{pd 1})$ whenever $Q_2^{pd 1} > Q_1^{pd 1}$, as shown in Figure 5.5a. To emphasize the difference in this idea, we have drawn the figure without economies of scale within any period. In Figure 5.5b, therefore, the average cost curve for period one is horizontal.

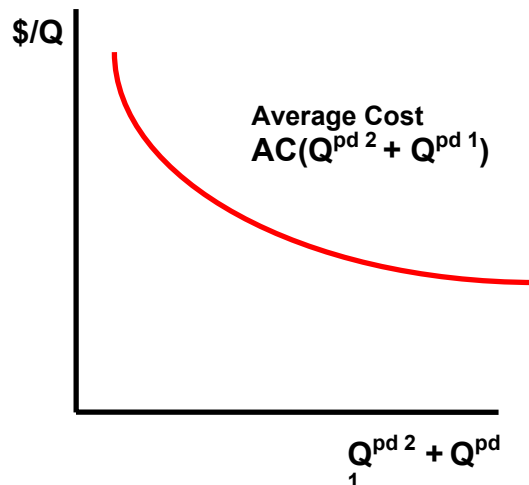
Figure 5.5a

Figure 5.5b



To show exactly how dynamic economies of scale reduce average costs in a manner analogous to Figure 5.4a, we can assume that what matters for average cost is cumulative output. This is different from the situation in Figure 5.5, since now there are economies of scale within a period as well. In Figure 5.6, we draw average cost as a function of cumulative output, $AC(Q^{pd 2} + Q^{pd 1})$. Now average cost is decreasing in cumulative output.

Figure 5.6



What is the source of these dynamic economies of scale? The justification is somewhat different from per period fixed costs. Instead, we can think of dynamic economies of scale as resulting from learning by doing, or specialization. The more you do something, the better you get at doing it. This improvement is reflected in lower costs (less time taken, better use of the machines required for production, and so on).

We have so far been thinking of these benefits as applying at the level of the firm. Instead, think of these factors working at the level of individuals. In this case the product may be some intermediate product in the firm's value chain. For example, an accountant who specializes in tax accounting gets better and better at performing the tasks that make up this subfield, the more she does. If being employed by a firm does not allow this kind

of specialization, it may make sense for these services to be provided by an independent contractor. Large firms can have activities where economies of specialization is important performed by employees, but small firms may find it better to outsource, since they will not be able to allow employees to specialize sufficiently to quickly move down the ‘learning curve’ of Figure 5.6, or to take continued advantage of such learning after they have accomplished it.

For the above argument to work, it must be that there is a sufficient number of firms that wish to outsource. In other words, if external providers of these specialized services are to take advantage of the economies of specialization, the market for their services must be large enough. Adam Smith put this idea as follows, in 1776: “The division of labor is limited by the extent of the market”. Finally, if there are also economies of scale in providing the outsourced function, a firm rather than an individual may take advantage of these by employing many individuals together.

Economies of Scope Now suppose that there are two different products that can be produced, either separately or together. The output levels are denoted Q_A and Q_B . Production is subject to **economies of scope** if the cost of producing these products separately is greater than the cost of producing them together. Mathematically, we can express this as $C_J(Q_A, Q_B) < C_A(Q_A) + C_B(Q_B)$. If other aspects of production are unchanged across these two possibilities, economies of scope will work in favor of the firm producing both products.

Again, we can think of an analogous argument at the level of individual activities. If a pair of activities can be performed jointly by one person, or separately by two people, economies of scope work in favor of the joint performance of the activities. Economies of scope thus help to define job assignments (what tasks are best done together). They can also work against outsourcing, since outsourcing just one task will lose the benefits of economies of scope, and outsourcing both tasks may require satisfying more demanding criteria. For example, it may mean exposing sensitive company information to outsiders, and not be desirable for that reason.

Economies of scope work against economies of specialization, though both may be present simultaneously. Even in this very simple framework, the implications for the organization of firms are not straightforward, since the two kinds of factors can vary in relative importance. In general, however, simple value chain functions, such as building maintenance, that are relatively independent of the firm’s core business are more likely to be outsourced to take advantage of economies of specialization or economies of scale. At the level of the firm overall, scale and scope are two defining characteristics of the nature of firms, and how they evolve. Horizontal integration, through mergers of firms producing the same product, can be thought of as attempts to take advantage of economies of scale. Diversification strategies, especially into related product lines, can be thought of as attempts to take advantage of economies of scope. In particular, different products may share portions of the value chain, such as brand management or marketing in general, or distribution. These activities may themselves be subject to economies of scale or scope, but they allow economies of scope to operate at the level of

the firm. This is illustrated for a simple value chain, which is a modification of that of Ekin, our mythical shoe company, in Figure 5.7. The company now produces sports shoes and sportshirts. These have different production requirements, but they can benefit from shared brand management and distribution efforts. These perspectives on economies of scale and scope will be useful as we examine the evolution of firms in e-commerce, in Chapters 10 and 11.

Figure 5.7: Economies of Scope in the Value Chain

Production operations: shoes	Combined brand management	Combined distribution
Production operations: shirts		

5.4 The Boundaries of the Firm

We now delve deeper into the existence of firms, and why their functions are not performed solely through market transactions. The answer has many dimensions, and no single explanation will suffice to capture the diversity of organizational forms. Clearly the benefits of ongoing activity, as illustrated in one way by dynamic economies of scale, are somewhat more easily captured by a continuing organization. On the other hand, long term contracts might achieve the same benefit. A further argument is that contracts may not be precise enough to govern all the contingencies that might arise over time. In a firm, residual authority resides at various levels of a hierarchy, and explicit contracts may not be able to substitute for the flexibility this provides. This ‘residual authority’ is related to concepts of ownership and control over assets. Efficient use of these assets may not be possible through market transactions.

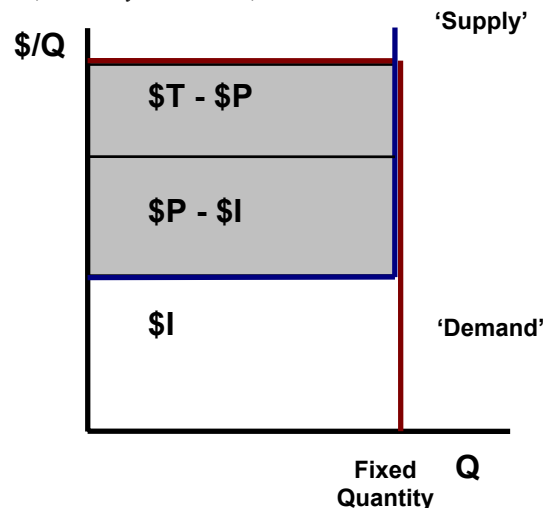
Digging still deeper, the reasons for the inability to use market contracts with perfect efficiency must lie with imperfections in information. These imperfections may be with respect to abilities, other characteristics, actions, or performance. The same kinds of imperfections may apply within a firm, but they may have different relative importance, and the set of feasible contracts that provide incentives for more efficient actions may be different within firms than in arm’s-length, market transactions. This section outlines these themes to understand the boundaries of firms, and how they may change as the economy changes.

Assets We begin by making two changes to our earlier descriptions of firms. Unlike our discussion of economies of scale and scope, we now consider revenues as well as costs, to focus squarely on the creation of value. Unlike our review of firms in Chapter 4, we consider the firm's inputs as assets, rather than just services of these assets that are contracted for period by period. This is important when assets are specific to a particular production activity and investment in the assets is irreversible. By an **asset** we mean anything that can generate a stream of returns over time, including physical equipment as well as intangibles such as knowledge and skills.

Consider a productive activity somewhere along the value chain, that produces an input into the final product. For example, this may be a computer hardware component, or an automobile engine or body component. We first consider a situation where the producer of the component is a separate firm from the user. Therefore the user buys the part from the producer, and we refer to them as the Buyer and Seller respectively. Initially, suppose that the total amount to be produced is fixed, so we will be able to focus only on investment and prices paid, without having to consider output decisions. Suppose that, in order to produce the part, the seller has to make a fixed, irreversible investment, which has a per period cost of $\$I$. Again to keep it simple, assume that there are no other costs of production. Suppose that the total value of the parts is $\$T$. The net value created is then $\$V = \$T - \$I$.

The amount $\$V$ is also the total (economic) rent in the relationship. In the Disney example of Chapter 4 (Figure 4.23b), this rent was distributed between the buyers and the seller by the willingness to pay of the marginal buyer. The seller's economic profit was its share of the rent (the rectangle in Figure 4.23b), while the buyers' share of the rent was the sum of all the differences between their willingness to pay and the price they paid (the area above the rectangle and below the demand curve). In the current example, the Buyer and the Seller do not have a unique price that equates supply and demand. The Seller can potentially accept any amount greater than or equal to $\$I$, while the Buyer is willing to pay anything up to $\$T$. If the agreed payment is $\$P$, then the Seller's rent is $\$P - \I , while the Buyer's rent is $\$T - \P . These two quantities add up to $\$V$. This situation is illustrated in Figure 5.8, which also shows that the 'supply curve' and 'demand curve' do not have a unique intersection, so there is no unique equilibrium price.

Figure 5.8



The payment, $\$P$, for the component will be determined in this case by the relative bargaining power of the Buyer and Seller. If the Buyer can use any one of several competing sellers, then $\$P$ may be as low as $\$I$, which gives the Buyer all of the net value created, $\$V$. This is an extreme case of value capture, much as we discussed in considering competition and monopoly in Chapter 4.

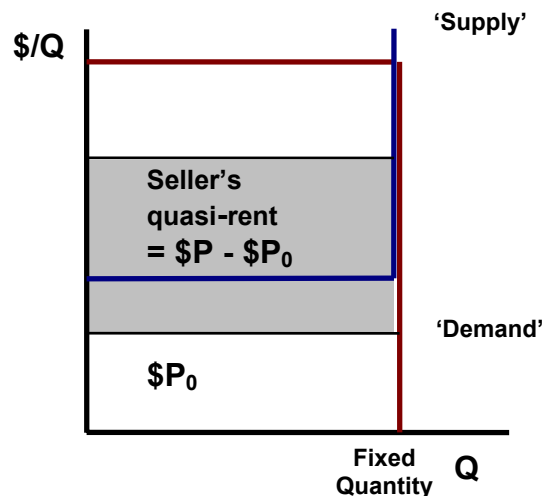
Transaction Costs We now illustrate a problem that may arise with this kind of market transaction. The sources of the problem are (1) the specific, irreversible investment by the Seller, that we have already introduced, and (2) the new assumption that the contract for payment $\$P$ is not strictly enforceable. The second assumption is an example of a **transaction cost**. Suppose that, after the Seller makes the investment, the Buyer claims that it needs to renegotiate the payment, and that it has a reason (weak demand for the product, potential low quality of the part, and so on) that will be difficult and expensive to challenge in court. Therefore, we assume that the Seller has no effective legal remedy.

If the Seller has no alternative use to which it can put the investment, the Buyer can reason as follows. The Seller's investment is a sunk cost, and therefore the Seller should be willing to accept any payment above zero for the component. Hence $\$I$ is no longer a floor on the payment, once the investment has been made. If the Buyer has enough bargaining power, it can get away with paying less than $\$I$, and saddling the Seller with a loss. If the Seller anticipates this opportunistic behavior by the Buyer, then it may not even enter into the transaction and make investment to begin with.

One way to avoid this 'hold-up problem' is to integrate the Buyer and the Seller into one firm. The payment for the component is then a transfer within the firm, and is irrelevant. The motive for opportunistic behavior disappears when a within-firm transaction replaces the market transaction. Thus this analysis provides one way of resolving the 'make-or-buy' decision. Other kinds of transaction costs have also been identified as favoring this outcome of **vertical integration**. These are costs of market contracting associated with lack of information required to specify contingencies or performance, or the ability to monitor and enforce contract performance. Ronald Coase

and Oliver Williamson have emphasized these kinds of transaction costs as leading to the existence of firms, rather than using markets for all economic exchange. The same explanation also helps define the boundaries of firms, as above. Before we provide counter arguments, we briefly describe how the above example can be generalized.

Figure 5.9



Suppose that the Seller is not totally without alternatives. It may, for example, be able to find an alternative buyer who will pay $\$P_0$. If $\$P_0$ is greater than $\$I$, then $\$P_0$ becomes the new floor on the payment the Buyer must make, either before or after the seller's investment. There is no room for opportunism. If, on the other hand, $\$P_0$ is less than $\$I$, then the Buyer can still behave opportunistically, trying to capture the post-investment surplus of $\$T - \P_0 . This quantity, $\$T - \P_0 , is called a **quasi-rent**. It is independent of the investment, $\$I$, which is now a sunk cost that can not affect future decisions. Recall that the rent was the net value created in the relationship, $\$V = \$T - \$I$. We can also distinguish the total quasi-rent, $\$T - \P_0 , from the Seller's quasi-rent, which reflects the initial contract, and is therefore $\$P - \P_0 (see Figure 5.9). Our earlier example was simply a special case, with $\$P_0 = 0$.

Another possible generalization of the hold-up problem has the Buyer also making relationship-specific investments. This now creates possibilities for opportunism on both sides, and merely reinforces the argument for vertical integration, already made. An example of such an investment by the Buyer would be in designs that use the component. Again, alternative sources of supply would limit the Seller's scope for post-investment opportunism, but might not eliminate it entirely. A third generalization of the example allows the production level of the component and the level of investment to be variables. Here again, opportunism still is a problem, though it now may lead to underinvestment rather than a complete breakdown of the market transaction.

The transaction cost approach to explaining the existence and boundaries of firms has been subject to the following counter arguments. First, it focuses on individual

transactions, and thus does not capture the concept of a firm as an entity that performs many different tasks over time. Second, repeated contracting and reputation effects will reduce the incentive for opportunistic behavior, by making it costly in terms of future losses. Thus, even in the absence of legal enforceability of some contract, markets may work fine. Japanese subcontracting relationships are a well-known example of this market-mediated enforcement mechanism. Nucor, a very successful US steel-maker that uses scrap in ‘mini-mills’, also maintains long-run supplier relationships without vertically integrating.² Finally, the transaction cost analysis assumes that within-firm transactions are not subject to opportunism or other incentive problems. As Harold Demsetz (1991) has emphasized, there is not a clear practical distinction between general transaction costs and management costs. In the Buyer-Seller example above, both the case of separate firms and the integrated case will involve costs of managing production and exchange. The relative magnitude of the sums of all these costs is what should matter for the boundaries of the firm. Treating this issue requires a more explicit examination of information and incentives, and we come to this after discussing ownership of assets, and coordination as factors in defining firms.

Asset Ownership The generalization of the transaction cost model to the case of variable investment is similar in spirit to a related approach to understanding firm boundaries. We will use the same basic framework as before. Suppose that the gross value of the parts that are produced, $\$T$, and the Seller’s alternative payment, $\$P_0$, may depend (positively) on the level of investment, I . This investment level is now a choice variable. Furthermore, it is not a value that can be specified in a contract, as the transaction cost approach sometimes assumes. To justify this, think of the investment as being more than physical plant and equipment. It also includes the knowledge acquisition by the Seller, or its employees, and this can not be specified precisely in a contract.

Suppose that the Buyer also has some alternative source, which will be worth a total of $\$T_0$. In the best possible outcome for the two entities, the investment I is chosen to make the net surplus or rent, $\$T(I) - \I , as large as possible. This will be when the marginal addition to T of an extra dollar of investment is itself a dollar. This is a special case of the general “marginal benefit equals marginal cost” rule. On the other hand, if the two parties bargain after the investment has been made, the surplus, or quasi-rent they will bargain over is $T(I) - P_0(I) - T_0$. Each gets its alternative payoff, plus its share of the quasi-rent. Suppose for simplicity that they split this quasi-rent equally. Then the Seller’s share of the quasi-rent is $\frac{1}{2} (T(I) - P_0(I) - T_0)$, and the net value captured by the Seller is this quantity plus $P_0(I)$, minus the investment, or

$$V_S = \frac{1}{2} [T(I) + P_0(I) - T_0] - I.$$

The Buyer receives the remainder of the economic rent, so captures net value $T(I) - I$ minus the Seller’s share, or

$$V_B = \frac{1}{2} [T(I) - P_0(I) + T_0].$$

You can easily check that $V_S + V_B = T(I) - I$.

² Both these examples are discussed further in Holmstrom and Roberts (1998).

The Seller, in choosing the level of investment, I , calculates its marginal benefit of increasing investment based on averaging the marginal benefit within the relationship, $\Delta T(I)$, and the marginal benefit of the alternative, $\Delta P_0(I)$. This may or may not lead to underinvestment relative to the level that maximizes the total net value created. At the margin, the value captured and the value created do not match, and this is the source of the possible inefficiency. It is somewhat analogous to the problem of the monopolist (Figure 4.21), for which the marginal revenue (marginal value captured) of production and sale is less than the marginal willingness to pay of buyers (marginal value created).

The problem is much worse if the Buyer makes the investment decision. The Buyer's marginal value captured is $\Delta T(I) - \Delta P_0(I)$, which is clearly lower than the marginal value created, $\Delta T(I)$. In this example, it is better for efficiency for the Seller to own the asset measured by I . If the assumptions are changed (for example, if I affects T_0 more than it affects P_0), then it may be better for efficiency if the Buyer owns the asset. Furthermore, if there is more than one asset, it is possible that one or other party in the transaction owns the entire group of assets. Ownership of the group of assets in this case represents a firm. In such an example, the nonowner of assets may be thought of as a contract employee.

This description of firms as combinations of asset ownership rights (due to Sanford Grossman, Oliver Hart, and John Moore) captures the idea that the key right provided by ownership is the right to exclude others from using the asset. The owner of assets thereby gains partial authority over those who use the assets for productive purposes. The asset ownership approach to explaining the existence and the boundaries of firms is subject to similar counter arguments as the transaction cost approach. It assumes that alternative incentive mechanisms to ownership are not available, when repeated relationships and reputation may serve quite well. Incentive issues within the firm are not really addressed in the asset ownership approach, though asset ownership provides greater scope for incentive provision. Finally, asset ownership may be separate from full vertical integration. For example, dies used in making some car parts are owned by the automobile manufacturer, even though they are located and used at the supplier's plant.

Coordination One factor that favors ownership of a group of assets is complementarity in the returns generated by those assets. Formally, this means that increasing the amount of one asset increases the marginal returns of an other asset. In this respect, the formal idea of complementarity is mathematically the same as complementarity in the inputs of the standard production function, $F(L, K)$, as well as the one underlying the idea of strategic complements (Chapter 4, Figure 4.12).

In the extreme case of complementarity, assets are worthless unless used together. The asset ownership theory of the implies that such assets should be owned together, and not separately, for efficient allocation. In this sense, coordination of assets is a useful way to think of the organization of firms. If we return to the value chain, and think of it being even more detailed, so that different stages of the production process itself are

spelled out, it seems reasonable to assume that stages of production that are close in terms of timing will be located together, and the assets required for these contiguous stages will be owned by one firm. In the absence of one stage of production, the output can not be produced at all, so this is an example of perfect complementarity, and the need for coordination at a very basic level.

Information and Incentives In the discussion of transaction costs and asset ownership theories of firm structure, incentives for efficient investment were a key idea. Different patterns of ownership, or the choice between market contracting and within-firm decisions were based on the relative effectiveness of incentive provision in the two cases. The problems of incentive provision could, in turn, be traced back to problems of information, with respect to measuring or verifying effort, performance or underlying factors that could affect performance.

In a world where the problems of imperfect information, and consequently imperfect incentives, are pervasive, the choice between allocation via firms and allocation via markets can depend on their relative effectiveness in incentive provision. For example, where control over assets is important for appropriate incentives, this may be achieved by contractual assignment of control rights, without transferring ownership. Exclusive dealing contracts, and licensing agreements are examples of this possibility. Holmstrom and Roberts (1998) emphasize the pervasiveness and sophistication of what they call **governance contracts**. Firms in this case are at the “center of a network of relationships, rather than ...owners of a clearly defined set of capital assets.” In some respects, this idea is actually a move toward the standard textbook model of the firm, which flexibly hires the services of assets as needed, to produce as needed. The difference lies in the detailed performance of those assets, where incentives must be provided: simple fixed per unit payments will not do. This idea is also similar to the recent concepts of the **virtual corporation** and the **network organization** (see Illustration Box).

Illustration Box

The Network Organization and the Virtual Corporation

In the early 1990's, well-known management guru Tom Peters pushed the idea of a network organization, using long-term relationships with suppliers to substitute for vertical integration, much as in the case of Japanese subcontracting relationships.

William Davidow and Michael Malone suggested an even more radical dismantling of firm hierarchies, what they called the virtual corporation. Even before the rise of the Internet, they identified advances in information processing, telecommunications, and computer-aided design and manufacturing as factors that allowed coordination to be achieved without integration of ownership. What was left unanswered in their discussion was the issue of relative effectiveness of incentive provision, and some of the more basic transaction costs of market contracting.

Sources:

Peters, T., *Liberation Management: Necessary Disorganization for the Nanosecond Nineties*, New York: Knopf, 1992.

Davidow, W.H., and M.S. Malone, *The Virtual Corporation*, New York: HarperBusiness, 1992.

Besanko, D., D. Dranove, and M. Shanley, *Economics of Strategy*, 2nd ed., New York, Wiley, 2000.

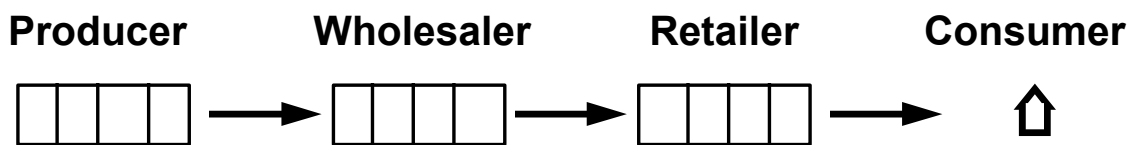
Other cases of firm structures as methods of incentive provision include the assignment of tasks to employees, such as sales people. Holmstrom and Milgrom (1991, 1994) use this example, where sales people carry out three tasks: current sales, developing long-term customer satisfaction, and collecting and communicating information on customer needs. While the latter two activities are hard to measure, they are important, and the need for balanced incentives may require weaker incentives for current sales. Otherwise current sales will be stressed to the detriment of the activities with longer-run benefits to the firm. These weak incentives may not work for non-employee sales people, however, because they may engage disproportionately in other activities, including selling other firms' products. An employee is constrained in this respect, more than an independent contractor, and she or he can also be given incentives for career advancement not available to the outsider. In sum, the need to provide incentives across a range of activities pushes out the boundary of the firm, allowing the use of a set of complementary incentive mechanisms. This explanation extends the asset ownership approach, rather than replacing it.

A final example of incentive provision comes from comparing fast food restaurants and grocery stores. Fast food restaurants are often franchises. In other words, they pay the parent corporation for the brand name, advertising, and basic business process guidelines, but they own and manage the physical assets and hire the employees directly. Grocery store chains, on the other hand, are more vertically integrated, and do not use franchises. The difference, according to Maness (1996), does not lie in asset specificity or non-contractible investments. Instead, cost control and service are difficult to monitor and reward in the fast food restaurant case, leading to decentralized ownership as a means for providing strong incentives. In the case of grocery store chains (and department stores), inventory control is more important for cost control, and it can be more easily monitored and controlled centrally. This favors vertical integration, since store managers do not need the same incentives as fast food restaurant operators. Note that franchisees can still enjoy the benefits of economies of scale in areas such as advertising and input purchasing, so this traditional reason do not explain the difference in firm boundaries.

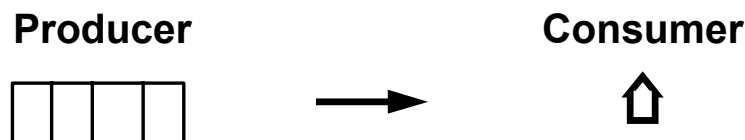
Information also matters for firm boundaries beyond its role in shaping incentives. Information technology and more basic telecommunications tools such as the fax machine have allowed the coordination and scheduling of related production activities across space and time. Computer-aided design and manufacturing have helped reduce the switching and setup costs associated with changes in product designs or production runs. These technological changes have tended to reduce economies of scale and scope, and therefore act in favor of smaller, more specialized firms. Just-in-time inventory management is also made possible by the same technological advances, and this, too, permits firms to operate along only a part of the value chain.

E-Commerce and Firm Structure E-commerce, in its B2B manifestation, aims to make market contracting more transparent and more efficient. Direct savings in transaction costs through reduction in the rents extracted by toll-taking intermediaries (see Chapter 4), better matching (Chapters 8 and 11) and reductions in the cost of adjustment in response to changing external conditions all appear to make market transactions more efficient, tilting the scales toward more disintegration of firms along the value chain, and even reconfiguring some value chains and eliminating others (see Figure 5.10 and Chapter 11). However, the same forces are at work for within-firm transactions. Hence, it is not immediately clear that markets are favored by online transactions. At the same time, the broader technological changes described in the previous paragraph do seem to support less integrated firms. Where markets do have an advantage is in transparency and competition. If B2B exchanges provide more competition, more repeated transactions, and easier reputation-building, hold-up problems and other forms of opportunism may diminish in importance. That will indeed favor markets over within-firm transactions. We return to these possibilities in Chapter 11.

Figure 5.10
Intermediaries



Disintermediation: Reconfiguring Value Chains



5.5 Location and Distribution

Figure 5.10 shows a traditional sequence of firms bringing the product to the final consumer. In Chapter 8, we shall examine the role of wholesalers and retailers in greater detail, as part of a general discussion of intermediaries. Here we focus on the location,

transportation and distribution decisions that the movement of a product from manufacturer to wholesaler (which we shall treat as equivalent to ‘distributor’) to retailer represents. Traditional economies of scale in production are clearly an important starting point for looking at location. Economies of scale were an important reason why production moved out of the household in the first place. The shift from cottage production of textiles (the ‘putting out’ system) to factory production during the Industrial Revolution is an example. To connect this with our discussion of the boundaries of the firm, note that the different needs of monitoring and the different scale of assets caused independent contractors to be replaced by mill employees. Thus changes in the needs of asset ownership and incentive provision shifted the boundaries of the firm, along with the effects of economies of scale and concentration of location.

In short, firms are larger than households, and their location is necessarily more concentrated. Therefore moving products from manufacturers physically closer to households is an important aspect of the economy. Wholesalers and retailers serve this function, as well as simultaneously aggregating ranges of products. Depending on the kinds of factors discussed in the last section, wholesalers and retailers may be independent firms, as shown in the top of Figure 5.10, or they might be owned by the manufacturer, or the retailer and wholesaler may be integrated, as in the example of Wal-Mart, which owns the distribution warehouses that supply its retail stores.

The bottom panel of Figure 5.10 shows an extreme alternative to the traditional three-stage sequence. This is direct sale by the manufacturer to the consumer. This is probably the most popular example of the impact of e-commerce. Of course it predates the Internet. For example, Dell Computer began direct sales over the telephone, years before it took to the Web. Its direct sales experience made the move online natural, of course. We will discuss the factors that made direct sales possible in Chapter 11: they included cheap communications and computerized inventory management. Here we emphasize that the logistics of direct delivery are different from those of the manufacturer-wholesaler-retailer model. In the latter case, typically trucking companies, either independent or owned by the buyer or seller, move the products from one location to another. In the case of direct delivery, it is companies such as United Parcel Service (UPS) that are set up to make house-to-house deliveries. Not surprisingly, UPS’s model is being imitated and complemented as e-commerce increases the ease and popularity of direct sales.

In the traditional distribution framework, retailers have chosen where to locate based on the locations and characteristics of households. For example, Wal-Mart is well-known for having initially pursued a strategy of targeting small towns in the southern United States. These towns had been considered by existing retail chains to not have enough purchasing power to make location there worthwhile. However, when Wal-Mart started on this course, the region was about to begin a substantial economic expansion that helped make the company successful. Wal-Mart located its distribution centers in the center of groups of towns, and built the stores that would be served by these centers to circle them. Location decisions were crucial to Wal-Mart’s success.

We can illustrate some of these ideas quite simply³. Suppose that customers are located along a straight road, denoted by the straight line in Figure 5.11a. There are seven customers, whom we denote by the letters A through G (this is to protect their privacy). Each customer demands exactly one unit of the product, and all pay the same price for it. Transportation costs are proportional to distance traveled. Where should the seller locate to minimize transportation costs (and hence maximize profits)? It turns out that the best location for the firm is where the middle customer, Ms. D, is located. Moving a little bit away from this location (but staying on the road) increases the distance traveled by four customers by that amount, and reduces the distance traveled by only three customers by the same amount. The move raises total distance traveled and the total transportation cost. Clearly moving off the road would also be worse. The importance of the customer located in the middle, relative to the others, depends on the fact that they are otherwise identical. If Ms. A and Mr. B were much more valuable customers, locating near them might be optimal.

Figure 5.11a



Figure 5.11b

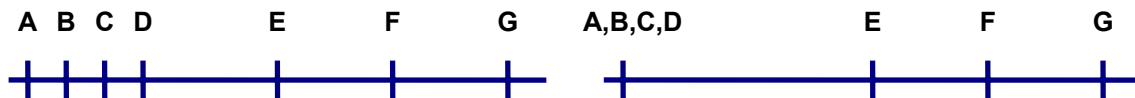


Figure 5.11b shows a slightly different situation, where consumers A through D are all in the same location. Now the firm will still choose to locate where the middle customer is, but this happens to be where four people are located. Economizing on transportation costs reinforces the concentration of location, and this idea is the germ of the growth of towns and cities. We can think of the joint location of the four customers in Figure 5.11b as an urban area. In the situation of Figure 5.11b, if any one of the four urban consumers moves to the ‘suburbs’ (to the right along the line), the firm will move as well, assuming that any moving costs are eventually outweighed by the savings in transportation costs.

Other factors may cause consumers and firms to locate in clusters. In particular, being close may have benefits beyond transportation cost savings. Cost reductions for firms or benefits to consumers may arise, not from the size of any single firm or household, but the overall size of the industry or the local market. This happens because increases in overall size make economies of scale possible for other activities. Once an urban area is large enough, it can support major league baseball, NFL or NBA teams, as well as cultural activities like the theater and the symphony. Input suppliers of a firm are able to achieve economies of scale more easily, and knowledge sharing and monitoring of trends is also easier. Besides the examples of Silicon Valley and its many high-tech clones, fashion designers in Milan and entertainment industry firms in the Los Angeles area are examples of industries where clustering makes following trends and getting skilled workers easier. These effects are called **economies of agglomeration**, and are a

³ The following examples are adapted from McDonald (1997).

special type of **external economies**, because they are external to any single firm or individual. Many of these effects are likely to persist well into the evolution of e-commerce, since they are driven by physical proximity.

We have focused on the costs for the firm of getting the product to its customers as the determinant of location. This assumes that there are no other differences as a result of location. The external economies argument, though, is one justification for why input costs may differ in different locations. This argument is more general. The costs of labor, capital, energy, raw materials and other inputs may all differ across locations, and have an influence on where firms decide to locate. Taxes and publicly-provided infrastructure may also vary from place to place, leading firms to choose one location over another. Movements of manufacturing from the US to Asian countries, or the choice of the American South over other US locations for opening Japanese automobile manufacturing plants are examples of location decisions affected by such factors.

The costs of transporting raw materials may also be balanced against the costs of transporting the finished product to consumers. The nature of the product, and whether it gains or loses weight in processing, will affect the outcome of this balancing act. Lumber processors will probably prefer to locate relatively close to sources of their raw material, since the finished wood is lighter than the logs, while Coca Cola is better off shipping its concentrated syrup to bottlers around the country, rather than adding the water first. Input and output transportation costs can easily be incorporated in the standard economic analysis of the firm, as sketched in Chapter 4. For example, the cost curve can be thought of as including transportation costs, with a different cost curve for each location the firm might choose. The location with the lowest overall costs (also maximizing profits) will be the one selected.

How have changing transportation costs affected the location of economic production activity? Physical transportation costs have fallen with better roads, greater efficiencies in loading and unloading products, and better transportation infrastructure in general, resulting in greater possibilities for geographic separation of production and consumption. However, the Internet and e-commerce do not directly affect the economics of such location decisions. They do make long-distance information flows less expensive and more effective, further allowing geographic separation, but they do not change the costs of moving physical products around. Thus the ‘death of distance’ has its limits, as we shall explore in Chapter 18!

5.6 Industry Structures and Strategy

In this chapter so far, we have added the following dimensions to Chapter 4’s story of what firms do: a disaggregation of the stages of productive activities (the value chain), a discussion of economies of scale and scope (the latter introducing the possibility of a firm producing more than one output), a consideration of how firm boundaries are drawn within the value chain, and some economics of how firms decide where to locate.

In this section, we use some of these elements to discuss the structure of industries, and the implications for firms' strategic decision-making. The next section continues the overview of firms' strategy options, building on the concepts introduced in Chapter 4, and laying more groundwork for the analysis of e-commerce strategies in Part III.

In Chapter 4, we explained briefly how economies of scale (increasing returns to scale) relative to the size of the market affect the structure of an industry. In that case, we assumed that the market was defined to include all demand for a homogeneous product. In practice, products are rarely completely homogeneous at the level at which general economic analysis is to be conducted. For example, we typically consider the domestic airline market and industry as a unit of analysis, even though the services offered are thousands of different flights, differentiated by starting point, destination, time of flight, and stops en route. The airline example also illustrates a difference between the terms 'market' and 'industry'. While all US airlines consider themselves as in the same industry, they compete differently in different markets, typically defined by geography. For example, the Boston-New York route constitutes a 'market' in this sense. A slightly broader market would be 'short-haul flights in the Northeast'. When one looks at industries, therefore, encompassing related but different products, economies of scope will go hand-in-hand with economies of scale in determining industry structure,

Since providing air travel involves investing heavily in capital equipment (airplanes), and since there are substantial economies in having a sizeable fleet flying many routes, one would expect the airline industry to be relatively concentrated. This is borne out by the figures in Table 5.2. The four largest airlines had a passenger revenue share in 1997 of 64%, while the share of the six largest was 79.2%. In fact, on particular routes between pairs of cities or regions, and in terms of flights into or out of particular cities, the concentration is even higher. For example, in 1998, Delta's share of seats into an out of Cincinnati was 90.8%.

Table 5.2: US Airline Revenues and Share

Airline	Passenger Revenue	
	Amount (\$ billion)	Share (%)
United Airlines	15.1	19.0

American Airlines	14.3	18.0
Delta	12.8	16.1
Northwest	8.7	10.9
US Airways	7.1	8.0
Continental	5.7	7.2
Southwest	3.6	4.5
Trans World	2.9	3.6
America West	1.8	2.3
<u>Alaska</u>	<u>1.3</u>	<u>1.6</u>
Industry Total	79.5	100.0

Source: *The Structure of American Industry*, Ch. 8, ed. Walter Adams and James W. Brock, 10th edition, New Jersey: Prentice Hall, 2001. Original data from Air Transport Association, *Annual Report*, 1998.

Note that while firms' revenues are a natural measure of their importance within the industry, revenues do not provide much insight into other aspects of industry structure. They do not tell us how integrated firms' value chains are, nor where the sources of value added are. The hypothetical example of the shoe manufacturer, Ekin, in Section 5.1 was precisely designed to examine those issues. Because airlines are always the final sellers, some of these issues do not enter into their industry. In the example of Ekin, however, the firm's revenue of \$1.81 million included sales at three different points along the value chain. In an example from another industry, an integrated oil company's revenue may include sales of crude oil, sales of various refined products to other businesses, and sales of gasoline and home heating oil to individuals through its own channels. Each of these categories lies at a different part of the value chain. The firm's revenue share of the overall petroleum market may not reflect its share of crude oil production, different refined products, or distribution of some products to consumers. Similarly, lack of concentration of revenues at the aggregate level may hide dominance at particular points along the value chain.

The value chain thus not only gives us a picture of the firm's activities, and a basis for an accounting of value added at various stages of the chain, but it points the way to identifying what we may call potential 'choke points'. This identification, in turn, becomes an important component of strategic thinking. Nowhere are these issues more important than in the 'new economy', including not just e-commerce, but all aspects of technology and information industries. The complexity of the value chains in these industries, the importance of complementarity of components, and the rapid technological change that these industries are undergoing all make this perspective on industry structure crucial for strategic thinking.

We will provide some basic facts on the computer industry. At one time, this term was easy to define. Computer makers were vertically integrated: hardware, software and after-sales service all were provided by the same company. In the 1960s,

IBM's market share of this monolithic industry ranged from 65.3% to 71.6%.⁴ IBM used its strength in some parts of the value chain to ensure that it maintained dominance of the whole chain. For example, while the possibility of separating the hardware and the software did not really exist at that time, due to technological complementarities, IBM used pricing and design strategies to prevent entry in after-sales service and peripherals such as tape drives.⁵ Rapid technological change, and IBM management's failure to adequately adapt their strategy in response, led to a substantial erosion of IBM's overall market share.

Thirty years after IBM's heyday, the computer industry landscape has changed dramatically. The industry has grown phenomenally, and this has supported the emergence of greater specialization along the value chain. Three major segments of what we can call the 'information technology (IT) sector' are hardware, software and services. IBM remains as one of the few integrated IT firms, and its revenues still place it in the lead, but it lags in profit and stock market value. The 1996 data presented in Table 5.3 list the top ten IT firms in terms of revenue but, by 1999, this picture had already changed dramatically. There were numerous other IT firms besides the ten listed here, but even as a share of the top 20 firms, IBM's revenue was less than 25% by 1996. Thus the industry has become considerably less concentrated.

Of the firms in Table 5.3, Compaq specialized in hardware, Microsoft in software, and EDS in services. Thus these companies staked out parts of the overall computing value chain. Digital Equipment was more vertically integrated, and in 1998, Compaq acquired it in an attempt to build a one-stop computing shop to challenge IBM. It is not clear, however, that this strategy has been particularly successful. For example, integrating Digital's service staff into Compaq has not been easy. This issue can again be thought of in terms of the value chain. Compaq's acquisition of Digital did not necessarily give it corporate assets with enough unique value. The company that controls the choke point in the value chain for much of the industry is in fact Microsoft, with its dominance of desktop operating systems. Microsoft has resisted vertical integration out of software, but it has expanded its product line with complementary products for the desktop market (applications software such as Microsoft Office) as well as products for new markets (server operating systems), building on its control of desktop operating systems (85-90% of that market). This control of a value chain choke point is hidden in the broad sectoral picture of Table 5.3. A final interesting fact that Table 5.3 illustrates is the greater horizontal diversification of the Japanese firms in the list: Fujitsu, Hitachi, NEC, and Toshiba. Of the US firms in the list, only HP was not 100 percent in IT, and that has now changed with its spinoff of Agilent, separating out its scientific instruments business.

Table 5.3: Global IT Firms' 1996 Revenue

⁴ Source: *The Structure of American Industry*, Ch. 6, ed. Walter Adams and James W. Brock, 10th edition, New Jersey: Prentice Hall, 2001.

⁵ *ibid.*

Rank	Company	IT Revenue (\$ billion)	IT Revenue as % of Total Revenue
1	IBM	75.9	100
2	Hewlett Packard	31.4	80
3	Fujitsu	29.7	63
4	Compaq	18.1	100
5	Hitachi	15.2	23
6	NEC	15.0	34
7	EDS	14.4	100
8	Toshiba	14.1	24
9	Digital Equipment	13.6	100
10	Microsoft	9.4	100

Source: *The Structure of American Industry*, Ch. 6, ed. Walter Adams and James W. Brock, 10th edition, New Jersey: Prentice Hall, 2001. Original data from *Datamation*, July 1997.

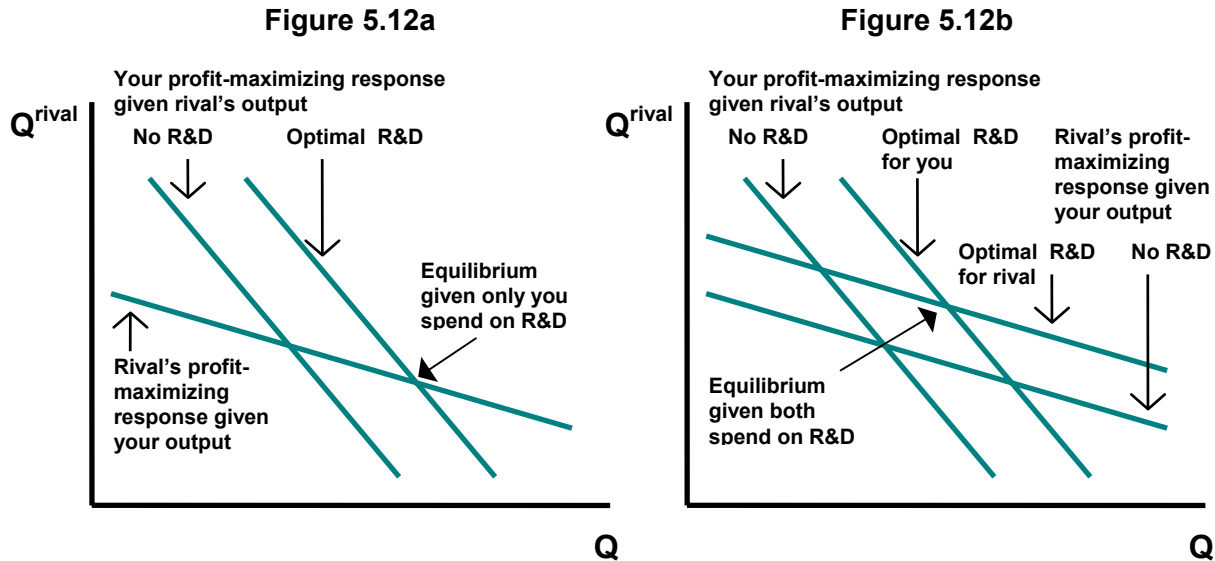
Ultimately, industry structure is determined mainly by two factors. One is the extent of economies of scale and scope (Section 5.3). The other is the control over unique assets (Section 5.4). If these unique assets are complementary to other assets, either in the same stage of the value chain, or at different points, then this control provides leverage. In the case of Microsoft, the fixed costs of R&D and marketing provide economies of scale and scope. Its intellectual property (the source code for its software) is Microsoft's unique asset.

The greater the extent of economies and control over unique assets, therefore, the more likely it is that the industry will be concentrated. Most industries are somewhere between the extremes of monopoly and perfect competition. Where they lie between these extremes depends on kinds of factors highlighted in this section, including the structure of the value chain, and the nature of competition along its stages. Since economies of scale and control over unique assets can be affected by firms' actions, they form a basis for focusing firms' strategic decision-making. We take this up next.

5.7 Dimensions of Firm Strategy

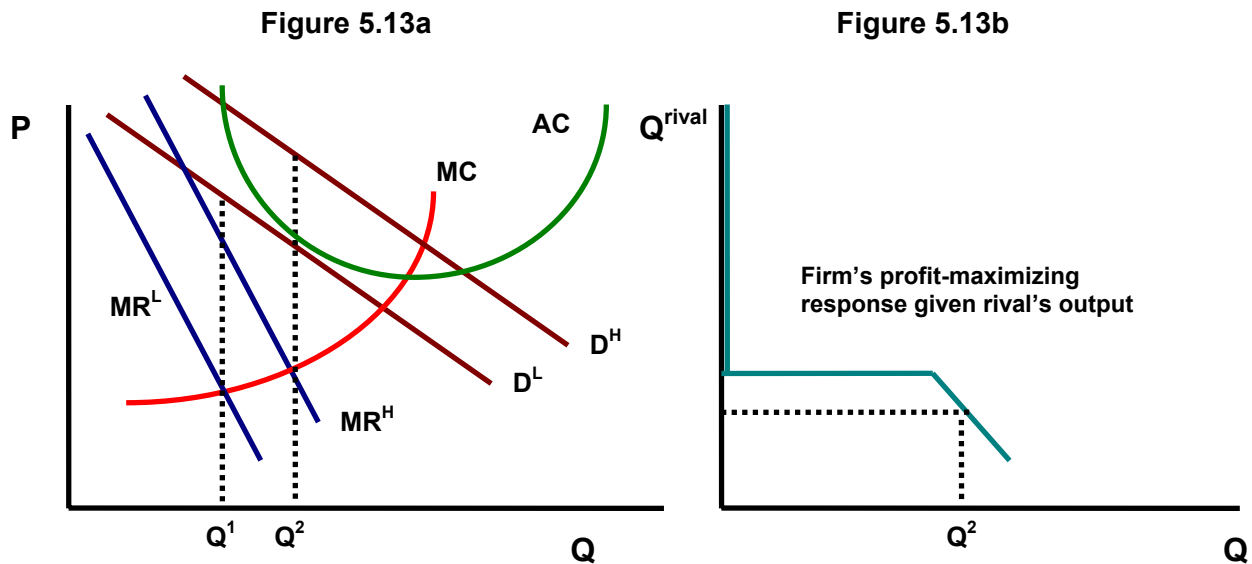
In Chapter 4, we introduced a basic approach to firms' strategic behavior and its outcomes. The key classification of firms' actions was into strategic substitutes (one firm's aggression caused its rival to accommodate) and strategic complements (a rival would match aggressive behavior). The examples of strategies we gave in Chapter 4 were choices of output or of price. The value chain view of firms confirms what we observe in practice, that they have a whole range of strategic choices to make, including finance, operations, distribution, marketing and R&D.

Many of these aspects of strategic choice will be taken up in Chapters 13 through 17, in the context of e-commerce specifically. Here we indicate how the framework of Chapter 4 can be used more broadly. Typically, other strategic variables besides price and quantity can be classified as strategic substitutes or complements. For example, increased expenditures on marketing and advertising by a firm may increase the marginal benefit of advertising for its rival. In this case, the rival will respond by increasing marketing expenditure itself. In this case, advertising expenditures, like prices in the example in Chapter 4, are strategic complements (see Figure 4.12). Just as there are price wars, there can be marketing wars as well!

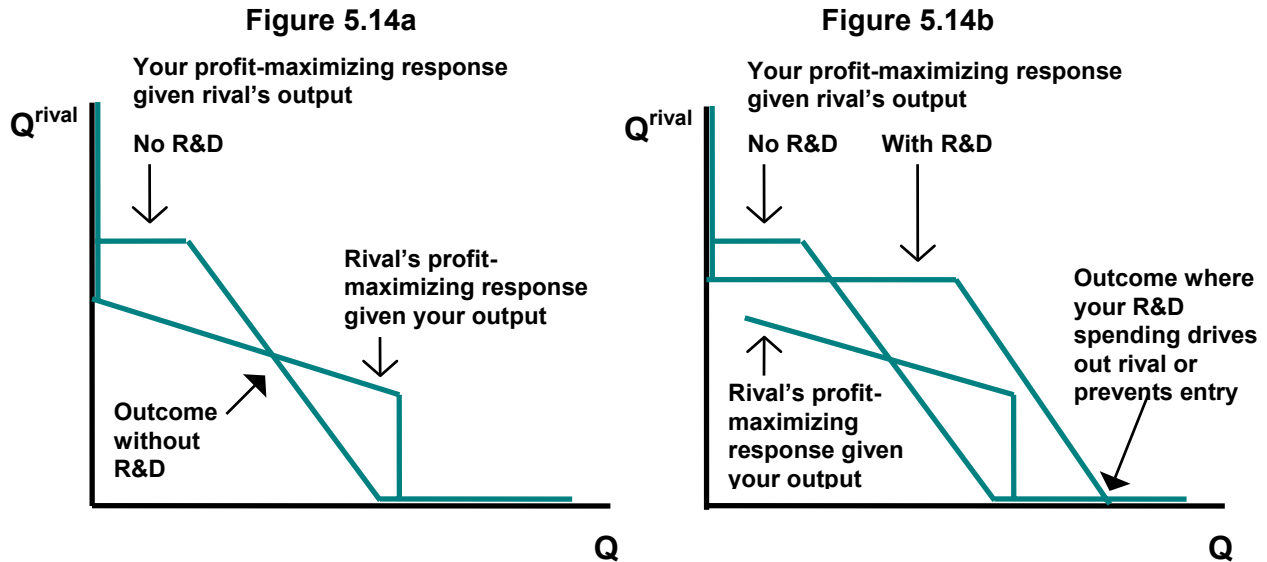


The fact that firms choose more than one strategic action can also be illustrated. Suppose that firms invest in cost-reducing R&D before they decide on their output levels. Thus R&D choices are followed by strategic decisions on production: the outputs are the strategic variables. Let us assume that if a firm spends more on R&D, it reduces its marginal costs. Thus this R&D results in **process innovation** rather than **product innovation**. A reduction in marginal costs increases the firm's profit maximizing output, whatever the rival happens to be producing. Thus, the firm's profit-maximizing response curve shifts outward as a result of its R&D (Figure 5.12a): the firm effectively behaves more aggressively. The expenditure on R&D makes the firm more aggressive in its competition with the rival, and increases its profits. Of course, the costs of the R&D itself will place a limit on what the firm does. The shift in Figure 5.12a assumes that the firm chooses its best level of R&D, where the marginal increase in profit equals the marginal cost of R&D. Figure 5.12b shows what happens when the rival has the same kinds of options available to it. The rival's profit-maximizing response also shifts out, as it becomes more aggressive. Both firms may end up with lower profits as a result of their trying to gain competitive advantage in their product market competition through competing in R&D spending for process innovation!

We may also consider in more detail the effect of fixed costs on the profit-maximizing response of a firm. If the rival's output is low enough, the firm's fixed costs will not matter: it will still make a profit, and produce. However, at a higher output by the rival, the fixed cost will cause the firm to produce nothing, because the best it can do by producing leads to a loss. This is shown in Figure 5.13a, which is a modification of Figure 4.10. With the addition of an average cost curve, it is apparent that at output level Q_1 (the best it can do given the rival's output), the price the firm can charge is lower than its average cost at that output, so it makes a loss. Thus, at some high enough output level by the rival (which will leave the firm with a residual demand sufficiently below D^H), the firm leaves the market. This is shown in Figure 5.13b.



Rather than contemplate this fate of shutting down for your firm, let us suppose that you can have the same effect on your rival also, but that neither of you will choose to drive the other out of the market. In Figure 5.14a, a Nash Equilibrium is shown that still involves both firms producing for this market. Now suppose that you are clever enough to invest in R&D that shifts your profit-maximizing response curve out. As in the situation behind Figure 5.12a, you lower your marginal cost and this makes you more aggressive. In the new situation, shown in Figure 5.14b, your rival is the one that leaves the market. If we imagine the situation as one where you initially had the market to yourself, then your R&D spending keeps your *potential* rival out of your market. Without that spending, your rival would enter, and you would have to share the market. This example illustrates how industry structure (in the example, one firm versus two in the industry) can depend on the strategic weapons available to the firms in the industry, and how they use them. In the case of blocking entry, we can think of your spending on R&D as **pre-emptive spending**. This idea will be useful in thinking about e-commerce, as much as it is for traditional industries or markets.



5.8 Conclusion

This chapter has built on the ideas covered in Chapter 4 to give a richer picture of what firms are, what they do, and what industry structures can result. Underlying all the diagrams and symbols are a couple of basic ideas. The firm is a collection of assets and a sequence of value-creating activities. The term 'assets' in this case refers to human capital and intellectual property as much as, if not more than physical assets. The unique or special assets of the firm not only create value, but help the firm capture that value. An exceptionally successful firm will be one that controls the combination of special assets that allow it to create and capture exceptional value. In devising a strategy to do so, the firm must also anticipate and respond to what rivals and potential rivals can and will do, all along the value chain.

These general principles are just as important for e-commerce as they are elsewhere in business. E-commerce heightens the importance of reconfiguring value chains. This is obviously driven by the broader technological changes associated with the rise of information technology. Knowledge assets are thus important in their own right, and as a driver of changes in how other assets and activities are combined and organized. Much of this book is concerned with applying the principles of this chapter to understanding current and future developments in e-commerce.

Summary

- The value chain provides an overview of the sequence of value-creating activities undertaken by a firm, including production, logistics, marketing, research and development, human resource management and financial management. Value chains may extend beyond firm boundaries.

- The supply chain is the portion of a firm's value chain that involves supply activities, either by the firm itself, or supplier firms, including production and purchasing of inputs and inbound logistics.
- Economies of scale and scope are related types of decreasing average cost of output, that favor larger size in firms. Economies of scale can be extended to include dynamic economies, where cumulative output matters rather than just output per period. Economies of specialization and learning are a type of dynamic economies of scale.
- The choices between organizing activities through markets, or within firms are governed by the impacts on value creation and value capture of asset ownership, transactions costs associated with market contracts, and asymmetries of information that affect incentives and the methods of providing them.
- Location and distribution decisions are an important part of firms' strategies, especially at a time where the costs of transportation, or the informational benefits of physical closeness, are changing rapidly.
- Industry structures are determined by economies of scale and scope, as well as by firms' ability to control or dominate portions of the value chain, and the strategies they use to try to do so.
- The value chain approach highlights the many dimensions of firm strategy, including marketing and advertising, R&D, and finance, as well as the typical dimensions of production and pricing decisions.

Questions

1. How do you think the value chain for the airline industry differs from the value chain for the computer industry? In what ways are they similar?
2. The United States is now a service economy: over 70% of GNP is attributable to services (the rest is mostly manufacturing, while agriculture is a small percentage). Are service firms smaller than manufacturing firms? Why or why not? What factors create economies of scale in service provision?
3. How do you think the significance of various strategic dimensions varies across industries? In what kinds of industries is distribution likely to be important? Where is R&D most important? Marketing and advertising? Are some dimensions of strategy always present? Are some absent in particular industries?